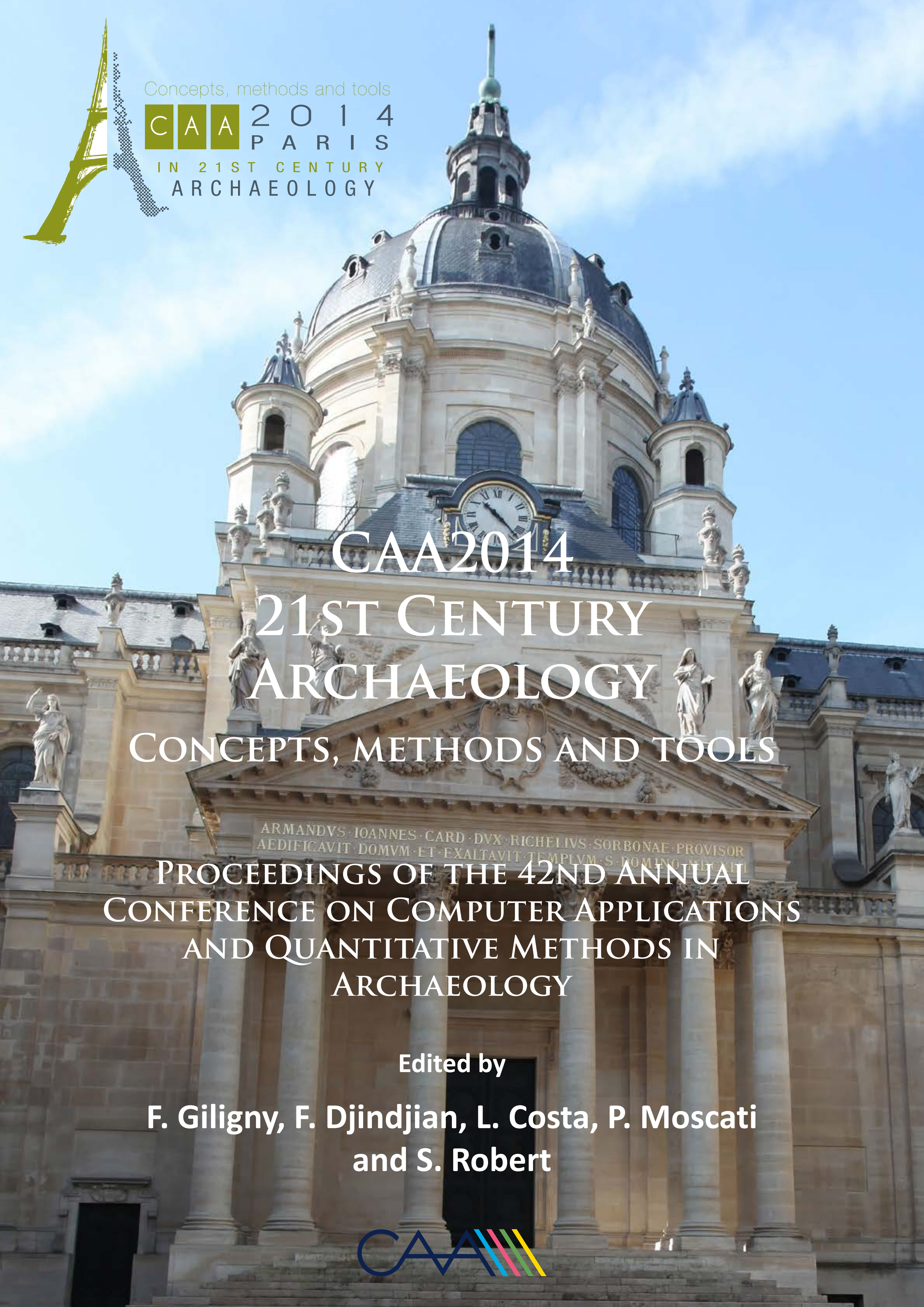
Concepts, methods and tools
CAA 2014 PARIS
IN 21ST CENTURY ARCHAEOLOGY



CAA2014
21st Century Archaeology

Concepts, methods and tools

Proceedings of the 42nd Annual Conference on Computer Applications and Quantitative Methods in Archaeology

Edited by

F. Giligny, F. Djindjian, L. Costa, P. Moscati and S. Robert

CAA

CAA2014
21st Century Archaeology

Concepts, methods and tools

Proceedings of the 42nd Annual Conference on Computer Applications and Quantitative Methods in Archaeology

Edited by

F. Giligny, F. Djindjian, L. Costa, P. Moscati and S. Robert

Archaeopress
Summertown Pavilion
18-24 Middle Way
Oxford OX2 7LG

www.archaeopress.com

ISBN 978 1 78491 100 3
ISBN 978 1 78491 101 0 (e-Pdf)

This book is available direct from Archaeopress or from our website www.archaeopress.com

UNIVERSITÉ PARIS 1
PANTHÉON SORBONNE

CNRS

L'ECOLE DES HAUTES ETUDES EN SCIENCES SOCIALES

Institut national de recherches archéologiques préventives
Inrap

MAIRIE DE PARIS

Trajectoires
De la sédentarisation à l'État

UMR 7041
ArScAn
Archéologies et sciences de l'Antiquité
CNRS - Univ. Paris X
Univ. Paris I - Ministère de la Culture

Contents

Chapter 3. Ontologies and Standards

Chapter 4. Internet and Archaeology

Chapter 5. Archaeological Information Systems

Chapter 6. GIS and Spatial Analysis

Foreword

This volume brings together a selection of papers proposed for the Proceedings of the 42th Computer Applications and Quantitative Methods in Archaeology conference (CAA), held in Paris (France) from 22nd to 25th April 2014.

The conference venue was Paris 1 Panthéon-Sorbonne University, in the main building next to the Panthéon. Workshops were held at the Institute of Art and Archaeology and the EHESS School. This was the first time in 42 years that the CAA had come to France, and we are proud to have hosted this important scientific event in Paris.

CAA2014 welcomed 477 participants from 39 countries. Altogether 397 papers were presented in 26 different sessions. The 5 round tables and 12 workshops also contributed to the success of the conference.

The program was divided into different themes and this structure has been maintained in the arrangement of articles in the various chapters of this book.

We are grateful to the following institutions which made the conference possible and supported it financially. Paris 1 Panthéon-Sorbonne University, the *Mairie* of Paris, the CNRS, the EHESS – Ecole des Hautes Etudes en Sciences Sociales, the INRAP – Institut national de Recherches Archéologiques Préventives, the research laboratories from the Maison de l'archéologie et de l'ethnologie, Nanterre – UMR Trajectoires & UMR Arscan. We would also like to thank the staff of the university and the student volunteers.

We hope that the congress participants, the contributors and all people interested in computing in archaeology will enjoy these proceedings.

Computers and Mathematics in Archaeology, Anatomy of an Ineluctable Success!

François DJINDJIAN

Université Paris1 Panthéon-Sorbonne & UMR 7041 Arscan, Paris, France
francois.djindjian@wanadoo.fr

Abstract

Over the last fifty years the use of computer science and mathematics in archaeology has undergone continuous development and to date it has become an almost indispensable tool at any stage of the archaeological procedure: documentation, planning and data recording during surveys and excavations, laboratory studies, landscape analysis, reconstruction of social systems, archiving, mediation within the scientific community and amongst a broader public. New technologies have revolutionised the discipline: archaeological information systems, data retrieval systems, geographic information systems, 3D, the Internet, multidimensional data analysis, mathematical modelling and multi-agent systems. But the most ambitious contribution lies in the field of the formalisation of a general theoretical framework of the discipline, the independence of which from any paradigm and ideology could raise archaeology to the level of the leading scientific disciplines.

Keywords: Computational archaeology, emergence, development, scientific context

1. Introduction

The increasing success over the last fifty years of scientific contributions related to computer applications and quantitative methods in archaeology may now be analysed from different technological and sociological points of view in order to understand the absolute relevance of such contributions to archaeology and the way in which the specialists of computational archaeology could play a major role in the future of modern archaeology.

2. The very beginnings of quantitative and computational archaeology

Several references mark the very beginnings of quantitative and computational archaeology which developed in the 1950s.

The book '*The Application of Quantitative Methods in Archaeology* ' edited by R. F. Heizer and S. F Cook (Heizer & Cook 1960) was the publication of the eponymous Wenner-Gren symposium organised at the Burg Wartenstein conference centre (Austria), from 1st to 9th July 1959, at which participants included J.D. Clarke, A. C Spaulding, A. C. Blanc, F. R. Matson, H. Vallois and W.W. Howells. The conference officialised the existence of quantitative archaeology, which had emerged about ten years prior, when Brainerd and Robinson published the first algorithm of seriation in 1951 and Spaulding the first statistical method for typometry in 1953.

In 1955 Jean-Claude Gardin, with the assistance of Henri Seyrig, then at the French Institute of Near Eastern Archaeology in Beirut, launched the first application of data retrieval systems in archaeology using mechanographic machines and later an IBM 650 computer. In 1957 he founded the CADA laboratory (*Centre d'Analyse Documentaire en Archéologie*) which set out the famous nine descriptive codes (1957-1969) marking the beginnings of semiotics and data banks in archaeology (Gardin 1956, 1976). From 20th to 30th June 1962 D. Hymes (Indiana University) organised a second Wenner-Gren symposium at the Burg Wartenstein conference centre entitled ' *The Use of Computers in Anthropology* ' in collaboration with P. Ihm, R. W Needham, J. Cl. Gardin, S. Ciccato, S. Lamb, Th. Sebeok (Hymes 1965).

3. Naming the discipline

The emergence of a new discipline requires a name. In our case a lot of names were used, following the evolution of computer sciences:

– *quantitative archaeology* (Heizer & Cook, 1960);
– *computer applications and quantitative methods in archaeology (*CAA);
– *data processing and mathematics applied to archaeology* (Djindjian and Ducasse 1982);
– archaeological computing (used as the title of the '*Archaeological computing Newsletter* ' (1984-2008) published by the Oxford University ; this name was used by the University of Southampton for its ' *Archaeological computing research group* '; and is also the name of the laboratory created by I. Johnson at the University of Sydney in 1992) ;
– *computational archaeology* (Wikipedia; seems to replace the previous name or to be specific to computerised mathematical modelling in archaeology);
– *digital archaeology/ digital antiquity* (used unfortunately by the web community for the history of the Internet instead of ' archaeology of the digital ', but also by the ' digital antiquity service tdar ' (digital archaeological record), by the center of digital heritage (University of York) and by DigAR Lab, the digital archaeology research lab (University of Washington, Seattle);
– *digital applications in archaeology and cultural heritage* (the name is used by the on-line peer review created by B. Frisher (DAACH);

– *archaeoinformatics* (name sometimes used by German speaking archaeologists).

The name ' *archaeological computing* ' seems to be the most popular although ' *computational archaeology* ' would be the most appropriate because there are equivalent uses in other sciences. The name ' *digital archaeology* ' or ' *digital antiquity* ' is a more recent coining, probably due to the success of the digital revolution in technology, by analogy with ' *digital heritage* ' and the current development of virtual reality and 3D applications.

4. Positioning the discipline

What is and what could be the academic position of our discipline ?

- A vertical specialty ?
The vertical organisation is the favourite one in the academic world, more particularly at university where the creation of a chair means that a discipline really exists! To my knowledge to date there is not a single chair in computational archaeology at any university in the world. Nonetheless M. Baxter was professor in statistical archaeology at Nottingham University (UK), J. Barcelo is associate professor in quantitative archaeology at the University of Barcelona (Spain) and I hold a chair of archaeological methods and theory at the Paris 1 Pantheon Sorbonne University. I suggest here that it is necessary to create a chair of archaeological methods and theory at all the universities hosting a department of archaeology, which would be in charge of archaeological skills training, and which should be headed by a specialist in computational archaeology.

Computational archaeology research laboratories (general purpose or devoted to GIS, virtual reality, archaeological data banks, etc.) are more frequent. A typical example is the laboratory led by J. Johnson at the Sydney University (see above) or the ' Archaeology Data Service ' at York University (UK) headed by J. Richards. The first was probably Irwin Scollar, head of the department of technical and computer methods in archaeology at the Rheinisches Landesmuseum in Bonn (Germany) from 1971 to 1991.

- A transversal way to improve archaeological knowledge?
Obviously the most effective way to develop computational archaeology applications would be to share projects and lectures with all the archaeological specialties which are most often structured according to geographic areas and temporal periods, whether at university or in research institutes. This implies some kind of matrix organisation which theoretically would be well adapted to the integration of a variety of disciplines related to different sciences (physics, geology, zoology, botanic, computing, mathematics, etc.). It is disappointing that this approach was unsuccessful in many conservative countries such as France, Italy and Spain, where the governance is in the hands of influential scholars (in their position or in the commissions they are members of) and does not result from a dialogue between all scientists that aims to define the global needs and the strategy of a department. Such an organisation, which is so efficient in the very dynamic and competitive world of industry and services, is unfortunately not possible in our academic world. As a consequence the progression of our discipline is slowed down by an institutional and sociological opposition, probably in the same way as for archaeometry.

5. Defining the professional skills and qualification as well as the academic position

Just as for any new scientific discipline, the pioneers of computational archaeology were scientists with a variety of rich backgrounds including hard sciences and human sciences. They trained many students who became the second and third generation of specialists of the discipline, for example John Willcock of the School of Computing at North Staffordshire Polytechnic or myself at the Paris 1 University.

With regard to the scientific production of our colleagues, it is possible to define different profiles of activities:

– Computer scientists or engineers who consider archaeology as a serious hobby. Some of them are very renowned specialists in their own discipline and their research in archaeology reaches a very high scientific level.
– Computer scientists who are interested in archaeology as a field of application for their new computing research. They may have the opportunity to obtain a grant from the European Union or from a national institution and they use archaeology as a display window for their advanced computerised projects. Current applications in digital heritage seem to belong to this category.
– Archaeologists who are involved in quantitative and computational techniques for their research (statistics, mathematical modelling, GIS, 3D, etc.).
– Archaeologists who are using modern archaeological methods (of course using computers).
– Archaeologists who are involved in the production of formalised (and then computerised) archaeological knowledge.

Specialists in computational archaeology occupy very different academic positions, sometimes depending on the sociology of archaeology in the different countries. Among them are:

– Archaeologists using modern techniques, methods and formalisation. In this case they may occupy major academic positions in archaeology, at universities or in research institutes. But they are first of all specialists in an archaeological period and a geographical area, as for instance F. R. Hodson (European protohistory), C. Renfrew, G. Cowgill (Mesoamerica), J. D. Clark (Neolithic), or myself (European Upper Palaeolithic) and many others!
– Technicians working for archaeologists who are not familiar with computational techniques.

A typical case is the French CNRS employing specialists in computational archaeology (for example specialists in archaeological data retrieval systems or geographic information systems) as engineers, technicians or administrative staff (ITA) and not as researchers. This means that the recruitment commissions of the CNRS (sections 31 and 32) reject all the applications with a computational archaeology profile. And as a consequence students do not choose to do a PHD in computational archaeology, because they know they will get a researcher job neither in the CNRS nor at universities (as was also the case for archaeometry). This situation seriously hinders the development of modern methods and techniques in French archaeology.

– A specialised laboratory as part of an archaeological institute

The decision to found a specialised laboratory is linked to major investment (equipment, salaries and office) necessary to its establishment and often results from a long struggle by the pioneer at the creation of the laboratory. By contrast to archaeometry laboratories (for example absolute dating or ceramology or geophysics), nowadays the creation of a laboratory of computational archaeology does not require a major material investment. It should even be stressed that such laboratories existed in France in the 1960s and the 1970s (CADA, LISH) in human and social sciences, but were finally closed. The context of the recent creation of a virtual reality laboratory in Bordeaux by R. Vergnieux (Archéovision) after a long struggle with the Ausonius archaeological institute is symptomatic of these difficulties.

However, multidisciplinary studies carried out in joint projects together with archaeologists are a prerequisite for these laboratories to guarantee their success. This seems to be a major difficulty in countries where the organisation of archaeology is very conservative. Otherwise they are forced to work separately and independently, as was the case for the department of use wear analysis in Saint-Petersburg created by Y. Semenov in the 1950s and it seems to be the case currently with the laboratories of paleogenetics.

-An institution providing specialised services
Examples of private laboratories are rare. Their existence depends on the attitude of governments applying liberal politics in archaeology as for example in rescue archaeology. Start-up companies have been created for 3D processing in archaeology and cultural heritage.

A major topic is the scientific background of students in archaeology. Archaeology is increasingly taught in Humanities, although it is a discipline that uses methods and techniques related to physics, mathematics, computing, earth sciences and others. Such a situation may depress the scientific level of the future generations of archaeologists.

Be that as it may, an archaeologist using computational archaeology will be considered as a statistician or a computer scientist! That is not news! Remember that A. Glory (a priest studying palaeolithic art) described A. Leroi-Gourhan as a technician in mechanics, because he used punched cards (and knitting needles) to study European Palaeolithic art! So don't worry if your colleagues say you are a system engineer because you are using Powerpoint! Or a driver if you have a driving licence! But this is not a sound reason not to encourage our students to become skilled archaeologists!

6. Academic reviews devoted to computational archaeology

The existence of academic reviews devoted to computational archaeology confirms its status as a scientific discipline. Several of them have disappeared, others still exist:

1965-1976 *Newsletter of Computer Archaeology* (Arizona State University)

1967 *Computers and the Humanities*

1982-1992 *Archéologie et ordinateurs* (CRA CNRS)

1984-2008 *Archaeological computing Newsletter* (Oxford University)

1990 *Archeologia E Calcolatori* (CNR)

Archeologia E Calcolatori has published one volume per year and special issues since 1990.

7. Conferences on computational Archaeology

As mentioned above, the first conferences on computational archaeology were held at the end of the 1950s, notably the Wenner-Gren symposia.

In 1970, the conference of Mamaia (Romania) 'Mathematics in the Archaeological and Historical Sciences' was the place where the most renowned statisticians met archaeology: Rao, Kruskal, Kendall, Sibson, La Vega, Lerman, Wilkinson, Solomon, Doran, Ihm, Borillo, Gower. And where archaeologists demonstrated that they were skilled in the use of statistics: Moberg, Spaulding, Cavalli-Sforza, Hodson, Orton, Hesse, Ammerman, Goldmann.

This conference probably influenced the organisation of the first CAA conference (Computer applications and quantitative methods in archaeology) in Birmingham in 1973. This was the beginning of an annual conference cycle, held for the first time in Birmingham, then elsewhere in the UK, after 1992 in Europe and from 2006 outside Europe. It describes itself as: '*CAA is an international organization bringing together archaeologists, mathematicians and computer scientists. Its aims are to encourage communication between these disciplines, to provide a survey of present work in the field and to stimulate discussion and future progress*'.

During its 1976 world congress in Nice (France) the IUPPS (International Union of Prehistoric and Protohistoric Sciences) decided to create scientific commissions for their inter-congress activities. Commission 4 '*Data* *management and mathematical methods in Archaeology*' was then launched. Over the last 35 years commission 4 organised one or more sessions at each congress every four years as well as inter-congress conferences (Amsterdam, Denver, Sydney, Scottsdale, Paris, Leiden, etc.). At the 2006 conference in Lisbon commission 4 decided to change its name to '*Archaeological methods and theory :formalization, quantification, mathematics and computerization*'.

National conferences on computational archaeology were held, in particular the ' *Workshop on Archaeology and Computers* ' in Vienna (Austria), organised each year since 1995 by W. Börner and the ' *Journées d'Informatique et Archéologie de Paris* ' (JIAP) in Paris held biannually since 2008.

CAA also has several national branches organising additional conferences.

Dedicated sessions were also organised at national (for example SAA in the USA) or international archaeological congresses (for example WAC or EAA) and it can be assumed that there is currently no archaeological congress without computational archaeology sessions.

The rapid development of virtual archaeology, for which cultural heritage is a more gratifying market of application than archaeology, prompted the creation of subdivisions in computational archaeology. The 1998 CAA conference in Barcelona gave rise to a separate publication (Barcelo, Forte and Sanders 2000). In 2000, the first VAST conference (International Symposium on Virtual Reality, Archaeology and Cultural Heritage) was organised in Arezzo (Italy) by F. Niccolucci and the 14th international VAST symposium was held in Pistoia (Italy). The biannual 'Virtual Retrospect ' conference has been held in Bordeaux (France) since 2003.

Virtual reality (VR) is a very dynamic field of research and its application to cultural heritage is very spectacular. For example, in 2009 as many as five VR conferences on archaeology were held: in Malta, Trento, Paestum, Bordeaux, Seville (' i*nternational charter of virtual archaeology* ') as well as over a dozen on general virtual reality! In 2013, all these events converged to become the ' *Digital Heritage* ' conference in Marseille (France).

The very rapid development of 3D archaeology, which is not only concerned with cultural heritage, may possibly revolutionise field archaeology as well as all data processing that takes place following excavations and surveys, and it could create the conditions necessary for the creation of further subdivisions in computational archaeology.

8. What techniques are used in computational archaeology?

Computational archaeology uses almost all the techniques and tools related to applied mathematics:

- Applied mathematics
- Sampling
- Statistics including graphics, elementary statistics, statistical tests, Bayesian statistics,
- Multidimensional data analysis
- Algorithms
- Graph theory
- Mathematical modelling
- Signal processing
- Image processing
- Multi-agent systems
- Etc.

Computational archaeology also uses almost all the computing applications:

- Semiotics
- Data Retrieval Systems
- Ontologies
- Data Base Management System,
- Archaeological Information System
- Geographic Information Systems
- Virtual Reality
- 3D recording and processing
- Publishing
- CAD (computer-aided design) / CAM (computer-aided manufacturing)
- Internet X.0
- Office automation
- Etc.

All these techniques are of course software programs that the archaeologist has downloaded onto his/her personal computer or workstation, which is powerful enough to run almost all of these applications. The long list reveals the great effort that archaeologists have to put into training in order to be able to use all the software tools. It confirms the need for courses in computing applications very early in academic archaeological training.

9. What archaeological methods use computing tools?

I have frequently pointed out the prime importance of separating the level of techniques (which are steadily improved and often replaced) from the level of methods or best practices (which are only optimised). The techniques are the field of competence of computer scientists or statisticians whereas the methods are the field of competence of archaeologists. For example, it is fundamental to distinguish sample dating (involving a laboratory's competence in radiocarbon dating) from site dating (involving an archaeologist's competence).

The following list of items, which are the chapters of an archaeological textbook, shows the major role of computational archaeology in the process of archaeological investigation (Djindjian 2011):

- Archaeological BPM (business process management)
- Survey

– Excavations
– Stratigraphy
– Sampling for Rescue archaeology
– Typology
– Stylistic analysis
– Seriation
– Culture historical identification
– Intra-site spatial analysis
– Metropolitan spatial analysis
– Archaeology of the construction
– Raw material and manufacturing characterization
– Landscape analysis
– Image processing for special archaeological purposes
– Environmental studies
– Exchange systems
– Virtual Archaeology
– Demography
– Food resource management system
– Transition models
– Collapse models
– Economic models
– Network analysis
– Complex system reconstitution
– Culture Resource Management (CRM)
– Epigraphy
– Etc. !

10. Towards a generalised cognitive framework?

The books ‘ Analytical Archaeology ‘ published by J. D. Clarke (Clarke 1968) and ‘ Archaeological constructs ‘ by J.-C. Gardin (Gardin 1979) reveal that the formalisation of an archaeological construct is also the aim of computational archaeology, just as in the case of artificial intelligence (AI).

Over the last fifty years of archaeology it is possible to distinguish two different approaches with regard to archaeological constructs:

– The constructs are embedded, explicitly or not, in an ideology, a paradigm or a theory (in the Anglo-Saxon sense). In fact, they may be considered as operating an a priori reduction of the range of possibilities: evolutionism, Marxism, Neo-Marxism, functionalism, cultural ecology, gender theory, substantivism, Marrism, Kossinnism, behavioural archaeology, evolutionary archaeology, symbolist archaeology, etc.

Such a reduction may only be useful if it is used not as a dogma (which generally defines an ideology) but rather as a driving force to explore the explanations which were deduced.

– The constructs must be formalised, backed by a cognitive framework (in other words, an epistemology), which is nothing other than a theory of knowledge. An example of this kind of approach based on Peirce’s logic is my paper ‘*Pour une théorie générale de la connaissance en archeologie*’ (Djindjian 2002).

In such a context the traditional opposition between processual archaeology (New Archaeology) of the 1960s and post-processual archaeology (or symbolist archaeology or post-modern archaeology) of the late 1980s, which is also a classical opposition between structuralism and hermeneutics, appears to be obsolete in comparison with the progress registered in cognitive sciences.

It is important to point out that specialists in computational archaeology have to be considered as being the best contributors to the renewal of the theoretical framework of archaeology.

An example of such a contribution is the topic of the complementarity between the data-oriented approach and process-oriented approach in archaeology. In computer sciences there is a classic complementarity between the data-oriented approach (date bases and data storage), the computation approach (algorithmic) and the process-oriented approach (real time).

Archaeology of the 19th and 20th centuries was mainly a ‘data-oriented archaeology ‘ following the Montelius typology and the corpus programs, renewed in the 1970s by data retrieval systems (archaeological data banks) and more recently by the Internet, making up the ‘ back office ‘ of archaeology.

The process-oriented approach concerns first the organisation of the archaeological activity (APM or archaeological process management), second the study of the reliability and the representativeness of the archaeological record (taphonomy) and third the systemic reconstruction of past societies (Djindjian, 2014).

11. A charter for ‘ 21st Century Archaeology’

The goal of computational archaeology is to act as a driving force for the creation of a present archaeology, able to generate improved knowledge and to reconstruct more reliable systems. The following charter of recommendations may illustrate such a programme:

– Relaunch thesaurus and ontology projects in all archaeological fields!
– Create multimedia data bases to save millions of archaeological slides and drawings!
– Write International archaeological standards!
– Start a 3D revolution in Archaeology!
– Think process and systems and renew all archaeological issues!
– Take part in multi-agent system simulations, processes of governance, social organisation, societal behaviours, beliefs, etc.!
– Ignore ideologies, paradigms, theories and the like!

– Carry out real multidisciplinary studies: be inspired by new computerised techniques and integrate but not blindly apply them because of the unreliability of the archaeological record which makes things more complex!
– Ensure high-level academic training for new generations of archaeologists!
– Create specialised laboratories in research institutes!
– Create chairs of ‘ Archaeological Methods and Theory‘!
– Etc.

12. Conclusions : the 21st century archaeologist

The 21st century archaeologists are no longer excavators: they integrate other disciplines such as history, epigraphy, geography, anthropology, ethnology, economy, agronomy, physics, chemistry, mathematics, computer science, etc. They shoulder the very difficult task of reconstructing the complex systems of past societies based on partial, biased and meaningless archaeological and epigraphic data. They also play a role in society: knowledge of the past enables them to understand the present and to anticipate the future. They are the only scientists to conceive the depth of the time. Computational archaeology may play a major role in such a challenge, but it needs a proper name, a proper scientific review and a proper conference, well accepted and open to all specialists.

References

BARCELO, J.A. FORTE, M. SANDERS, DH. (2000). *Virtual reality in Archaeology*, (BAR International series, 843), Oxford: Archeopress.

BRAINERD, GW. (1951). ‘The Place of chronological ordering in archaeological analysis’, *American Antiquity* 15, pp.293-301

CLARKE, D. L (1968). *Analytical Archaeology,* Methuen, London.

DJINDJIAN, F. (2002). Pour une théorie générale de la connaissance en archéologie. XIV° Congrès International UISPP, Liège Septembre 2001. Colloque 1.3. *Archeologia E Calcolatori* 13, pp.101-117

DJINDJIAN, F. (2011). *Manuel d’Archéologie*. Paris: Armand Colin.

DJINDJIAN, F. (2014). Simuler une ‘Artificial Society’: organisation sociale, gouvernance et attitudes sociétales. *Actes des troisièmes journées Informatique et Archéologie de Paris - JIAP 2012*, Paris, Juin 2012. *Archeologia e Calcolatori*, Supplément 5, 2014, pp.18-33.

DJINDJIAN, F. DUCASSE, H. (edts) (1987). *Data processing and mathematics applied to archaeology. Mathématiques et Informatique appliquées à l’Archéologie*, PACT 16, Conseil de l’Europe.

GARDIN, JCl. (1976). *Code pour l’analyse des formes de poteries, établi en 1956, révisé en 1974*. Paris: Éditions du C.N.R.S.

GARDIN, JCl. (1979). *Une Archéologie théorique*, Hachette, Paris

HEIZER, RF. COOK, SF (eds.) (1960). *The Application of Quantitative Methods in Archaeology*. Chicago: Quadrangle books.

HODSON, F. R., KENDALL, D. G., TAUTU, P. (eds.) (1971). *Mathematics in the archaeological and historical sciences. proceedings of the Anglo-Romanian conference, Mamaia 1970* / ed. F. R. Hodson, D. G. Kendall, P. Tautu. Edinburgh : Edinburgh University Press

HYMES, D. (ed.) (1965). *The use of computers in Anthropology*. Studies in general Anthropology, 2, London: Mouton.

ROBINSON, NS. (1951). ‘A Method for chronologically ordering archaeological deposits’, *American Anthropologist* 15, pp.301-313

Chapter 1 Historiography

Towards a History of Archaeological Computing: An Introduction

Paola MOSCATI

CNR – Istituto di Studi sul Mediterraneo Antico, Italy
paola.moscati@isma.cnr.it

Abstract
The international research project on the history of archaeological computing, conducted in cooperation with the Accademia Nazionale dei Lincei and the Italian National Research Council, was designed to investigate some methodological aspects associated with the introduction and development of computer applications in archaeology. The project is subdivided into two parts, one based on a systematic data collection, mostly carried out in Italian and foreign Archives and Libraries, and the other addressing the construction of a communication tool, i.e. the Virtual Museum of Archaeological Computing. By retracing the unfolding of international scientific events occurring in the first two decades of the second half of the 20th century and the lives of the scholars who promoted them – in particular Jean-Claude Gardin – the paper illustrates the methodological renewal process, which gave birth to the convergence of the so-called 'two cultures', opening new directions in archaeological research thanks to the integration of objectives and tools inherited from different disciplines.

Keywords: Archaeology, Archaeological computing, History of computer applications, International events, Jean-Claude Gardin

1. Introduction

In recent years, several international events have provided the opportunity to reflect upon the history of archaeological computing as an autonomous discipline and to reassess its theoretical basis by reconsidering the earliest applications. A key clue of this historiographical approach can be found in the exponential growth of technology applied to archaeological documentation procedures, which at present strongly characterise this discipline, overshadowing in many cases past theoretical and methodological advances.

We can recall various attempts made in this direction. First, during the international Symposium held in Rome at the Accademia Nazionale dei Lincei in 2008 that specifically addressed 'The Birth of Archaeological Computing' and whose Proceedings were published in the 20th issue of the Journal 'Archeologia e Calcolatori' (Moscati, 2009; http://soi.cnr.it/archcalc/indice/PDF20/AC20.pdf). Secondly, four years later, during the 2012 CAA Conference, with the session '40 Years of Computer Applications and Quantitative Methods in Archaeology' now published online in the website of the Personal-Histories Project of the Cambridge University (http://www.sms.cam.ac.uk/media/1357554). We can also mention, in the framework of the 2014 UISPP Burgos Congress, the session focused on 'The revolution of the Sixties in Prehistory and Protohistory', organised by Colin Renfrew, François Djindjian and Alessandro Guidi (http://www.burgos2014uispp.com/).

2. Towards a history of archaeological computing: data collection and dissemination

Given the subject of this session, it is deemed appropriate to briefly introduce a project focused on the 'The History of Archaeological Computing', which is currently in progress at the Accademia Nazionale dei Lincei. The Centro Linceo Interdisciplinare 'Beniamino Segre' and the Italian National Research Council are involved in this international initiative, in cooperation with the Italian Department for Education, University and Research and the Department for Cultural Heritage (see, lastly, Moscati, 2014).

The Centro Linceo (http://www.lincei.it/modules.php?name=Centro_Linceo) was established under the initiative of the Italian mathematician Beniamino Segre in 1971, with the aim of encouraging scientific interdisciplinary research in various fields of Sciences as well as in the Humanities. Over the years, it has promoted many important projects, seminars and publications, which today can testify to the development of computers and IT applications in the Humanities in general and in archaeology in particular. The Istituto di Studi sul Mediterraneo Antico of the Italian CNR (http://www.isma.cnr.it/), established in 2013 but with a long tradition of studies in new and less developed archaeological sectors of investigation, specialised in digital archaeology over a period of more than thirty years. This early commitment has resulted in the regular annual publication, starting in 1990, of the international scholarly Journal 'Archeologia e Calcolatori' (http://soi.cnr.it/archcalc/), which is now listed in the Directory of Open Access Journals (DOAJ).

The idea for this interdisciplinary project began to take shape during the above-mentioned Lincei Symposium. In the workshop section, the prototype of a Virtual Museum was presented to the public, with particular reference to the Protagonists and Methodologies. At the present stage, a new website is under construction. It is still a work in progress that will constitute the communication mirror of the historical research project (http://archaeologicalcomputing.lincei.it/). An open-source Content Management System has been chosen to ensure its implementation and maintenance: Museo&Web (http://www.otebac.it/index.php?it/205/prototipo-museoweb/). Promoted as part of the European Minerva Project, whereby a core set of guidelines on cultural websites quality were issued, Museo&Web is a best practice, produced and funded by the Italian Department for Cultural Heritage especially for the benefit of cultural institutions, such as museums, libraries, and archives (Natale, Saccoccio, 2010).

In the home page of the Virtual Museum of Archaeological Computing, the female figure of Penelope oversees and accompanies visitors, as she did in the presentation brochure of the Centre d'Analyse Documentaire pour l'Archéologie, edited by J.-C. Gardin at the end of the Fifties (*Centre d'Analyse*, 1959). The banner presents the following links: History, Protagonists, Institutions, Projects, Events, Methods, Techniques, and Documents. The History section is subdivided into decades, while Protagonists include Precursors, Pioneers, and Promoters. Institutions, Projects and Events will document the main evolutionary phases, through specific case studies, characterised by their duration and continuity in scientific endeavours.

Some cultural itineraries have also been planned, according to a different architectural criterion: some scholars have been invited to deeply investigate some specific topics, and their reviews will be made available through interactive web pages. Being aware of the upcoming open science era, these pages will also contain the comprehensive multimedia documentation gathered in the open access digital repository of 'Archeologia e Calcolatori' during 25 years of publishing activity, together with the valuable documents produced on the occasion of the international scientific events organised and promoted by the Accademia dei Lincei. In this way, scholars will be able to share experiences, to reconstruct the establishment of institutions and/or of ad hoc research laboratories dedicated to developing archaeological computing, and to propose innovative research routes.

During the phase of data collection – more specifically the long and fruitful archival research carried out mostly in Italy, France and the UK (Moscati, 2013)[1] – a large body of documents centred on the earliest applications of computers in archaeology led us to focus primarily on the first two decades: the Fifties and the Sixties, when all foundations for future developments were skilfully laid and the New Archaeology movement was born. However, even after restricting the chronological range, we feel here compelled to further narrow down the field of investigation in order to better show how complex and ramified the beginnings of what we now call 'digital archaeology' were.

Past and present applications, the global circulation of information, the frequency of conferences and seminars, the dynamic library formed by manuals and specialised journals, all together undoubtedly give today a strongly positive image of the activity that has been taking place in archaeological computing. Therefore, even confining the analysis to the first twenty years of research, any descriptive attempt would be quite a feat. For these reasons, I will focus on a single subject: the universe of international conferences held in the first two decades of the second half of the Twentieth century, a truly virtuous period in terms of cultural ferment.

In those years, the scientific character of all the historical disciplines was reinforced. In archaeology, this gave rise to the intensive use of an expanding armoury of scientific techniques and instruments, from the statistical treatment of data to the application of archaeometric analyses and geophysical prospection techniques, from information retrieval systems to the setting up of data banks, from aerial photography to remote sensing, etc. Such quantitative innovative approach provided archaeological research with the possibility of building up an area of 'objective' data, which it had never had in the past, and of experimenting with computation, simulation, and graphic analysis techniques.

3. A timeline of international events

An interactive timeline seemed to be the best tool to introduce a virtual presentation of international conferences. For this reason, we chose a new software tool, strictly related, in its internal structure, to the Museo&Web implementation: the CMS MOVIO (acronym of Virtual Online Exhibitions). MOVIO (http://www.movio.beniculturali.it/) is an open source kit intended for both public and private cultural institutions that are implementing strategies to promote and enhance knowledge through temporary exhibitions or web applications specifically designed to cultural heritage.

By retracing the unfolding of international events, it is also possible to learn about the history of pioneering scholars, who were individually involved as full participants and as such became promoters of innovative methods of research within a broader movement of ideas that helped to change the theoretical and methodological attitude of archaeological scholarship.

Therefore, starting at the end of the Fifties, we can identify several conferences intended to both data automation and the application of quantitative methods in the Humanities and Social Sciences. In general, we can clearly distinguish two separate types of phenomena: first, the distinction between the automatic processing of scientific information

[1] In this paper, we will mostly refer to the documents consulted in Nanterre at the Archives of the Maison Archéologie & Ethnologie, René-Ginouvès ('Fonds Équipe Archéologie de l'Asie Centrale et Jean-Claude Gardin'; from now on JCG), and in Paris at the Archives of the Bibliothèque Claude-Lévi-Strauss, Laboratoire d'Anthropologie Sociale (Archives du LAS (1960-1982). Direction Lévi-Strauss, Dossier Jean-Claude Gardin, fr/cdf/las/FLAS.D.04.05). Significant data originate also from the documents and interviews provided by James Doran and Roy Hodson, the authors of the first textbook on archaeological computing (Doran, Hodson 1975).

and the application of statistical techniques. This basic distinction – that Jean-Claude Gardin, by giving birth to a specific research unit of the CNRS (URADCA) directed by Mario Borillo, so masterfully defined as the *analyse documentaire* and the *calcul*, respectively – allows us today to provide a comprehensive overview of the achievements made in those days. With specific reference to archaeology, a third field of applications should be added, more oriented towards the processing of data coming from systematic geophysical prospecting on archaeological sites and from chemical-physical analyses designed to detect artefacts composition and chronology.

Secondly, the boundaries between all the disciplines that characterise the Humanities and Social Sciences appeared still blurred in their relationship with new technologies. Literature, linguistics, psychology, history, anthropology, archaeology, all made progress together towards new solutions offered by the application of quantitative methods and computer science. This is why we will focus our attention not so much on archaeological international conferences, in which we can sporadically find the earliest contributions illustrating isolated research projects, but primarily on events centred on issues pivoting around new research methods.

3.1. Automatic documentation systems

By following the first path of our itinerary, i.e. that of the development of automatic documentation systems, the starting point can be found in some international events organised in response to the growing interest in the dissemination of information for the sake of the scientific community. At the end of the Forties, when the Royal Society convened the 'Scientific Information Conference', some basic principles were stated. In a globalising perspective, aimed at sorting, storing, and searching scientific information, one of the most important assertions made was that dissemination required orderly procedures and systems that will make scientific communication widely available, raise awareness of information accessibility, and provide prompt access to any desired information.

Some conferences followed, but we have to wait until 1958 before we can register the presence of an archaeologist attending these meetings: Jean-Claude Gardin. During the 'International Conference on Scientific Information', organised in Washington D.C. by the National Science Foundation, Gardin presented a paper in which he explained the methods of analysis of archaeological data with which he was experimenting at the Centre mécanographique de documentation archéologique in Paris (Gardin, 1959). The following year, Gardin participated also in the 'International Conference on Information Processing', organised in Paris by UNESCO. The organisation was very impressive and only a part of the contributions were published (*Information Processing*, 1960). The subject presented by Gardin, together with Paul Braffort, within a specific Symposium on the Use of Automatic Computation in the Social Sciences, focused on the 'Determination and Mathematical Development of the Structures in Human Phenomena: Anthropological and Mathematical Angle' (a typewritten document can be downloaded from the data bank UNESDOC: http://unesdoc.unesco.org/images/0015/001574/157450eb.pdf).

In this period the link between Gardin – who was then developing the Syntol documentation language in cooperation with the EURATOM Laboratory at Ispra, Italy – and UNESCO intensified. After his participation in 1960 in the 'Colloque sur la coopération internationale en matière de documentation et d'information scientifiques et techniques', organised in Milan by the Centre français d'échanges et de documentation techniques, under the auspices of the Italian National Research Council and the Federation of Scientific and Technical Associations, in Moscow Gardin was one of the coordinators of the Working Party on 'Automatic Documentation: Storage and Retrieval' (http://unesdoc.unesco.org/images/0014/001491/149107eb.pdf).

The Russian cultural activity in the research area of automatic documentation, and more strictly the involvement of the Russian Academy of Science Archaeology Institute in the application of exact sciences in archaeology, are also reported by Gardin during the Steering Committee Meeting of the Centre d'Analyse Documentaire pour l'Archéologie (CADA), which in 1964 had been moved from Paris to Marseille, following the decentralising policy of the French CNRS (JCG 2/1, 6 January 1966). In the Report section illustrating foreign initiatives and the exchanges entertained with CADA, some Moscow meetings and the recent works by V.A. Ustinov (1964) and B.A. Kolčin (1965) are mentioned. The following year, a series of Gardin's conferences in Moscow, Leningrad, Kiev and Novosibirsk is also reported, to disseminate and make public the methods of analysis and automatic documentation applied to scientific literature in general, and archaeological objects in particular (JCG 2/1, 6 May 1967).

Coming back to the progress of international events, in the Archives of the Bibliothèque Claude-Lévi-Strauss, in a typewritten document (Dossier Jean-Claude Gardin, *Propositions pour une politique de l'organisation documentaire - Sciences Humaines*), Gardin, quoting two other international Conferences held in 1959 in Frankfurt ('Automatische Dokumentation in der Praxis/Automatic Documentation in Action') and Cleveland ('Standards on a Common Language for Machine Searching and Translation'), started out by saying that 'Les problèmes de l'information scientifique sont à l'honneur'. He also added that the debate was still opposing 'advocates' and 'opponents' on the use of machines; in fact, most of the documentation studies were based on the assumption, made by the former and rejected by the latter, that the huge amount of documents that should be handled to establish scientific bibliographies could not be mastered without the aid of automatic processing.

This conflict did not last long, and was obviously resolved in favour of the advocates of the use of machines. Since the early Sixties the interest for new documentation systems spread exponentially and Gardin himself, in the above-mentioned CADA Report, defined 1965 as the year characterised by a great proliferation of conferences on the use of computers in the Humanities. In fact, in the USA, thanks mostly to the financial support of IBM that was also sponsoring the newborn journal 'Computers and the Humanities', a series of conferences was held in six different universities: Rutgers University, Yale University, University of California, Georgetown University, Purdue University, and Boston University.

3.2. Quantitative methods

Meanwhile, quantitative methods continued to flourish together with their applications to archaeological data with obvious consequences in the context of international meetings (out of the many historical reviews, see in particular Wilcock, 1999 and, more recently, Djindjian, 2009). A very important event, at least in the premises contained in its title, was the Symposium on 'The Application of Quantitative Methods in Archaeology' held in Burg Wartenstein Castle in 1959 and organised by the Wenner-Gren Foundation (Heizer, Cook, 1960). Albert Spaulding's paper focused on 'Statistical Description and Comparison of Artifact Assemblages'. This paper, which follows the most famous article published in 1953 in 'American Antiquity' on 'Statistical Techniques for the Discovery of Artifact Types' (Spaulding, 1953), is generally referred to as one of the first methodological assessments of the role of statistical techniques applied to archaeology. As a matter of fact, and as recognised also by Gardin (Gardin, 1971, 189), Spaulding can be considered one of the most convinced promoters of the intensive use of statistical techniques in archaeological analysis. Moreover, by suggesting that statistics offered not only a collection of methods, but also a way to represent and express the ideas of archaeologists more accurately, Spaulding set forth a line of thought that would characterise the New Archaeology movement, which emphasised the need for rationalisation and computation in order to make archaeological assumptions increasingly formal and explicit.

The Wenner-Gren Foundation was very active in those years, and in 1962 a new Symposium was organised in Burg Wartenstein with the title 'The Use of Computers in Anthropology', whose proceedings were later edited by Dell Hymes (Hymes, 1965). During the Conference, that officialised the entrance of computers in anthropological and archaeological studies, besides presenting his own research experience on reconstructing an economic network in the Near East, Gardin was asked to make a typological summary of computer uses in anthropology.

His overview is a very lucid one. Between two more general criteria of presentation, one based on illustrating the various branches of anthropology and the relative computer applications, and the other on classifying the applications according to the methods of data processing – a dichotomous approach that will long be adopted in the reviews of the years to come – Gardin chose a third way, taking into account both data typology and the nature of computer operations. The result is a classification based on three different types of data (natural languages, special codes and physical events) which, intertwined with two main groups of operations (compilation and systematization) led to nine different application fields corresponding 'to an increasing degree of complexity and ambition in the kind of data processing which is entrusted to the machine, to extend our knowledge of man' (Gardin, 1965, 117). In effect, they proceed from pure retrieval operations and automatic documentation procedures to more complex automatic classifications as a basis for elaborate theoretical constructions, i.e. from a static (description of a system) to a dynamic (behaviour of a system) systematisation.

As part of the session devoted to special research areas, the theme of classification and grouping was addressed, with particular reference to automatic classification, which in those days was gaining ground (see in particular Ihm, 1965, who assembled the principal methods under four headings: automatic grouping analysis in metric spaces; factor analytical methods; clustering by use of a density function; maximum likelihood method). Thanks to the documents that James Doran kindly provided us, we were able to retrace the history of the Classification Society, which this year has celebrated its 50th Anniversary. The Society's foundation in 1964 arose out of a symposium of ASLIB (Association of Special Libraries and Information Bureaux) on 'Classification: An Interdisciplinary Problem', organised in London in 1962. The first meeting convened by the Society – which was intended to promote co-operation and interchange of views between experts interested in the principles and practice of classification in a wide range of disciplines – was held at the Institute of Archaeology, London University, in 1965 and registered the presence, among others, of Roy Hodson. Four years later, in 1969, at King's College, Cambridge, the same Society organised a meeting in which David Clarke discussed problems of 'Archaeological Taxonomy'. The Society gave also birth in 1966 to the 'Classification Society Bulletin', editorially coordinated by the microbiologist Peter Sneath, who had just co-authored the famous book on numerical taxonomy (Sokal, Sneath, 1963).

Some of Doran's notes also deal with the preparation and implementation in 1967 of the Seminar on 'Statistical Methods in Archaeology', organised by the Biometric Society and introduced by David Kendall, who at the beginning of 1962 had moved from Oxford to Cambridge as Professor of Mathematical Statistics and Director of the Statistical Laboratory, in which many experts in statistical applications to archaeology, like e.g. Clive Orton and Roy Hodson, were to be trained (for 1967 group picture see http://www.statslab.cam.ac.uk/Dept/Photos/pic67.html).

3.3. Archaeology and computers

The first international Conference in which a specific section, coordinated once again by Gardin, was fully dedicated to archaeology was the one presented during the 'International Symposium on Mathematical and Computational Methods in the Social Sciences', organised in July 1966 by the International Computation Centre (ICC), an intergovernmental centre, headquartered in Rome and established under the aegis of UNESCO. The Conference was subdivided into four main sections, focused on anthropology, archaeology, psychology, and sociology. The Proceedings were published two years later with the title 'Calcul et formalisation dans les sciences de l'homme', inaugurating the crucial editorial activity of the CNRS for the promotion of events debating computer applications in the Humanities (Gardin, Jaulin, 1968).

Anecdotally, we can recall that Roy Hodson, during the home interview he generously granted us, describes the conference in Rome as one of the first meetings on archaeological computing, in which he met many interesting people, but academically, as in most of these early conferences, he noticed two completely separated sets of people: 'very able mathematicians who could understand the niceties of Principal Components Analysis, but could not understand what archaeologists were trying to do, and archaeologists who could not understand anything about the mathematics and could not present the data in a form that a mathematician could use'. James Doran, in turn, still has some letters exchanged with Gardin as he was unable to attend the conference, but was very interested in deepening their scientific exchange.

In the Proceedings, the archaeology section – mostly dedicated to the use of mathematical techniques to automate classification issues – was represented by five scholars: Robert Chenhall, about whom we will discuss later on; Vadime Elisseeff, who at the end of the Fifties had supplied the collection of Bronze Age Eurasian axes analysed by Gardin and Ihm; Juliette de la Genière and W. Fernandez de la Vega, who approached the classification of the grave goods found in the excavations of the Sala Consilina necropolis, in Southern Italy; and Bohumil Soudský, a scholar who can rightfully be ranked among the pioneers of computer applications in archaeology. Among the participants, there were also Irwin Scollar and Richard Linington, who were experimenting with new methods of archaeological prospecting.

Before moving to the first international event actually discussing the use of computers in archaeology, I would open a brief parenthesis on the Symposium held in New York in 1968 and entitled 'Computers and their Potential Applications in Museums'. Organised by the Museum Computer Network and the Metropolitan Museum of Art, with funding from IBM, the conference dealt with another sector of applications that, in the years to come, would have gained considerable success (*Computers and their Potential Applications in Museums*, 1968). Among the speakers, in addition to Gardin, the presence of Robert Chenhall was very significant, due to his direct involvement in computer applications since the early Sixties. He was also responsible for editing the 'Newsletter of Computer Archaeology', the first one on this subject, published by the Department of Anthropology of Arizona State University, and for promoting the use of data banks in archaeology (Chenhall, 1971).

Finally, in 1969, in Marseille, the time was ripe to extract archaeology from the other Humanities disciplines and measure it against the advent of computers. From the 7th to the 12th of April 1969 Jean-Claude Gardin organised the 'Colloque international sur l'emploi des calculateurs en archéologie: problèmes sémiologiques et mathématiques' at the CADA headquarters. In the presentation brochure of the Symposium, Gardin grouped the subjects to be discussed under two headings: A) Symbolic problems raised by the descriptive analysis of archaeological material, and B) Mathematical problems involved in the building of symbolic systems, for taxonomic purposes in particular. This subdivision was further articulated in the publication of the Proceedings: Formulation des données et des types: problèmes théoriques; Techniques documentaires et classificatoires; Méthodes de taxinomie numérique; Etudes mathématiques de problèmes classificatoires; Discussions finales (Gardin, 1970).

The event had been planned since 1967, when Gardin proposed to the CADA Steering Committee to choose Marseille as the best venue for the Symposium that had been solicited by the Russian colleagues. They would have liked to hold it in Novosibirsk at the Laboratoire de Calcul pour les Sciences historiques, a place, however, that Gardin considered too decentralised, especially for American participants (JCG 2/1, 6 May 1967). Unfortunately, the Soviet scientists themselves were ultimately unable to attend the Symposium. Gardin and his research team convened on that occasion more than 100 participants from all over the world, giving rise to a lively debate, as recalled by Jean-Paul Demoule, in a poignant article, in which he emphasises the centrality of the period that revolves around 1968 (Binford, Binford, 1968; Clarke, 1968), and highlights Gardin's role as a reformer and a magnet for young archaeologists who were looking for a mentor in approaching archaeological data formalisation methods (Demoule, 2012). It is not possible to dwell upon each paper, but I would like to recall the names of some of the scholars who paved the way for the future developments of computer archaeology. I refer in particular to George Cowgill, who in 1967 had already written a well-researched article on 'Computer Applications in Archaeology' (Cowgill, 1967) and to James Doran, who was already approaching the theme of 'Archaeological reasoning and machine reasoning' and was also given the task of reading for the absent David Clarke his paper on 'Towards analytical archaeology. New directions in the interpretive thinking of British archaeologists'.

4. The early Seventies: new directions

Getting to the end of the Sixties, in order to conclude my session introduction, I would like to hint at two other international Conferences, held in the early Seventies, which represented a turning point in the history of archaeological computing from a methodological point of view: the Mamaia Conference in 1970 and the Marseille Conference in 1972, dedicated to 'Mathematics in Archaeological and Historical Sciences' and 'Les banques de données archéologiques', respectively (Hodson, Kedall, Tautu, 1971; Borillo, Gardin, 1974). Between these two Conferences, two other seminars are referred to by James Doran in his home interview as important milestones in the evolution of computer applications in archaeology: the 'Research Seminar in Archaeology and Related Subjects', organised in 1971 at the University of Sheffield by Colin Renfrew (Renfrew, 1973; cfr. also http://www.webofstories.com/play/colin.renfrew/37) and the Seminar on 'Les méthodes mathématiques de l'archéologie', convened again in Marseille in 1971 by CADA.

Thanks to the Mamaia and Marseille international meetings not only did the role of computers in archaeological research intensify, but, as clearly demonstrated in the title of the Conferences, the dichotomy that would long characterize archaeological computing materialised. As stated by Djindjian, the Mamaia Conference was the venue where famous statisticians encountered archaeology, and where archaeologists showed they knew how to use statistics (Djindjian, 2009). In Marseille, computer applications officially opened up to Classical Archaeology (see in particular Ginouvès, 1974; Christophe, Guimier-Sorbets, 1974) and to the problems raised by Cultural Resource Management (see in particular Chouraqui, 1974).

Of course, I would like to close by recalling the birth of the CAA Conference, inaugurated in Birmingham in 1973, which gave rise to the most important regular meeting focused on archaeological computing (Wilcock, 1973). But this is a new chapter in the history of archaeological computing.

References

BINFORD, S.R. and BINFORD, L.R. (1968) *New Perspectives in Archaeology*. Chicago: Aldine.

BORILLO, M. and GARDIN, J.-C. (eds) (1974) *Banques de données archéologiques. Actes du Colloque, Marseille, 12-14 juin 1972*. Paris: Éditions du CNRS.

Centre d'Analyse Documentaire pour l'Archéologie (1959). Paris: Éditions du CNRS.

CHENHALL, R.G. (1971), *The Archaeological Data Bank: A Progress Report. Computers and the Humanities*. 5 (3). p. 159-169.

CHOURAQUI, E. (1974) Le système d'exploitation automatique de l'Inventaire général des monuments et richesses artistiques de la France. In: BORILLO, M. and GARDIN, J.-C. (eds) *Banques de données archéologiques, Actes du Colloque, Marseille, 12-14 juin 1972*. Paris: Éditions du CNRS. p. 147-158.

CHRISTOPHE, J. and GUIMIER-SORBETS, A.-M. (1974) Travaux préparatoires à l'établissement d'une banque de données pour les mosaïques grecques et romaines. In: BORILLO, M. and GARDIN J.-C. (eds) *Banques de données archéologiques, Actes du Colloque, Marseille, 12-14 juin 1972*. Paris: Éditions du CNRS. p. 237-242.

CLARKE, D.L. (1968) *Analytical Archaeology*. London: Methuen.

Computers and their Potential Applications in Museums, New York, 15-17 April 1968 (1968). New York: Arno Press.

COWGILL, G.L. (1967) Computer applications in archaeology. *Computers and the Humanities.* 2 (1). p. 17-23.

DEMOULE, J.-P. (2012) Quarante ans après : une génération. In: GIRAUD, J. and GERNEZ, G. (eds) *Aux marges de l'archéologie. Hommage à Serge Cleuziou.* Nanterre: Maison René-Ginouvès, Archéologie et Ethnologie. p. 19-33.

DJINDJIAN, F. (2009) The Golden Years for Mathematics and Computers in Archaeology (1965-1985). In: MOSCATI, P. (ed) *La nascita dell'informatica archeologica. Atti del Convegno Internazionale, Roma, Accademia Nazionale dei Lincei, 24 ottobre 2008. Archeologia e Calcolatori*. 20. p. 61-73.

DORAN, J.E. and HODSON, F.R. (1975) *Mathematics and Computers in Archaeology*. Edinburgh: Edinburgh University Press.

GARDIN, J.-C. (1959) On the Coding of Geometrical Shapes and other Representations, with Reference to Archaeological Documents. In: P*roceedings of the International Conference on Scientific Information, Washington, D.C., 16-21 November 1958.* Washington, D.C.: National Academy of Sciences. p. 107-124.

GARDIN, J.-C. (1965) A Typology of Computer Uses in Anthropology. In: HYMES, D. (ed) *The Use of Computers in Anthropology.* London-The Hague-Paris: Mouton & Co. p. 103-117.

GARDIN, J.-C. (ed) (1970), *Archéologie et calculateurs. Problèmes sémiologiques et mathématiques. Actes du Colloque international, Marseille, 7-12 avril 1969*. Paris: Éditions du CNRS.

GARDIN, J.-C. (1971) Archaeology and Computers: New Perspectives. In: *Use of Computers, Documentation and the Social Science. International Social Science Journal*. 23 (2). p. 189-203.

GARDIN, J.-C. and JAULIN, B. (eds) (1968) *Calcul et formalisation dans les sciences de l'homme*. Paris: Éditions du CNRS.

GINOUVÈS, R. (1974) Projets de banques de données archéologiques à l'Université de Paris X. In: BORILLO, M. and GARDIN, J.-C. (eds) *Banques de* *données archéologiques. Actes du Colloque, Marseille, 12-14 juin 1972*. Paris: Éditions du CNRS. p. 221-224.

HEIZER, R.F. and COOK, S.F. (eds) (1960) *The Application of Quantitative Methods in Archaeology*, Viking Fund Publications in Anthropology, 28. Chicago: Quadrangle Books.

HODSON, F.R., KENDALL, D.G. and TAUTU, P. (eds) (1971), *Mathematics in the Archaeological and Historical Sciences. Proceedings of the Anglo-Romanian Conference, Mamaia, September 1970*. Edinburgh: Edinburgh University Press.

HYMES, D. (ed) (1965) *The Use of Computers in Anthropology*. London-The Hague-Paris: Mouton & Co.

IHM, P. (1965), Automatic classification in anthropology. In: HYMES, D. (ed) (1965) T*he Use of Computers in Anthropology*. London-The Hague-Paris: Mouton & Co. p. 358-376.

Information Processing (1960) *Proceedings of the International Conference on Information Processing, Paris, UNESCO, 15-20 June 1959*. Paris, München, London: UNESCO, R. Oldenbourg, Butterworths.

KOLČIN, B.A. (ed) (1965) *Arheologija i estestvennye nauki*, Moskva: Nauka.

MOSCATI, P. (ed) (2009), *La nascita dell'informatica archeologica. Atti del Convegno Internazionale, Roma, Accademia Nazionale dei Lincei, 24 ottobre 2008*. *Archeologia e Calcolatori*. 20.

MOSCATI, P. (2013) Jean-Claude Gardin (Parigi 1925-2013). Dalla meccanografia all'informatica archeologica. *Archeologia e Calcolatori*. 24. p. 7-24.

MOSCATI, P. (2014) Parcours culturels pour une histoire de l'informatique appliquée à l'archéologie. In: COSTA, L., DJINDJIAN, F. and GILIGNY, F. (eds) *Actes des 3èmes Journées d'Informatique et Archéologie de Paris, JIAP 2012, Paris, 1-2 juin 2012*. *Archeologia e Calcolatori*, Suppl. 5. p. 9-17.

NATALE, M.T. and SACCOCCIO, R. (2010) Museo and Web: un kit pratico per le istituzioni culturali che vogliono realizzare un sito web di qualità. *Archeologia e Calcolatori*. 21. p. 27-47.

RENFREW, C. (1973) *The Explanation of Culture Change: Models in Prehistory*. London: Duckworth.

SOKAL, R. and SNEATH, P.H.A. (1963) *Principles of Numerical Taxonomy*. San Francisco: W.H. Freeman.

SPAULDING, A. (1953) Statistical Techniques for the Discovery of Artifact Types. *American Antiquity*. 18. p. 305-313.

USTINOV, V.A. (1964) *Primenenie vychislitel'nykh mashin v istoricheskoĭ nauke*. Moskva: Mysl'.

WILCOCK, J.D. (ed) (1973) *Computer Applications in Archaeology 1. Science and Archaeology*. 9.

WILCOCK, J.D. (1999) Getting the Best Fit? 25 Years of Statistical Techniques in Archaeology. In: DINGWALL, L., EXON, S., GAFFNEY, V., LAFLIN, S. and VAN LEUSEN, M. (eds) *Archaeology in the Age of the Internet. CAA97. Computer Applications and Quantitative Methods in Archaeology. Proceedings of the 25th Anniversary Conference, Birmingham, April 1997*. BAR International Series 750. Oxford: Archaeopress. p. 35-52.

A Retrospective on GIS and AIS Platforms for Public Archaeology in Italy Searching Backward for Roots and Looking Onwards for new Methodological Road-Maps

Mirella SERLORENZI, Ilaria JOVINE, Giorgia LEONI, Andrea DE TOMMASI, Andrea VARAVALLO

Special Superintendence for Colosseum, National Archaeological Museum and Archaeological Heritage of Rome
Ministry for Cultural Heritage and Tourism of Italy
mirella.serlorenzi@beniculturali.it, ilaria.jovine@beniculturali.it, giorgia.leoni@beniculturali.it,
adt.andrea.detommasi@gmail.com, andrea.varavallo@beniculturali.it

Abstract

The state of the art of Geographical Information Systems and Archaeological Information Systems developed with public resources for the archaeological knowledge management, calls common attention to a deeper understanding of new AIS web platforms as 'places' of multi-domain integration and collaborative processes.

This paper presents the preparatory phase of an analytical work arised in the context of the SITAR Project, the first digital archaeological cadastre of Rome, and focused on the evolution and revaluation of AIS in the Public Archaeology domain.

Some early considerations are proposed with respect to: a philological retrospective on public AIS, with a particular attention to the italian scenario; a first evaluation of real correspondance levels between typical cartographic/GIS environments, semantic/ interpretative tools, and systems for analysing and mapping scientific data and informations; a first AIS subdomain basic ontology; finally, a proposal for a new AIS platforms declension with regard to their roles in Public Archaeology.

Keywords: Geographical Information Systems, Archaeological Information Systems, Public Archaeology, Archaeological Heritage of Rome, Participatory Archaeological Knowledge.

1. Introduction

The state of the art of Geographical Information Systems (GIS) and Archaeological Information Systems (AIS) developed with public resources to manage complexity of archaeological knowledge, calls attention of institutional, academic and professional communities to a deeper and more aware understanding of the cultural values – including also open issues[1] – acquired in last decades by these specialized Information Systems. Effectively, a new attention seems to be paid to knowledge processes refinements, methodological evolutions, technological enhancements and innovations that have began to characterize Public Archaeology and its social roles since the last decade, also thanks to a wide diffusion of AIS platforms as new 'places' of integration between scientific domains and territorial managing and planning actions.

In this sense, our preparatory work looks at the epistemological opportunity to reconsider ideas, concepts, ontologies, methods, technologies, know-how, opening vs. closing trends, all underpinning both to fallen or still alive projects, pilot experiences or simple feasibility studies. Moreover, a particular interest is dedicated to those projects promoted by public institutions. Therefore, this early contribute tries to 'travel' through historical and theoretical backgrounds, different cultural policies, scientific and technological references, logical and procedural architectures, concrete achievements and their effects on the evolution of archaeological research and knowledge, both in past and present time.

This work arises specifically in the context of the SITAR Project,[2] the experience of public archaeological knowledge management that has been promoted since 2007 by the Special Superintendence for Colosseum, National Archaeological Museum and Archaeological Heritage of Rome (SSCol), a territorial institute of italian Ministry for Cultural Heritage and Tourism of Italy (MiBACT). Carried out to realize and maintain the first digital archaeological cadastre for the metropolitan territory of Rome (Serlorenzi and Leoni, 2015), from the beginning SITAR has been implemented as an AIS web platform. Actually, it is being developed and populated in order to rapidly reorganize, manage and represent the complexity of knowledge about Rome archaeological heritage, in multiple ways, for many different users and through suitable web media.

These are fundamental reasons for which our project workgroup's, first of all, has to become more aware of direct and indirect derivations of SITAR theoretical framework from mentioned past experiences, as its roots. In this sense, the purpose is to better steer all continuous methodological and technological developments of

[1] In this sense, 'concepts' themselves are seen and employed as epistemological means bearing hypothesis and not only investigated as subjects of the research (Margiotta, 2011, p. IX-X).

[2] The Italian acronym SITAR means 'Archaeological Territorial Infomative System of Rome'; see at http://sitar.archeoroma.beniculturali.it/.

SITAR, also to share our approaches and observations with other researchers and collegues.

In that direction, this paper presents some early considerations articulated in the following paragraphs with respect to: a proposal of a potential approach to a philological retrospective on public AIS systems in Italy; a first evaluation of their role in theoretical and methodological innovation of Public Archaeology; the definition of an early basic ontology for AIS domain, in order to reconsider main knowledge elements to be analyzed in each project; finally, a brief proposal for a new AIS platforms conceptual declension, towards a clearer recognition of their common theoretical and methodological roots and specific roles in Public Archaeology domain.

2. Towards a retrospective on italian public AIS: an approach proposal

The attention paid within SITAR Project to continuous innovations in GIS and AIS platforms development and their integration into other cultural domains, represents the starting point for our retrospective approach. In this sense, it seems to be relevant the active participation of SITAR Project in some italian and european archaeological networking experiences, such as: committees and workgroups promoted and coordinated by MiBACT since 2007, in the field of GIS, AIS, Spatial Data Infrastructures (SDI), Open Data and Preventive Archaeology; the two ongoing european projects ARIADNE – 'Advanced Research Infrastructure for Archaeological Database Networking in Europe' (Niccolucci, 2014)[3] and DCH-RP – 'Digital Cultural Heritage Roadmap for Preservation' (Justrell and Fresa, 2014);[4] and also some other valuable cooperations with Universities and Research Institutions as the National Research Council of Italy - Department for Social Sciences, Humanities and Cultural Heritage (CNR-DSU); the italian Agency for New Technologies, Renewable Energies and Sustainable Development (ENEA); the Consortium GARR, the managing body of 'GARR-X', the Italian National Research and Education Network (NREN).

Due to this specific institutional perspective, our early analyses will pay a special attention to those initiatives promoted in last decades by MiBACT and implemented in the two prevalent paradigms of the so-called Cultural Resources Management (CRM) systems and more recent AIS platforms. Our work will begin from those projects undertaken since mid 1980s, on the wave of so-called 'giacimenti culturali' (cultural deposits), in light of the legacy left by those pioneers experiences to descendant initiatives and early applications in the public archaeological sector. The aim is to revaluate various effects of those projects on involved communities as, first of all: the 1990s GIS and AIS outbreak and diffusion, the domain language formalisation and the birth of last generation of public information systems (fig.1).

[3] See at http://www.ariadne-infrastructure.eu/.

[4] See at http://www.dch-rp. eu/.

Therefore, a foundamental step for this work is to trace now – and to deepen in future – the origins of GIS and AIS

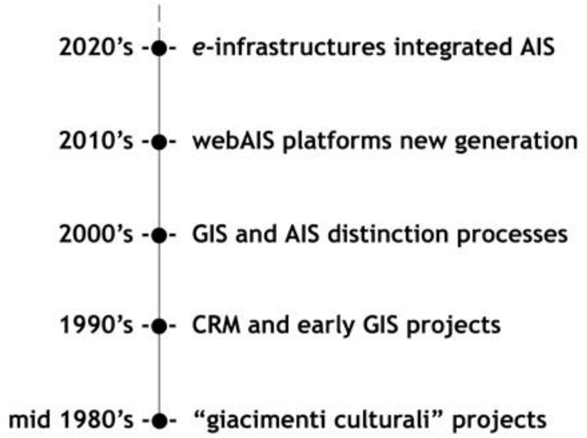


Figure 1: Time-line of GIS and AIS evolution path in italian Public Archaeology (source: SITAR Project).

applications in public italian archaeology, experimenting a philological approach: dealing with various projects, experiences and feasibility studies, each one quite as a 'textus', the relationships between epistemological/ methodological evolution of Public Archaeology and specific coeval technology scenarios, may be analyzed and mapped to point out the most relevant elements of knowledge. In this direction, our work will attempt to an early discriminating evaluation of failures and successes, sharing and diversifications of approaches and methods, natural aggregations between institutions, expertises specialization processes, and so on, all of them as perceived through literature analysis and, above all, by a direct observation of the Public Archaeology current scenario. In our actual perspective, the mentioned cultural and technological elements to be analyzed, may be primarily observed across some paradigmatic case study such as national cultural Information Systems, regional or local AIS and also SDI, like those ones developed in many cities and regions of Europe, in last years. Nevertheless, other GIS and AIS projects, even if less complex, could be considered as precious contributes to our early reflections.[5]

For the purposes of this paper, we look at CRM, GIS and AIS applications - already well known subjects of thematic literature, specifically observed in specific surveies and analysis since the 1990s (Moscati, 1998; Djindjian, 1998; Scianna and Villa, 2011) - as three items integrating each other (figg.2, 3) and, in some cases, as consecutive stages of GIS approach evolution in Public Archaeology domain (Harris and Lock, 1995).[6] So, with the acronim AIS we

[5] In view of deeper analysis of scientific and technical literature, all resources and papers consulted in this phase have been primarily selected among those ones directly available from following on-line journals and repositories: Archeologia e Calcolatori, CAA Proceedings web site, Journal of Computing in Cultural Heritage, ResearchGate, Academia. edu, Fasti on-line. The queries for resources selections have been based on some basic keywords as 'GIS', 'AIS', 'Archaeological GIS', 'GIS+Archeologia', 'SIT' (in italian: 'Sistema Informativo Territoriale'), 'IDT' (in italian: 'Infrastruttura Dati Territoriali'), 'SDI', 'Spatial Archaeology', 'Spatial Analysis', 'Digital Libraries', 'Open Data', 'Public Archaeology'.

[6] With respect to the evolution of these information system declensions, the statistics on use of such terms as 'GIS' and 'CRM' supplied by

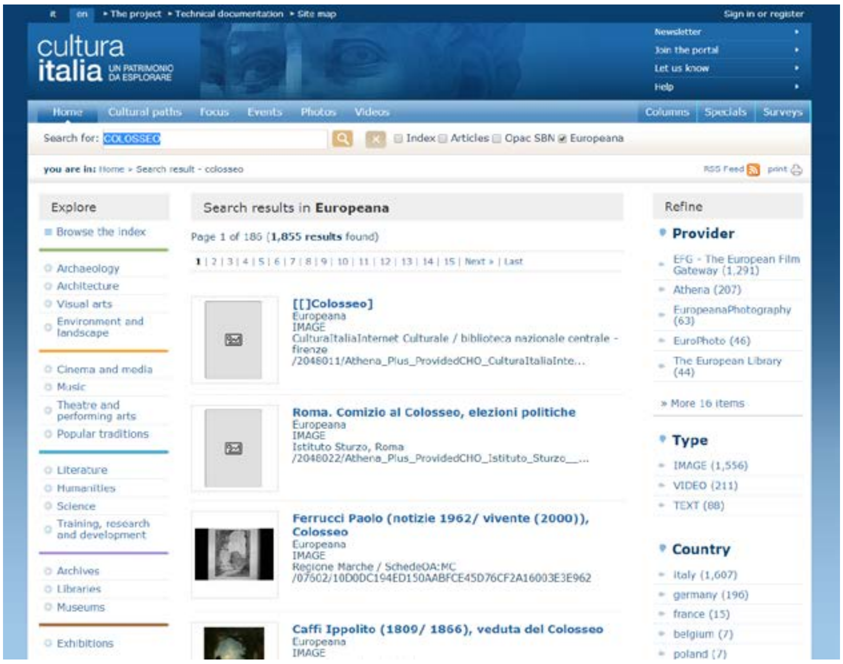


FIGURE 2: AN EXAMPLE OF INFORMATION COMPARISON BETWEEN CRM AND AIS: A QUERY FOR 'COLOSSEO' WITHIN CRM SYSTEM 'CULTURAITALIA.IT' (SOURCE: MINISTRY FOR CULTURAL HERITAGE AND TOURISM OF ITALY).

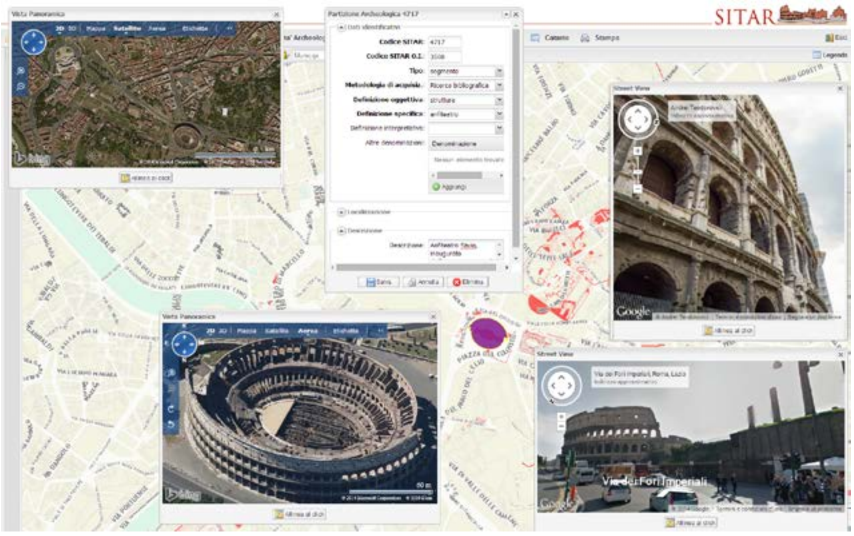


FIGURE 3: AN EXAMPLE OF INFORMATION COMPARISON BETWEEN CRM AND AIS: MULTI-REPRESENTATION OF COLOSSEUM SPATIALIZED AND DESCRIPTIVE DATA IN SITAR WEBAIS PLATFORM (SOURCE: SITAR PROJECT).

will generally refer both to so-called archaeological GIS applications and more complex public archaeological information systems. In particular, among various definitions of AIS, as avalaible in literature, we look at those ones explicited in Gillings and Goodrick (1996), Arroyo-Bishop (1998) as 'the use of the GIS structure to base the Archaeological Information System (AIS)', Djindjian (1998, 2012), Cavulli and Grimaldi (2005) and also to the reflections formulated by Carver (2005).

In such retrospective, it seems to be fundamental the adoption of a mapping approach among: theoretical paradigma; logical, phisical and semantical objects; technical tools and systems; digital infrastructures;[7] and so on. Indeed, our first aim is to describe an early basic ontology of the 'historical AIS landscape' and analyse its current multiple representations in form of various digital platforms dedicated to Public Archaeology, with a particular attention for recent spreading and evolution of public AIS web platforms. In other words, it will be attempted to implement a metaphorical 'GIS of AIS' by means of which to map and query some primary 'knowledge layers' related to ideas, concepts, know-how, approaches, methods, ontologies, technologies, persons, etc., involved in this specific kind of technological applications.

the online journal Archeologia e Calcolatori, appear very interesting to suggest the underpinning trends. See at http://soi.cnr.it/archcalc/overview_classification.htm.

[7] For this specific topic, the outcomes of the survey 'Research Infrastructures for Cultural Heritage in Italy - 2014' carried out jointly by MiBACT and Consortium GARR, will appear very relevant. See at http://www.garr.it/a/comunicazione/notizie-dal-garr/news/743/.

3. 'What, where, when' isn't ever visually equal to 'who, why, how'

The literary metaphor offered by José Saramago through his tale 'O Conto da Ilha Desconhescida' seems to be a nice figurative pass key to introduce our early rereading of GIS and AIS history in Public Archaeology, and also to approach to the definition of a basic ontology for these applications. In the mentioned tale, a no-named main character – who, metaphorically, could be also an archaeologist – asks to a 'king' – which may be seen as a personification of any epistemological value and issue – of an elsewhere 'kingdom' – for our perspective, the archaeological domain – for a 'boat' – it may be seen as an expert system based on methods, processes and tools – to reach an 'unknown island' – that could represent finalities of the research itself, as they are not ever completely claryfied or defined from the beginning. That island seems not to be reported on kingdom official maps and therefore not to exist, as the vexed king answers to the protagonist. Indeed, that is really the unique reason for which it is 'unknown', as the protagonist replies to the king; but, at the same time, according to him that element doesn't appear enough to declare the unknown island not existing. Seeing this scene, others people begin to shout to the king 'Give him the boat!' and so finally the man achieves his purpose and can begin preparations for sailing away from the kingdom harbour, neither having a crew on board nor being himself both a sailor and an expert captain. Then, the tale switches to a dreamlike dimension that makes the protagonist understand the 'real' unknown island he was looking for, it is effectively the boat itself. Suddenly, in the dream, the little ship begins to change itself in a sailing flourishing garden.

For our analysis context, that pleasant literary metaphor seems to offer two main suggestion elements. The first one is the focus oriented just so on the rediscovery of richness and potentiality of the 'medium' itself, once the researcher has gone on board, obviously. Indeed, that seems to be happened also in the case of AIS platforms, if we consider the path of their breakthrough and widespread success in archaeological domain. Looking from this metaphorical perspective, therefore, the AIS diffusion could be better understood also beyond pervasive availability of software and hardware, and ICT, 2D/3D web mapping and other abilitating technologies.

At the same time, the metaphor suggests another fundamental issue related to ambiguity of the concept of 'official map', today more than in the past: it is clear, indeed, that the same 'institutional map' couldn't ever represent enough informations for all users, their own 'mind road-maps' and imaginations; nevertheless, as a 'public datum' the official mapbases should be ever considered as a topic starting point to discover – or better, re-discover – all unmapped items and create different visual representations of new data, theories, scientific discoveries and 'consciousness layers', we can say; in other words, to materialize all new personalized and shareable 'unknown islands'.

With regard to both these reading levels and questions underpinning to title of this paragraph, it may pointed out, on one hand, the importance to revaluate roles of advanced information systems in Public Archaeology evolution path, in processes of circular knowledge creation and in continuos innovation of historical disciplines; on the other hand, new declensions of AIS platforms have to address the current gaps between needs of an 'official representation', in geographical terms, of Public Archaeology and characteristics of 'personalized maps' being produced by different users.

Particularly, the latter question refers specifically to real correspondance levels between typical visual/cartographic systems and semantic/interpretative tools/system for mapping and analysing activities, especially needed in social sciences and humanities. Indeed, while the first ones are essentially based on typical graphic combination of three primary topics as 'what', 'where', 'when' and conventional data representations, the latter instruments are more extended across interpretative concepts as 'who' (persons/societies), 'why' (functions/cultures), 'how' (technologies/processes), moreover across all different epochs (fig.4). This is an epistemological issue, furthermore, that has been already noticed and discussed by many authors (Castelford, 1992; Arroyo-Bishop and Lantada Zarzosa, 1995; Barceló and Pallarés, 1996; Gardin, 2002; Conolly and Lake, 2006, pp.8-10; Constantinidis, 2007; De Runz *et al.*, 2011; Kondo *et al.*, 2011; Desjardin, Nocent and De Runz, 2012; Djindjian, 2012) and has been stressed also in the political geographical domain by means of some interesting theoretical assertions (Forest, 2004).

4. Searching backward for roots through a basic ontology of AIS subdomain

Starting our first overview on AIS applications from these premises, the metaphor kept from Saramago's tale may give us a simple but useful suggestion to approach and describe a basic ontology through which analyse AIS case studies and trace our early theoretical reflections. In order to point out the more interesting key defining features and knowledge elements of that 'landscape', the following aspects have to be considerd for describing a primary analytical matrix:

– the specific archaeological domain, obviously, containing and identifing theoretical and methodological needs and approaches, processes and procedures, scientific finalities, values and issues, tangible and intangible archaeological heritage items and all their translations into digital objects; this domain appears as the 'kingdom' in the mentioned literary metaphor, representing the 'territorial context' for our retrospective and so the 'spatial extension' of our attempted 'GIS of AIS';
– the human actors, involved and interacting in AIS applications in both institutional/collective and individual forms, with their own 'mind road-maps' and specific needs to be expressed and satisfied within new partic-

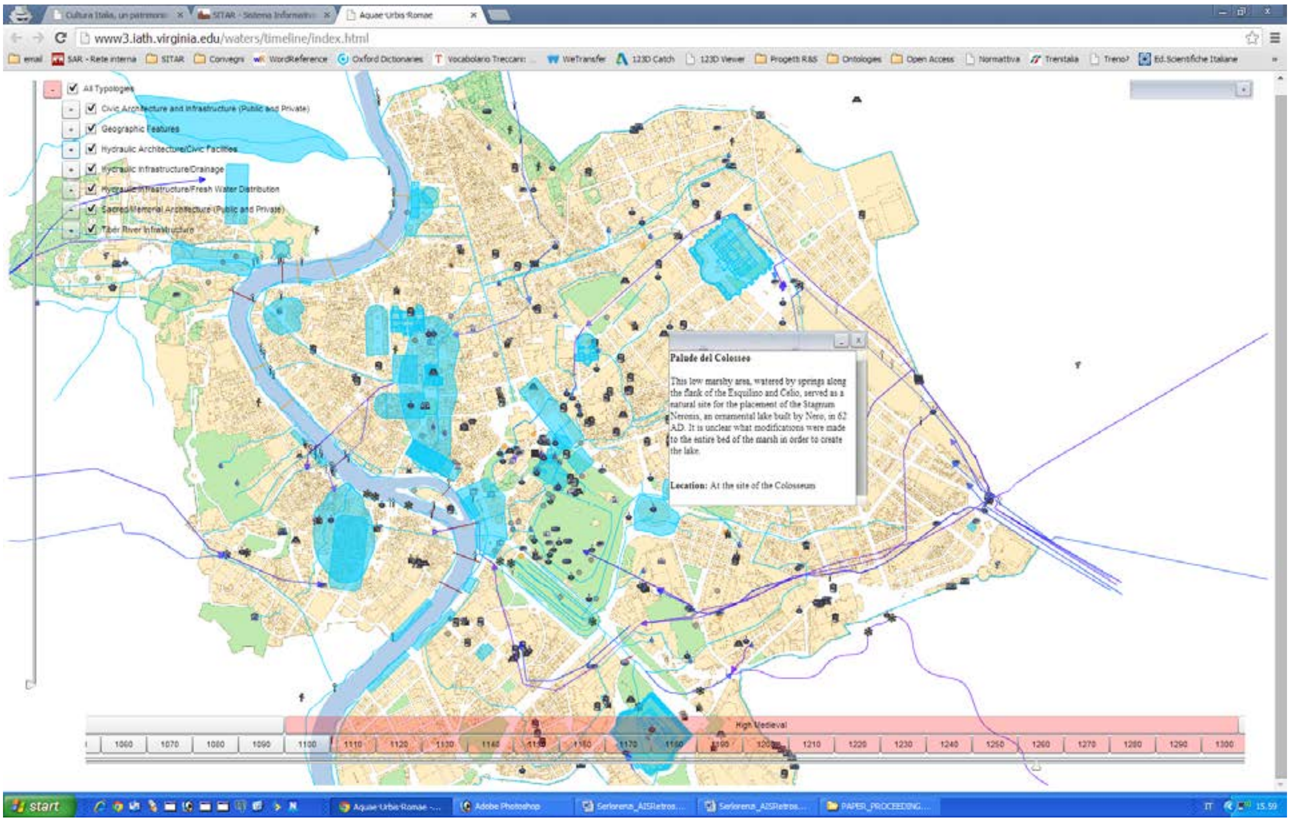


Figure 4: Mapping temporal dimensions: a thematic web map of hydraulic infrastructures and facilities in early medieval period at Rome, represented through the webGIS of 'Aquae Urbis Romae' Project (source: Aquae Urbis Romae Project, http://www3.iath.virginia.edu/waters/).

ipatory knowledge processes (as for protagonist and people of the tale);

– the epistemological values and open issues, as factors impeding or stimulating and allowing new research, achievements and awareness (as the king makes, before denying unknown island existence, then conceding the boat);
– the abilitating technologies and methodologies, as media allowing to us to reach new scientific and cultural outcomes, and becoming themselves new research subjects (as it happens metaphorically to the boat in the tale);
– the seeked object, as data, informations and more generally knowledge, 'waiting' for being achieved, analyzed, organised, shared and re-thought through scientific and cognitive processes, towards new augmented and participatory knowledge (as the unknown island, not ever officially mapped, but anyway existing and reachable).

As said above, the definition of that simplified ontology appears as a mandatory activity to approach and steer our preparatory work to better describe and analyse all the different conceptual instances and characteristic of AIS subdomain.

5. Towards a GIS-oriented collation of public AIS platforms

In our perspective, it may be attempted a GIS-based collation of primary case studies to be analyzed among public italian AIS projects. So, being also inspired by the initiative for the creation of an 'Observatoire des pratiques géomatiques dans les organisations de l'Archéologie' in France (Costa, 2012), the first step in implementation path of our expected 'GIS of AIS' is represented by an usual census survey, newly started and still in progress at this moment. In this direction, the real effort should be to go beyond the typical data-entry of identifying informations, both spatial and descriptive, and towards a deeper understanding of cultural elements represented by each project and experience. So, for purposes of this paper it seems more useful to express some theoretical considerations rather then to examine in technical details our collation work just started, that would be detailed in next publication places.

As in a usual GIS application, first of all it is necessary to set the extension of the context in which to map, analyze and aggregate all interesting 'datasets' corresponding to each project, experience or feasibility study, both fallen and still alive. In this sense, the 'territory' to be analyzed is effectively the Public Archaeology scenario, notoriously studded with different institutional mission profiles, professional figures, methods and above all with heterogeneous visions. Already noticed by Wilcock in 1973 as a ‘‘bridge subject’ between the two cultures of arts and sciences’, Public Archaeology is now characterized by many different and shared instances of improvement and exploitation of its social finalities and values, particularly aimed to align archaeological discipline and professional sector with the digital society rapid evolution, fluidity and strategic challenges.

Moreover, new trends and applications are bringing the whole sector to a stronger comparison with knowledge society trends, towards new declensions of archaeological discipline, such as the ‘open archaeology’ (Serlorenzi, 2013; Costa et al., in press). These most innovative items are, e.g., 3D GIS (Harris and Lock, 1995; Scianna and Villa, 2011) and 4D GIS (Castelford, 1991; Constantinidis, 2007; Johnson, 2008), preventive archaeology researches, ‘archaeological potentialities’ analysis and mapping (Cavazza, 2014), interdisciplinary integration e innovation (Campana and Forte, 2006), shared re-thinking of approaches, methods and procedures (Niccolucci, 2014), and also many valuable inputs and opportunities coming from socializing web platforms and tools. Furthermore, in the last decade the role played in the GIS applications field by ‘industrial’ actors in managing and supplying georeferenced, personalized and socialized knowledge contents, has pulled also archaeological communities towards new processes and ways of data production, sharing and dissemination, often speeding up digital growth of institutions and involved persons - especially in terms of competences and outcomes - and engaging of new audiences.

Within this reference framework, the effort of ‘collecting-for-rethinking’ public AIS applications assumes a greater relevance also in order to deal with some historical distinctions between various declensions of specialized informative systems. For example, some residual methodological differences between CRM projects and more specialized GIS and AIS applications, require for a necessary integration on behalf of wider and more flexible platforms really capable to address new common needs of knowledge management, along with cultural and social values enhancement, for many institutional, academic and research actors and other stakeholders (Moscati, 2009; Aubry and Ferjani, 2012; Costa, 2012; Djindjian, 2012; Hofmann and Mani, 2012; Jensen, 2012, p. 212). Therefore, in this scenario the themes of organization and multi-representation of fluid and collaborative knowledge appear as two real priorities to be strongly developped indeed by means of a new generation of advanced information web platforms that may be derived from the full integration of GIS, AIS and SDI with Digital Libraries and Open Access repositories, as new forms of so-called ‘Geo-Digital Libraries’ (Pozzo and Virgili, 2013; Pozzo, in press).

In the case of italian archaeology, public AIS projects appear to be still characterized by the legacy grown around experiences promoted by MiBACT in mid 1980s’ and early 1990s’, that have drawn the first operative framework in which traditional archaeological processes have gradually met and used potentialities of computer science.[8] In that same period, involved human actors have began to acquire new technological competences and public instances of Cultural Heritage management have been coupled gradually with professional/technical skills and competences of researchers, academic experts and private companies, giving birth to an important bi-directional comparison, even if not always complete and balanced. Anyway, those projects gave great impulse to circulation and adoption of multiple forms of data organization, knowledge representation and digital contents publication, particularly thanks to early descendant solutions based on web applications and GIS approaches and developed in various technological versions, firstly such as commercial desktop ones and more recently as free/open source and software-as-web-service platforms (Cantone, 2013; Serlorenzi, 2013).

With regard to this evolution framework, it may be generally observed that different MiBACT projects have maintained their own development paths across last decades and they are not still properly unified neither in a true ecosystem of web informative platforms, nor in a full integrated ‘cultural meta-system’. This situation may be due to different development perspectives related to various involved scholarships, workgroups and especially to a certain fragmentation of methods, resources, systems and tools, that often have affected the final compliance of these systems with each project premises, real needs and instances of different users. In this sense, the fundamental role of post-implementation reviews (Clubb and Lang, 1996b; Arroyo-Bishop, 1999) has not been widely applied in these projects development stories, and consequently final users have not been ever well involved in fundamental assessment processes.

In our opinion, two fundamental reasons for projects success or failure cases seem to be recognized: the first one can be referred to same internal relationships between specific executive workgroups; the second one may be located at the level of crucial interactions between these actors and real final users of each on system. Furthermore, this fragmentation phenomenon seems to have also limited the attention paid to dialogue and relationships between various institutional, research and academic levels, two elements seen as allowing or impeding factors for a successful implementation of public information systems, as observed and stressed in recent outlines of two MiBACT Committees on National AIS development (Serlorenzi and Jovine, 2013), as well as in other european contexts during last decades (Clubb and Lang, 1996a; Arroyo-Bishop, 1998; Costa, 2012, p. 265).

[8] A syntetic visual summary of that pioneering period has been traced by Biallo (2009), very useful to describe at a glance the main relationships between those early MiBACT experiences and primary descendant public AIS projects.

Starting from these common premises and, less or more, parallel paths, the italian scenario has been enriched with many new methodological approaches and innovations that have been achieved and specialized within some AIS projects carried out in last years. Some of them have been extended up to a regional or wider scale (Hiebel and Hanke, 2008; Miele, 2011; Cavazza, 2014), while others have been limited to specific case studies of historical urban centres, or again just to single archaeological sites (Lazzeri, 2011; Keay and Earl, 2013). So each project results more or less focused on a specific topic, for example the so-called 'archaeological risk' or the 'archaeological potentialities' analysis and mapping (Cavazza, 2014), or advanced data management. Particularly, with respect to AIS platforms dedicated to historical urban centres, there are some interesting examples to be mentioned for addressing crucial themes of updating, sharing and dissemination of data and knowledge, since the beginning of their development. These are distinctively the 'SITAVR Project', the first digital archaeological cadastre for the urban center of Verona (fig.5), derived from SITAR data model and operational paradigma (Basso et al., in press); the 'MAPPA Project', a stimulating institutional/academic experience focused on management, web-sharing and dissemination of archaeological dataset, 'archaeological potential' analysis and scientific knowledge about the ancient centre of Pisa; and, again, the 'SIURBE Project' focused on integrated geo-archaeological knowledge of the historical center of Benevento, vehiculated through an AIS web platform (Santoriello, Rossi and Rossi, in prep.).

Obviously, as said in the premises, many other italian relevant projects actually stud the national AIS constellation and all of them will contribute to our analysis and identification of cultural values and open issues related to the evolution of public archaeological knowledge platforms.

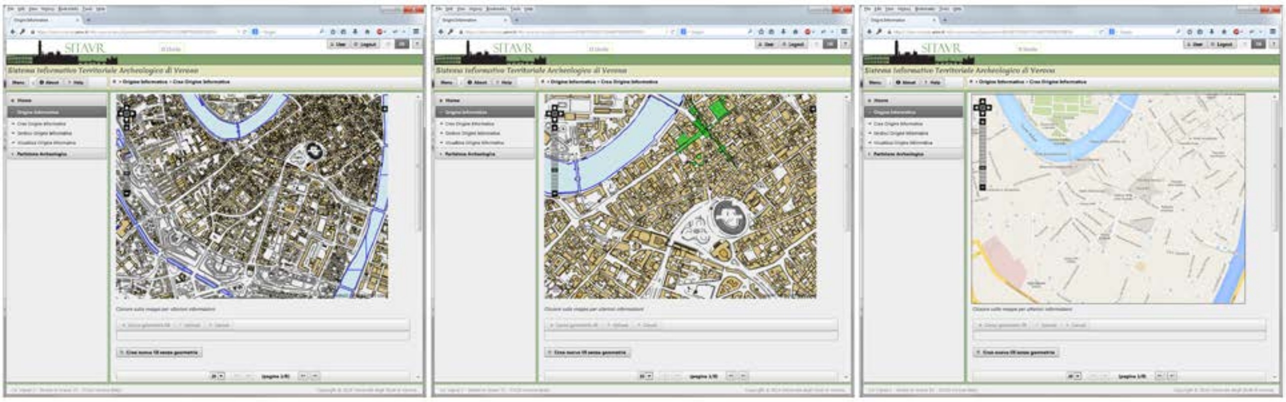


Figure 5: SITAVR project web platform and GIS environment (source: SITAVR Project).

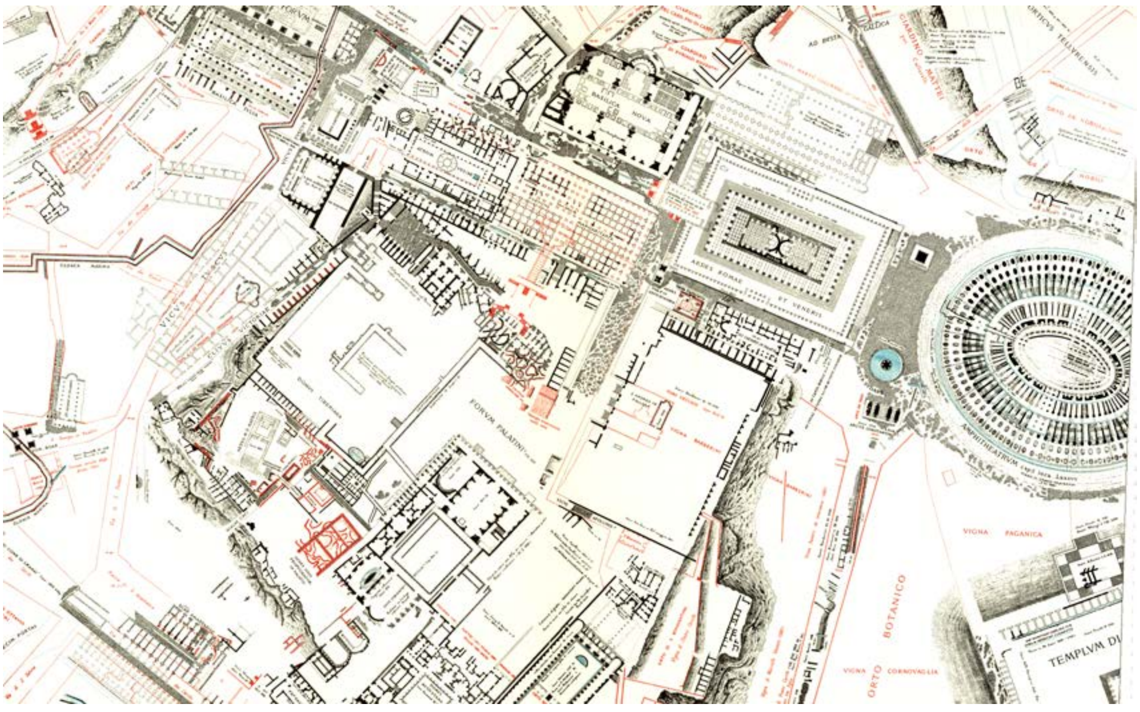


Figure 6: A detail of the Forma Urbis Romae by Rodolfo Lanciani: the archaeological area around the Colosseum (source: R. Lanciani, Forma Urbis Romae, table 29).

6. The case study of Rome

For our anlysis purposes, the case study of Rome appears as a 'bridge' between the past of archeological mapping tools and the future of archeological knowledge management systems. Indeed, it represents a paradigmatic case study in light of its mosaic of different AIS applications developed in last decades by various institutional, research and academic actors. Undoubtedly for its own ancient history, wide archaeological heritage well distributed in every part of the metropolitan territory and real complexity of its rapid urban, social and economical transformations, Rome has been the privileged subject of many archaeological mapping projects, since by those ones started and implemented by Rodolfo Lanciani between 1893 and 1901, by means of his famous 'Forma Urbis Romae' (Lanciani, 1981) (fig.6).

Following that precious 'archetypal map', some later updating activities have been promoted by University of Rome, until mid 1980s, and others have been undertaken by ex-Ministry for Education - Superintendence for Antiquities of Rome (then become the Special Superintendence for Colosseum, National Archaeological Museum and Archaelogical Heritage of Rome). Complessively, those activities have been extended from 1947 to 2005, through different, not ever continuous initiatives of archaeological mapping based on bibliographic notices and archive data, as available in different periods. Moreover, since 1995 this 'maps mosaic' has been enriched also by preparatory works of a Commission in charge of the updating of Lanciani's 'Forma Urbis Romae', promoted by the Council of Rome and its own Sovraintendenza Comunale office, in cooperation with the University of Rome 'Sapienza'.

On one hand, some results of these shared activities of data collecting, updating and checking, along with some early proofs of databases integration, flowed into the wider municipal informative system and linked with other cultural data on museums, art galleries, heritage items and territorial sites; on the other hand, many data acquisitions have been recently revised and collected into the 'Atlante Storico di Roma Antica' (Carandini, 2012). Moreover, for modern historical periods some very valuable data banks have been implemented and also supplied of GIS components and web applications, such as the 'Imago II Project' promoted by the Archivio di Stato di Roma, Council of Rome - Sovraintendenza Comunale office, University of Rome 'Torvergata' (Micalizzi *et al.*, 2011; http://www.cflr.beniculturali.it/), and those ones realized by CROMA, the Center for historical economic urban studies of University of Rome 'Torvergata'. In particular, those projects have dealed with digitization and georeferencing of some topic historical cartographic resources such as the 'Carta di Roma' edited by G.B. Nolli (1748), the so-called Pio-Gregoriano cadastral maps ensemble, drawn during last decades of the roman Papato authority (Lelo and Travaglini, 2013), and all the other maps series derived from it (fig.7).

All together, these older maps and more recent archaeological mapping projects represent the scenario in which different cultural institutions have promoted their own latest initiatives, too often without a unified vision on epistemological issues and real needs of different communities of users, underpinning to archaeological knowledge mapping and managing. Unfortunately, this situation has brought to a clear lack of shared purposes and methodological/technological solutions, instead of promoting a single public point-of-access to archaeological field informations, resources libraries and data banks.

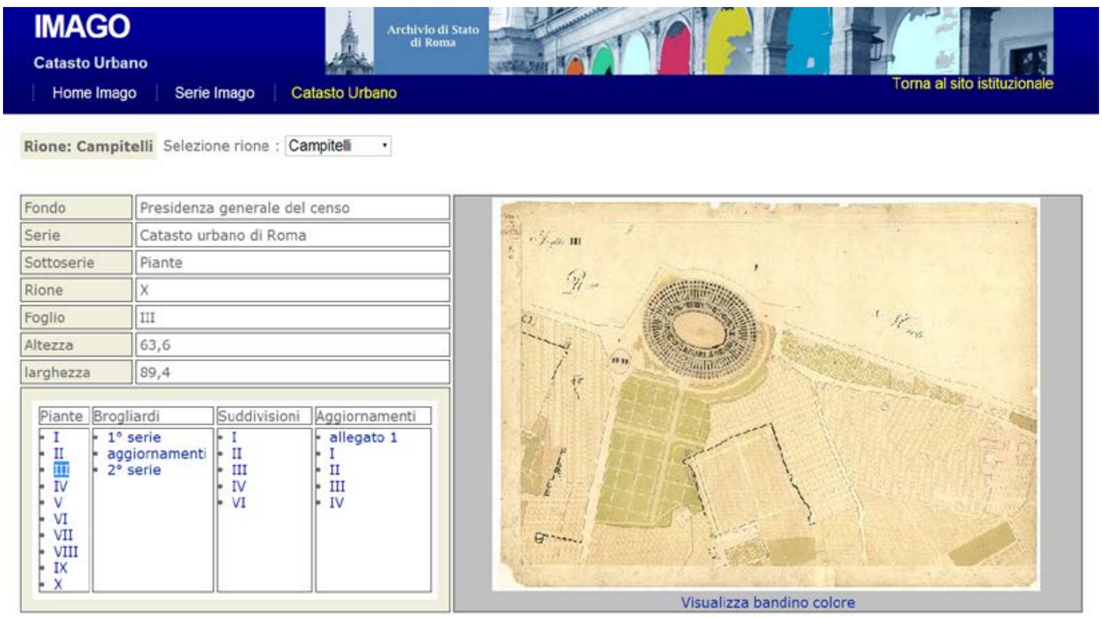


Figure 7: A screenshot from 'Imago Project' web site: the cadastral map of Colosseum area, derived from 'Pio-Gregoriano' cadastre of XIX century (source: Imago Project, http://www.cflr.beniculturali.it/).

Since 2007, this scenario has been further enriched thanks to development and maintenance of the first digital archaeological cadastre for the metropolitan territory of Rome, undertaken through SITAR Project and specifically its AIS web platform. This effort has been motivated also by SSCol accountabilities and competences in terms of a rapid re-organization, fluid managing and correct multi-representation of Rome complex archaeological knowledge, that is being translated from field and archive data/documents into more accessible and useful metadata and digital objects, on behalf of many different SITAR users and through specific web collaborative instruments. In this direction, SITAR platform aims to offer an effective decision-support-system for territorial and urban planning and sustainable development processes, whenever it is necessary to cooperate between SSBAR and other public administrations and local government bodies for new action plans delivery.

6. Looking onwards for a common 'AIS road-map'

In light of all these premises and theoretical reflections, also in view of next phases of our retrospective work, it seems possible to highlight some early open conclusions. First of all, it may be pointed out the importance of more clear and shared purposes in development processes of ongoing public AIS platforms, as well as in creating the new ones. Among the other key features, following seem to be firstly put in evidence:

- it should be constantly guaranteed full access and sharing of methodological approaches to all the interested researchers and final user, in order to improve a virtuous circuit of participatory procedural and technological enhancements;
- all the promoting institutions should support concrete availability of archaeological data and knowledge through different suitable digital interfaces;
- these user interfaces should be better tailored on specific users requirements and with full respect of preservation of data and knowledge themselves;
- the knowledge platforms should have to be easily accesible for, explained and delivered to everybody, including obviously non-specialist publics, especially through well personalized accesses.

In conclusion, the new frontier of public AIS development seems to be an advanced and e-infrastructures-based 'cultural meta-system' that should be seen as a constellation of renewed AIS, potentially named PArKS as an acronym for 'Public Archaeology Knowledge System'. For the evolution of such a digital cultural infrastructures will be necessary a clearer and shared awareness of common epistemological, methodological, cultural roots and social roles, even if in full respect of specific purposes of each experience and above all on behalf of Public Archaeology domain.

In this perspective, new and ongoing AIS development projects should have to take into account the current complex scenario, in order to be more deeply inspired by new Public Archaeology instances and compliant with critical societal challenges.

Acknowledgements

We are very grateful to:
Paola Moscati (CNR-ISMA, Editor of the Journal 'Archeologia e Calcolatori') for all valuable and continuous support to these reflections;
Anna Conticello (MiBACT – General Secretariat) and Alberto Bruni (MiBACT) for their precious suggestions in reconstructing the history of GIS and AIS projects promoted by MiBACT in last decades;
Rossella Caffo (Director of MiBACT - ICCU), Sara Di Giorgio and the Colleagues of ICCU and others Instituts of MiBACT, for shared experiences in the ongoing european projects ARIADNE and DCH-RP;
Riccardo Pozzo (Director of CNR – SSH Department) and the Colleagues of CNR-SSH Department for their valuable reflections and contributes in inspiring also this paper;
Patrizia Basso, Piergiovanna Grossi and Alberto Belussi (University of Verona) for sharing with us the experience of the new development of SITAVR web platform;
Andrea Scianna and the Colleagues of GISLab (University of Palermo – DICAM and CNR – ICAR) for shared reflections about 3D GIS data modelling in archaeology;
Francesca Cantone (CNR – IRAT) for the sharing of many ideas on cultural GIS and especially the updated survey on GIS/AIS in the territorial context of Campania, Italy;
François Djindjian (University of Paris – Panthéon-Sorbonne) as chairman of the session at CAA 2014 in which these early reflections have been firstly presented, and for his enlightening literature contributes;
and to all the other Colleagues and all SITAR Project Workgroup, for their shared suggestions, ideas, user cases kindly provided to this preparatory work.

References

ARROYO-BISHOP, D. (1998). GIS and archaeology in France. *Archeologia e Calcolatori* 9. p. 31-45. Available from: http://soi.cnr.it/archcalc/indice/PDF9/09_04_Arroryo.pdf. [Accessed: 15th July 2014].

ARROYO-BISHOP, D. (1999). From Earth to Cyberspace: the unforeseen evolution. *Archeologia e Calcolatori* 10. p. 7-16. Available from: http://soi.cnr.it/archcalc/indice/PDF10/10_01_Arroyo.pdf. [Accessed: 18 March 2014].

ARROYO-BISHOP, D., LANTADA ZARZOSA, M.T. (1995). To be or not to be: Will an object-spacetime GIS/AIS become a scientific reality or end up an Archaeological Entity?. In Lock, G. and Stancic, Z. (eds.) *Archaeology and Geographic Information Systems. A European Perspective.* London: Taylor and Francis. p. 43-54. Available from: http://books.google.it/books?id=ZPTLjMLSc-UC. [Accessed 17 July 2014].

AUBRY, L., FERJANI, S. (2012). Le système d'information du programme 'Archéologie du Bassin parisien'.

Entre SIG et SGBD, vers un applicatif Open Source adapté. *Archeologia e Calcolatori*. Supplemento 3. p. 75-95. Available from: http://soi.cnr.it/archcalc/indice/Suppl_3/06-aubry-ferjani.pdf. [Accessed: 13 July 2014].

BARCELÓ, J.A., PALLARÉS, M. (1996). A critique of G.I.S. in archaeology. From visual seduction to spatial analysis. *Archeologia e Calcolatori*. 7. p. 313-326.

BASSO, P. , BELUSSI, A., GROSSI, P. , CAVALIERI MANASSE, G. (in press). Il work in progress del SITAVR: primi passi metodologici, modello dati e documenti condivisi del sistema informativo territoriale archeologico. In: Serlorenzi, M. and Leoni, G. (eds.). *III Convegno SITAR. Il SITAR nella Rete della Ricerca Italiana. Verso la conoscenza archeologica condivisa. Roma 2013.* Archeologia e Calcolatori. Supplemento [still to be defined]. Firenze: All'Insegna del Giglio.

BIALLO, G. (2009). I sistemi informativi a riferimento geografico del Ministero per i Beni e le Attività Culturali. [Online presentation slides]. Available from: http://www.uniroma2.it/didattica/sist_inf/deposito/progetti_nazionali__beni_culturali_marzo_2009_leggera.pdf. [Accessed: 14July 2014].

CAMPANA, S., FORTE, M. (eds.) (2006). *From Space to Place: 2nd International Conference on Remote Sensing in Archaeology. Proceedings of the 2nd International Workshop. Rome 2006.* BAR International Series. 1568: Oxford, Archeopress. Available at: http://www.academia.edu/2465816/FROM_SPACE_TO_PLACE._IInd_International_Conference_on_Remote_Sensing_in_Archaeology. [Accessed: 15 July 2014].

CANTONE, F. (ed.) (2013). *ARCHEOFOSS. Open Source, Free Software e Open Format nei processi di ricerca archeologica. Atti del VI Workshop. Napoli, Italy, 2011.* Napoli: Naus Editoria.

CARANDINI, A. (ed.) (2012). *Atlante di Roma Antica. Voll. 1-2.* Roma: Electa.

Carver, G. (2005). Archaeological Information Systems (AIS): Adapting GIS to archaeological contexts. In: W. Börner and S. Uhlirz (eds.). *Workshop 9 – Archäologie und Computer 2004. Wien, 2004.* Wien: Phoibos. Available from: https://www.academia.edu/173585/Archaeological_Information_Systems_AIS_Adapting_GIS_to_archaeological_contexts. [Accessed: 16 July 2014].

CASTLEFORD, J. (1992). Archaeology, GIS, and the Time Dimension: an Overview. In: Lock, G. and Moffett, J. (eds.). *CAA91. Computer Applications and Quantitative Methods in Archaeology 1991.* BAR International Series. S577. Oxford: Tempus Reparatum. p. 95-106 Available from: http://proceedings.caaconference.org/paper/13_castleford_caa_1991/. [Accessed: 16 July 2014].

CAVAZZA, E. (2014). Linee guida per l'elaborazione della Carta delle potenzialità archeologiche del territorio. Ministero dei Beni e delle Attività Culturali e del Turismo - Regione Emilia Romagna. Available from: http://territorio.regione.emilia-romagna.it/paesaggio/pubblicazioni/LGpotarc.pdf/at_download/file/LG-pot-arc.pdf. [Accessed: 20 July 2014].

CAVULLI, F., GRIMALDI, S. (2005). To see or not To see. Archaeological data and surface visibility as seen by an AIS (Archaeological Information System) approach. In: Figueiredo, A. and Leite Velho, G. (eds.) . *The world is in your eyes. CAA2005. Computer Applications and Quantitative Methods in Archaeology. Proceedings of the 33rd Conference. Tomar 2005.* Tomar: CAA Portugal. p. 413-420. Available from: http://proceedings.caaconference.org/paper/61_cavulli_grimaldi_caa_2005. [Accessed: 12 July 2014].

CLUBB, N.D., LANG, N.A.R. (1996a). Learning from the Achievements of Information Systems; the Role of the Post- Implementation Review in Medium to Large Scale Systems. In: Kamermans, H. and Fennema, K. (eds.). I*nterfacing the Past. Computer Applications and Quantitative Methods in Archaeology CAA95. Vol. I. Analecta Praehistorica Leidensia. 28.* Leiden: Institute of Prehistory-University of Leiden. p. 73-80. Available from: http://proceedings.caaconference.org/paper/09_clubb_lang_caa_1995/. [Accessed: 13 July 2014].

CLUBB, N.D., LANG, N.A.R. (1996b). A Strategic Appraisal of Information Systems for Archaeology and Architecture in England – Past, Present and Future. In: Kamermans, H. and Fennema, K. (eds.). *Interfacing the Past. Computer Applications and Quantitative Methods in Archaeology CAA95. Vol. I. Analecta Praehistorica Leidensia. 28.* Leiden: Institute of Prehistory – University of Leiden. p. 51-72. Available from: http://proceedings.caaconference.org/paper/08_clubb_lang_caa_1995/. [Accessed: 12 July 2014].

CONOLLY, J., LAKE, M. (2006). *Geographical Information Systems in Archaeology.* Cambridge Manuals in Archaeology. Cambridge: Cambridge University Press. Available from: http://www.scribd.com/doc/139092272/Conolly-James-and-Lake-Mark-Geographical-Information-Systems-in-Archaeology. [Accessed 12 July 2014].

CONSTANTINIDIS, D. (2007). TIME to Look for a Temporal GIS. In: Figueiredo, A. and Leite Velho, G. (eds.). *The world is in your eyes. CAA2005. Computer Applications and Quantitative Methods in Archaeology. Proceedings of the 33rd Conference. Tomar 2005.* Tomar: CAA Portugal. p. 407-411. Available from: http://proceedings.caaconference.org/paper/60_constantinidis_caa_2005. [Accessed: 12 July 2014].

COSTA, L. (2012). La mise en place d'un observatoire des pratiques géomatiques dans les organisations de l'archéologie. *Archeologia e Calcolatori*. Supplemento 3. Firenze: All'Insegna del Giglio. pp. 265-278. Available from: http://soi.cnr.it/archcalc/indice/Suppl_3/20-costa.pdf. [Accessed: 17 July 2014].

COSTA, S., BEZZI, A., CANTONE, F., DE FELICE, G., GROSSI, P., LOTTO, D. (in press). The Italian job: open source in Italian archaeology. In: *AA.VV. Open Source Archaeology – Ethics and Practice.* Warsaw: Versita.

De Runz, C., Desjardin, E., Piantoni, F., Herbin, M. (2011). Towards Handling Uncertainty of Excavation Data into a GIS. In: Jerem, E., Redő, F. and Szeverényi, V. (eds.). *On the Road to Reconstructing the Past. Computer Applications and Quantitative Methods in Archaeology (CAA). Proceedings of the 36th International Conference. Budapest, 2008.* Budapest: Archeaeolingua. pp. 187-191. Available from: http://proceedings.caaconference.org/paper/cd15_derunz_et_al_caa2008/. [Accessed: 16 July 2014.

Desjardin, E., Nocent, O., De Runz, C. (2012). Prise en compte de l'imperfection des connaissances depuis la saisie des données jusqu'à la restitution 3D. *Archeologia e Calcolatori*. Supplemento 3. p. 385-396. Available from: http://soi.cnr.it/archcalc/indice/Suppl_03/AC_Suppl03_2012.pdf. [Accessed: 17.02.2014].

DJINDJIAN, F. (1998). GIS usage in worldwide archaeology. *Archeologia e Calcolatori*. 9. pp. 19-30. Available from: http://soi.cnr.it/archcalc/indice/PDF9/09_03_Djindjian.pdf. [Accessed: 05 July 2014].

DJINDJIAN, F. (2012). L'approche par les processus en archéologie. Archeologia e Calcolatori. Supplemento 3. p. 279-297. Available from: http://soi.cnr.it/archcalc/indice/Suppl_03/AC_Suppl03_2012.pdf. [Accessed: 17 July 2014].

FOREST, B. (2004). Information sovereignty and GIS: the evolution of 'communities of interest' in political redistricting. *Political Geography*. 23. pp. 425-451.

GARDIN, J.C. (2002). Les modèles logico-discursifs en archéologie. In: Djindjian, F. and Moscati, P. (eds.). *XIV UISPP Congress. Proceedings of Commission IV Symposia. Data Management and Mathematical Methods in Archaeology. Liège, Belgium, 2001.* Archeologia e Calcolatori. 13. Firenze: All'Insegna del Giglio. p. 19-30. Available from: http://soi.cnr.it/archcalc/indice/PDF13/01Gardin.pdf. [Accessed: 20 July 2014].

GILLINGS, M. and GOODRICK, G.T. (1996). Sensuous and Reflexive GIS: exploring visualisation and VRML. *Internet Archaeology*. 1. Available from: http://intarch.ac.uk/journal/issue1/index.html. [Accessed: 15th July 2014].

HARRIS, T.M., LOCK, G.R. (1995). Towards an evaluation of GIS in European archaeology: the past, present and futures of theory and applications. In: Lock, G.R. and Stancic, Z. (eds.). *Archaeology and Geographical Information Systems. A European perspective.* London: Taylor and Francis. p. 349-365. Available from: http://books.google.it/books?id=ZPTLjMLSc-UC. [Accessed: 17 March 2014].

HIEBEL, G., HANKE, K. (2008). Concept for an Ontology Based Web GIS Information System for HiMAT. In: Jerem, E., Redő, F. and Szeverényi, V. (eds.). *On the Road to Reconstructing the Past. Computer Applications and Quantitative Methods in Archaeology (CAA). Proceedings of the 36th International Conference. Budapest, 2008.* Budapest: Archeaeolingua. p. 243-250. Available from: http://proceedings.caaconference.org/paper/cd32_hiebel_hanke_caa2008/. [Accessed: 16 July 2014].

HOFMANN, E., MANI, C. (2012). De la découverte au SIG: l'exemple de la base terrain du service archéologique de la Ville de Lyon. *Archeologia e Calcolatori*. Supplemento 3. pp. 141-152. Available from: http://soi.cnr.it/archcalc/indice/Suppl_03/AC_Suppl03_2012.pdf. [Accessed: 17 July 2014].

JENSEN, P. (2012). Event-based Archaeological Registration Principles. In: Zhou, M., Romanowska, I., Wu, Z., Xu, P. and Verhagen P. (eds.). *Revive the Past. Computer Applications and Quantitative Methods in Archaeology (CAA). Proceedings of the 39th International Conference. Beijing, China, 2011.* Amsterdam: Pallas Publications. pp. 209-216. Available from: http://proceedings.caaconference.org/paper/23_jensen_caa2011. [Accessed: 12 July 2014].

JOHNSON, I. (2008). Mapping the fourth dimension: a ten year retrospective. *Archeologia e Calcolatori*. 19. pp. 31-44. Available from: http://soi.cnr.it/archcalc/indice/PDF19/3_Johnson.pdf. [Accessed: 10 March 2014].

JUSTRELL, B., FRESA, A. (2014). A Roadmap for Digital Preservation of Cultural Heritage Content. [Draft version 0.4.1 of the handbook published online at DCH-RP Project web site]. Available from: http://www.dch-rp.eu/getFile.php?id=397. [Accessed: 30 October 2014].

KEAY, S., EARL, G. (2013). Portali di Archeologia 1: Progetto Porto e le sue connessioni. In: Serlorenzi, M. and Jovine, I. (eds.). *SITAR. Sistema Informativo Territoriale Archeologico di Roma. Atti del II Convegno. Roma, 2013.* Roma: Iuno Edizioni.

KONDO, Y., MATSUMOTO, G., SEINO, Y., AKO, T., FUKUI, W., SUGIURA, M., UOZU, T., YAMAGUCHI, H. (2012) .Achievements of Archaeo-GIS Workshop 2007-10. In: Zhou, M., Romanowska, I., Wu, Z., Xu, P. and Verhagen, P. (eds.). *Revive the Past. Computer Applications and Quantitative Methods in Archaeology (CAA). Proceedings of the 39th International Conference. Beijing, China, 2011.* Amsterdam: Pallas Publications. p. 334-342. Available from: http://proceedings.caaconference.org/paper/37_kondo_et_al_caa2011/. [Accessed: 12 July 2014].

LANCIANI, R. (1981). *L'antica Roma*. [Republication of Lanciani, R. (1889) *Ancient Rome in the light of recent discoveries.* Boston]. Roma - Bari: Laterza Editori.

LAZZERI, E. (2011). Un modello multi-hazard per la valutazione del rischio dei beni archeologici: il caso sperimentale della Villa Adriana di Tivoli. *Archeologia e Calcolatori.* 22. p. 365-389. Available from: http://soi.cnr.it/archcalc/indice/PDF22/AC_22_Lazzeri.pdf. [Accessed: 12 July 2014].

LELO, K., TRAVAGLINI, M. (2013). Rome in 18th century: a GIS application in the field of urban history. *Historická Geografie.* 39/2. Available from: http://www.researchgate.net/publication/259399379_Rome_in_the_18th_century_a_GIS_application_in_the_field_of_urban_history. [Accessed: 10 March 2014].

MARGIOTTA, U. (2011). Prefazione. In: Banzato, M., *Digital Literacy.* Cultura ed educazione per la società della conoscenza. Milano: Bruno Mondadori.

MIELE, F. (2011) La conservazione della memoria materiale e immateriale. Sistemi informativi di catalogo e territoriali in Campania. *Archeologia e Calcolatori.* 22. Available from: http://soi.cnr.it/archcalc/indice/PDF22/AC_22_Miele.pdf. [Accessed: 15 July 2014].

MICALIZZI, P. , MAGAUDDA, S., BUONORA, P. , SASSO D'ELIA, L. (2011) L'informatizzazione della cartografia storica di Roma. Aspetti urbanistici, archeologici, archivistici e tecnici. In: Serlorenzi, M. (ed.). *SITAR – Sistema Informativo Territoriale Archeologico di Roma. Atti del I Convegno. Roma, 2010.* Roma: Iuno Edizioni. pp. 67-75. Available from: http://beniculturali.academia.edu/ProgettoSITAR/Atti-I-Convegno-SITAR-2010. [Accessed: 15 July 2014].

MOSCATI, P. (1998) GIS applications in italian archaeology. *Archeologia e Calcolatori*. 9. pp. 191-236. Available from: http://soi.cnr.it/archcalc/indice/PDF9/09_10_Moscati.pdf. [Accessed: 05 July 2014].

MOSCATI, P. (ed.) (2009). *La nascita dell'informatica archeologica. Atti del convegno internazionale. Roma 2008.* Firenze: All'insegna del Giglio. Available from: http://www.progettocaere.rm.cnr.it/databasegestione/open_block_pages.asp?IDyear=2009-01-01. [Accessed: 31 June 2014].

NICCOLUCCI, F. (ed.) (2014). *ARIADNE - The Way Forward to Digital Archaeology in Europe.* Firenze: ARIADNE Research Consortium. Available from: http://www.ariadne-infrastructure.eu/.

POZZO, R. (in press). Le iniziative del CNR e il Progetto Nazionale sui Beni Culturali. In: Serlorenzi, M. and Leoni, G. (eds.). *III Convegno SITAR. Il SITAR nella Rete della Ricerca Italiana. Verso la conoscenza archeologica condivisa. Roma, 2013.* Archeologia e Calcolatori. Supplemento [still to be defined]. Firenze: All'Insegna del Giglio.

POZZO R., VIRGILI V. (2013). *Position Paper on Social Sciences, Humanities and Cultural Heritage in Horizon 2020.* Roma: CNR.

SANTORIELLO, A., ROSSI, A., ROSSI, P. (in prep.). SIUrBe 2.0: il Sistema Informativo del patrimonio archeologico Urbano di Benevento tra open data e open source. In: Stanco, F. and Gallo, G. (eds.). *ARCHEOFOSS. Free, Libre and Open Source Software e Open Format nei processi di ricerca archeologica. Atti VIII Workshop.* Catania, 2012. [in preparation]

SCIANNA, A., VILLA, B. (2011) GIS applications in archaeology. *Archeologia e Calcolatori*. 22. Available from: http://soi.cnr.it/archcalc/indice/PDF22/AC_22_Scianna_Villa.pdf. [Accessed: 15 July 2014].

SERLORENZI, M. (ed.) (2011). *SITAR – Sistema Informativo Territoriale Archeologico di Roma. Atti del I Convegno. Roma, 2010.* Roma: Iuno Edizioni. Available from: http://beniculturali.academia.edu/ProgettoSITAR/Atti-I-Convegno-SITAR-2010. [Accessed: 15 July 2014].

SERLORENZI, M. (ed.) (2013). *ARCHEOFOSS. Free, Libre and Open Source Software e Open Format nei processi di ricerca archeologica. Atti del VII Workshop. Roma, 2012. Archeologia e Calcolatori. Supplemento 4.* Firenze: All'Insegna del Giglio. Available from: http://www.progettocaere.rm.cnr.it/databasegestione/open_block_pages_sup. asp?IDyear=2013-01-01. [Accessed: 20 July 2014].

SERLORENZI, M., JOVINE, I. (eds.) (2013). *SITAR. Sistema Informativo Territoriale Archeologico di Roma. Atti del II Convegno. Roma, 2013.* Roma: Iuno Edizioni.

SERLORENZI, M., LEONI, G. (eds.) (2015). *III Convegno SITAR. Il SITAR nella Rete della Ricerca Italiana. Verso la conoscenza archeologica condivisa. Roma, 2013.* Archeologia e Calcolatori. Supplemento [still to be defined]. Firenze: All'Insegna del Giglio.

WILCOCK, J.D. (1973). A general survey of computer applications in archaeology. In: Wilcock, J.D. (ed.) *Computer Applications in Archeology 1. Science and Archaeology. 9.* Stafford: George Street Press. p. 17-21. Available from: http://proceedings.caaconference.org/paper/02_wilcock_caa_1973. [Accessed: 12 July 2014].

WEB SITES

SITAR Project web site: http://sitar.archeoroma.beniculturali.it/

SITAR web pages on Academia.edu platform: https://beniculturali.academia.edu/ProgettoSITAR/

SITAR Knowledge Base: http://sitarmoodle.archeoroma.beniculturali.it/

SITAR Knowledge Base on Academia.edu:

https://beniculturali.academia.edu/SITARProjectKnowledgeExperienceBase/

Art History of the Ancient Near East and Mathematical Models. An Overview

Alessandro DI LUDOVICO[1] and Sergio CAMIZ[2]

[1] Dipartimento di Scienze dell'Antichità, Sapienza Università di Roma, Italy
[2] Dipartimento di Matematica Guido Castelnuovo, Sapienza Università di Roma, Italy
alediludo@gmail.com, sergio.camiz@uniroma1.it

Abstract

The use of mathematical models in the art history of the pre-classic Near East is still comparatively little popular, partly because of cultural as well as technical and logical problems. In the history of research, such kind of approaches have been especially focussed on glyptic productions, but they mainly lack continuity and diffusion in use. Nor is the use of specific models widely spread through scholars originating from different academic institutions. The authors of this contribution review the different uses of quantitative models in Near Eastern art history and offer a summing up and an overview of the underlying approaches and methods. The deep examination of the models and approaches that have appeared up to now in the scientific literature is a first step toward the outline of possible future research courses that could become part of a wide debate on the topic of art history and mathematical models.

Keywords: Art History of ancient Western Asia, Mathematical Models, Figurative Languages, History of Research.

1. Introduction

Beyond the ordinary difficulties and incomprehensions which often affect scholars who make attempts to use quantitative methods in the humanities, the art history of ancient Western Asia traditionally offers further problems. These are partly related to the specific nature of this field of study, and partly connected to the approaches deep-rooted in principal academic institutions.

On the one hand, the figurative languages of pre-classical Western Asia mostly pertain to cultures that are largely unknown and obscure to present-day scholars. A large part of them had already faded into oblivion in ancient times, and very often the documentation is still so scarce that a number of gaps and hard to solve doubts and uncertainties persist nowadays in the cultural basis and achievements of comparatively young disciplines (Sumerology, Assyriology, Archaeology of Western Asia, History of Religion of Western Asia, etc.). On the other hand, a great number of the most traditional approaches of these disciplines either have a Biblical inspiration or are based on perspectives deeply permeated by the logics of Classical studies, alphabet cultures, or similar.[1] Hence, each scholar who prefers, explicitly or not, to follow the *safe way* of adopting so-called *traditional perspectives*, also accepts its logical and cognitive consequences. That the majority of other scholars would agree with such a view only confirms how deeply those habits are rooted. The situation of disciplines related to Classical cultures could be meaningfully different, since they can take advantage of an extended historical continuity in the studies of the products of Classical cultures and of recognisable traces of cultural continuity between Classical cultures and later European ones.

These basic differences can give a particular value to the use of quantitative methods within the disciplines which deal with the ancient Near East, and especially with reference to the field of art history.

2. The problem of coding and the early approaches

For the reasons mentioned above, the relative familiarity of some iconographies or visual themes pertaining to pre-Classical Western Asiatic cultures may be misleading for the scholar, while the effort in looking for a proper coding of those figurative languages in view of their quantitative description or study can disclose interesting and probably fruitful critical courses in the research into their interpretation. To devise a coding system is, in fact, in itself a critical research activity which repeatedly compels the scholar to question and analyse his or her view of the ancient cultures and the relevant products. It is therefore unsurprising that the first innovative uses of quantitative methods in the archaeology and art history of Western Asia originated with the methodological reflections of Jean-Claude Gardin and his colleagues of the Institute Français d'Archéologie at Beirut.

Following theoretical efforts dealing with methods of classification and interpretation and paying great attention to the most recent observations stemming from the international (especially English-speaking) level (Gardin 1958: 335-336; 1967: 13), Gardin developed some proposals for the use of automatic and quantitative methods in the study of various kinds of archaeological artefacts. From the nineteen-fifties, and thoughout his rich career, he thus proposed the use of universal codes which could help scholars in representing and comparing the

[1] Digital culture and the noteworthy qualitative differences between its typical ways of thinking to written texts, and handling them and those of non-digital or oral cultures, or cultures that have other approaches to writing and managing texts should also be remembered here.

content of archaeological reports in a useful manner.[2] His great challenge was to outline a code based on the logics of the *analyse logiciste* in order to describe the relics of ancient cultures fairly precisely in a universal language (see, e.g., Gardin 1958). The same strategies could then also be used to translate the content of archaeological reports into a uniform language (Gardin 2002: 19-21). In Gardin's view, a coding system had to be devised in order to describe systematically all information relevant to a class of artefacts, so that a reduced number of elements could be used in a scientifically satisfying formalised representation of archaeological findings. Of course, each category of relics needed its own code, but the general approach of the logicist was based on assumptions which were fit for universal adoption. Such assumptions were inspired by linguistics, and many explicit linguistic parallels and examples have been recalled by Gardin, while explaining his theoretical perspective (for example, in Gardin 1966; 1967: 18-26).

In Gardin's view, each code had to take into account three features of the object that needed to be described through a coding system: orientation, segmentation and differentiation (Gardin 1967: 13-18). Once defined, these three basic concepts outlined the relations existing between them, Gardin proposed an alphanumeric coding system which could describe the parts which made up the object and the ways in which these parts were related to each other in a concise manner. The material tools used to manage the codes and the coded objects were computer machines and punched cards.

Those efforts and thoughts led to, amongst other results, the development of a computational model in the treatment of ancient Near Eastern figurative languages (first outlines are in Gardin 1967: 21-26), the proposals and first concrete expressions of which concerned the iconographies of glyptic. A number of scholars of French archaeological research institutions located in France and abroad took soon part in the relevant projects, providing them with a considerable specialised scientific support (Moscati 2013: 10-11).

Thus, in the field of art history, Gardin first looked for proper solutions to the descriptive and syntactic problems of the iconography of cylinder seals, experimenting the proper procedures for a formalised representation of the relevant information. The system that resulted was then adopted to carry out the unprecedented *Répertoire Analytique des Cylindres Orientaux* (Digard 1975), a large catalogue of sheets collecting data on cylinder seals that had been published in different journals. Since it was the outcome of the uneven work of the publications of different authors in different periods, this heterogeneous corpus of seals was particularly suitable both to test the methodology and to produce a concrete tool which was ready for use.

[2] The education and research experience of J.-C. Gardin has been described recently and in-depth in Moscati 2013.

The *Répertoire* was designed consistently to summarise in a single publication the general information about those seals and also to provide instruments for automated search and comparisons. It was a model for other initiatives of the same kind dealing with other classes of artefacts and was obviously open to any kind of future enlargement owing to new discoveries. The work was reasonably well-received by specialists, but its authors were already aware that it would have had no follow-up: it was both the turning point and the end of an ambitious project. There were many reasons for this, and some of them had been foreseen years before by Gardin: scholars are often very jealous of their data, and it is rare for a specialist to dedicate his or her time and efforts to build an ordered archive of published materials for the scholars' community.

Gardin's system was largely based on linguistic logics,[3] and it needed to undergo adaptation in order to be applied to a wide range of iconographies. Nevertheless its creator paid great attention to reducing the opportunity for the personal interpretation of iconographies in the coding as much as possible (Gardin 1967). Years before he had expressed an important observation concerning the value and usefulness of such coding systems when he defined their development and use as something very similar to an explicit and conscious version of the usual implicit reasonings of archaeologists looking for typologies.[4]

3. American Experiences since the Seventies

During the nineteen-seventies, a new experience of use of quantitative methods in the study of the iconography of cylinder seals was developed in the United States, at UCLA. A project led there by Marylin Kelly-Buccellati aimed at planning a coding of data related to cylinder seals, and especially to their iconography, in order to allow statistical investigations and cross comparisons of traits or themes at various levels of logics or detail. The coding was based on a number of binary recordings of features chosen to minimise the subjective view and interpretation as much as possible, and to give a set of descriptions of the artefacts that could be not only quite rigorous, but also suitable for computer processing (Kelly-Buccellati 1977).

The digital catalogue obtained in this way would then also include the net of logical relations that had been recognised among those features. The project began within the perspective of the use of statistics as a tool for research in Near Eastern art history (Kelly-Buccellati and Elster 1973; Kelly-Buccellati 1977: 42), and then proposed some initial coding strategies, which included the choice of specific tools, such as photogrammetry for the graphic rendering and preliminary study of each seal.

Following these initial steps, the authors of this research project looked at the study of cylinder seals not just as a

[3] Gardin explicitly compares his coding signs to the phonemes or the graphemes (e.g. in Gardin 1958: 350-355). After all linguistics played a remarkable role in Gardin's multifaceted education (Moscati 2013: 7).
[4] Gardin 1967: 28-29. Gardin kept in considering the opportunity of formally coding also the way archaeologists think and explain their thoughts: see, e.g., Gardin 1997; 2002.

means to understand ancient iconographic cultures, but also with the view to exploring other historical and cultural phenomena, such as the geographic diffusion of cultural traits, social relations or similar phenomena.[5]

Their approach was deliberately different from that which had underpinned the great catalogue prepared by the researchers of the *Institut Française d'Archéologie*: it actually dealt with a specific glyptic production, that of Old Babylonia,[6] and had been thought to be available in the form of a collection of retrievals (coming from computer processes which had already been performed), explaining the types, numbers and the possible contexts of use of the different motifs observed on the seals, rather than being an interactive tool to be used with a computer (Kelly-Buccellati 1979-1980).

That interesting and ambitious project was part of a wider research program looking for a systematic examination through Information Technology of ancient Mesopotamian cultural products, including historical and linguistic traits and relics. The larger project was entitled 'Computer Aided Analysis of Mesopotamian Material', and the series of publications *Cybernetica Mesopotamica* had been prepared in order to publish from time to time the relevant outcomes (Kelly-Buccellati 1977: 45, fn. 2). As far as we know, the various branches of that program, including that concerning Old Babylonian glyptic, were halted before time: four issues of *Cybernetica Mesopotamica* were published[7], the last of which was in 1984, and all dealt with topics relating to the epigraphic and philologic study of cuneiform. One can thus imagine that the part of the project dedicated to the quantitative study of ancient art ended at the beginning of the nineteen-eighties. Perhaps very close to that project is a further experiment that was proposed in more recent times, but it probably had no actual developments.[8]

4. Glyptic Iconographies and Applied Mathematical Models

In the early nineteen-nineties a new research strategy on glyptic figurative languages was devised at the University of Rome 'La Sapienza' during a cooperation between Elena Rova and Sergio Camiz (Rova 1994). In this case, the chronological and geographical scale was reduced in comparison to that of the experiment of the French logicist school described above, while the aims, the type of mathematical models, and their use, were very different; however, the basic logical references for devising a proper coding strategy had quite a lot in common (Rova 1995). On the other hand, with the approach proposed by Kelly-Buccellati this experience shares the limitation of the corpus to a single cultural-historical period, but not the targets and just very few of the basic principles.

Elena Rova investigated with Sergio Camiz the glyptic iconography of the early historical periods of Lower Mesopotamia, with the aim of locating compositional patterns, as well as geographical or diachronic variation features, looking for connections between 'specific iconographic and compositional features' and geographical or chronological variation, and outlining thematic classification.[9]

The tools used in this research are Textual Correspondence Analysis, Multiple Correspondence Analysis, Hierarchical Ascendant Classification, and also Principal Component Analysis. The iconographies were coded according to three types of description: textual, symbolic and presence/absence (Camiz and Rova 2001; 2003; Camiz et al 2003; 1998). The textual coding was the most complex and was investigated on three main levels: distinct icons, sub-patterns of the composition - that is small sets of elements - and whole image syntax.

In the case of this research project the use of quantitative methods and logics is thus embedded in specific ways of investigating figurative languages: it is a tool for searching for answers to specific scientific questions about the glyptic iconography of a specific historical period. This means that in this case the model interacts more dynamically with the data set, the structuring of which can be well adapted to scientific requirements before being formally adjusted according to the adopted algorithms. Furthermore, here the aim is the interpretation of a figurative language and the reconstruction of its functioning. It is an experimental piece of research, both in its methodology and in its pursued results, rather than a formalised systematisation of materials or a tool for other researchers. The procedure adopted can, in fact, be correctly used as a tool for other investigations only following a proper adaptation of the specific needs of each item of research and of the characteristics of the studied corpus.

Similar principles inspired the research project on presentation scenes in third millennium Mesopotamian glyptic which have been developed since the early 2000s by Alessandro Di Ludovico. It was originally a partly analytical and partly qualitative study aiming at outlining the historical developments of the theme of presentation in the glyptic of Lower Mesopotamia from the Akkadian period to the end of Ur III. The first approaches through which the research was undertaken were mainly inspired by structural linguistics, and led to an initial analogue classification and a genealogy of the iconographic traits

[5] Kelly-Buccellati and Elster 1973: 199-200. This also provides the reasons for which seals are especially important in a quantative study in the history of art of the Ancient Near East: the *corpus* is wide, and thus 'quantifiable', the basic stylistic chronology is known, many of these can be dated with a certain precision, thanks to the imprints on tablets and references in the inscriptions, their publication is usually 'fuller than other classes of artefacts'.

[6] Old-Babylonian glyptic was chosen for the wide range of material datable with a certain precision. (Kelly-Buccellati 1977, p. 45).

[7] See Undena Publications Website: Data Sets - Cuneiform Texts, URL: http://128.97.6.202/up/dsc.html; Graphemic Categorization, URL: http://128.97.6.202/up/gc.html.

[8] F. Buccellati, Constructing Art Historical Definitions through a Comparative Database: The Evidence of Old Assyrian Glyptics, oral paper presented at the 36th CAA, Budapest 2008.

[9] Camiz and Rova 2001. It is important to remember here the theoretical contribution provided by Sergio Camiz along this research course (Camiz 2004) where he expressed his view as a mathematician on the coding of archaeological finds.

and the compositions of the scenes (the whole work has been summarised in Di Ludovico 2005). The fuzzy alphanumeric codings were later translated into binary, so that investigations through Artificial Neural Network algorithms could be carried out (Di Ludovico and Ramazzotti 2008). The translation also involved some adaptations of the coding logics which were further refined in the following experiments, the data set of which was largely renewed and centred only on the presentations of the Ur III period (Di Ludovico 2011; Di Ludovico and Pieri 2011a). In parallel to this, further experiments carried out with Giovanni Pieri and the testing of the linguistic logics of the compositional patterns of presentation scenes were performed, through which it was possible to develop machine automated simulations of the structuring of such scenes, which could be profitably used for a graphic rendering of the results of the Artificial Neural Network processing (Di Ludovico and Pieri 2011b).

These investigations allowed a better outlining of the inner relations existing through the compositional patterns, and gave clues for the interpretation of some motifs and their historical developments. Further contributions to these interpretation processes then came from the cooperation with Sergio Camiz, in which the Textual Correspondence Anaysis and the Hierarchical Ascendant Classification, which had been profitably used by Rova and Camiz, were employed (Di Ludovico et al 2013; Di Ludovico and Camiz in press). This part of the research project on Ur III presentation scenes is still in progress, and is producing inspiring results.

These research courses refer to a more limited historical and cultural context than that of Rova's works, and the corpus on which they focus is much more homogeneous. However, besides sharing some of the basic principles and some applied methodologies, the two projects follow partly different epistemological approaches; for instance the more recent one is much more heavily based on the topological relations between the compositions and the peculiar development of the cylindrical surface that serves as the field on which they are depicted.

In recent times digital technologies for image acquisition and handling have been dynamically employed with cylinder seals both by an Italian-French international team (Pitzalis et al 2008) and by Paul Boon and Martine de Vries-Melein (2013; both scholars work for Dutch institutions). Both experiences are of great interest because they can provide a remarkable help in the documentation and study of seals in a perspective which is more respectful of their basic physical features and shape. The two different types of three-dimensional modelling that these experiments propose can thus document well the information that is essential for quantitative studies of the iconography of seals, and deserve to be applied on a large scale.

5. Iron Age Ivories and Machine Learning

Outside the realm of cylinder seals, a very interesting use of quantitative methods for research in iconography and artisan traditions is that developed by Amy R. Gansell, with the technical contribution of some specialists of disciplines relating to natural sciences (Gansell et al 2007; Gansell et al 2014). The subject of this research course is a part of the famous corpus of ivory carvings of Levantine production which were found in the excavation of different Neo-Assyrian settlements. The project is still in progress, but has already provided useful results. In the investigation a classification of the ivory fragments through machine learning is attempted, so that, on one hand, the past hypotheses and interpretations of the corpus - mainly based on stylistic analysis - could be critically examined and, on the other hand, new proposals and observations on the corpus itself could be developed. Through the adopted algorithms, a number of features which can be decisive for outlining the classes were identified (Gansell et al 2014: 202). The methodology is of statistical kind, and the basic logics are not far removed from that used for some of the studies on early historic seals by Rova and Camiz, however the whole process and the coding distinguish themselves for paying special attention to metrological features, thus giving an important role to both the artisan's perception and the handling-related characteristics of the artefacts.

The starting point of this investigation was a long and well-established tradition of art historic analyses, documentation, and classification of the corpus. Statistical tools and reasoning were here used to locate the features of resemblance or dissimilarity between classes of artefacts, and to test classifications or trace new classification proposals and relations. The investigation of the cultural (or spatio-temporal) network relations and the enlargement of the corpus to other materials are two basic proposals that will characterise the continuation of this project (Gansell et al 2014: 203), that promises further fruitful outcomes.

6. Conclusion

Research in the art history of ancient Western Asia has been only very slightly and irregularly affected by the use of quantitative and computer-aided methods. Most experiences of this kind derived from the initiative of single scholars or small groups of scholars, who were differently fostered by the institutions to which they belonged. Until now, the most systematic and officially supported experience seems to be that developed and carried out by J.-C. Gardin within French institutions.

This notwithstanding, the results that have been obtained through the different research experiences with which we dealt in this paper lead us to take an optimistic view on possible further developments, both in the field of methodologies closely embedded in research courses, and in that of quantitative approaches to documentation and representation. As explained in the introductory notes of this contribution, the field of Western Asiatic art history not only requires coding and description strategies that must deal with very complex objects; it also mostly deals with products of cultures that are quite distant from those pertaining, or somehow related to, the perceptive and linguistic spheres of most scholars, and as such is particularly exposed to the risk of symmetrisation in the interpretation process (Di Ludovico in press). The experiences with quantitative methods in this specific context give particularly meaningful clues which reveal that the two problems of documentation and (more or less experimental) research are not really distinct, nor complementary: they should be rather thought as integrated parts of the same process.

Of course, as J.-C. Gardin observed many years ago, the problem of involving scholars in actively sharing their efforts to collect data and allowing them to be available for the whole scientific community can hardly find a true solution. Public institutions, at least, should help researchers by giving them free access to their archives of published artefacts (which would also help in preserving their archives), simplifying the bureaucratic steps needed to use technologies in collecting data by the artefacts, and, in general, affirming and promoting the logics of free access.

From the scientific point of view, the basic problems that emerge from an analysis of the experiences with quantitative methods in art history have to do with the coding techniques and logics, all strictly related to the scholar's perception of the artefact. As such they require the scholar to reflect critically on the categories and values that are implicit in his or her way of observing it. In Western Asiatic art history quantitative approaches are still used very little, in comparison with the great potential that their employment showed in the past, but perhaps the reduced costs of digital tools and logistical support could help the development of new ambitious projects, of the kind attempted in the earliest approaches, in the near future.

Bibliography

BOON, P. and de VRIES-MELEIN, M. (2013) Cylinder Seals Revealed. In: Contreras, F., Farjas, M. and Melero, F.J. (eds.). *Fusion of Cultures. Proceedings of the 38th Annual Conference on Computer Applications and Quantitative Methods in Archaeology, Granada, April 6-9 2010*. British Archaeological Reports – International Series 2494. Oxford: Archaeopress. p. 511-517.

CAMIZ, S. (2004) On the Coding of Archaeological Data. *Archeologia e Calcolatori*. 15. p. 201-218.

CAMIZ, S. and ROVA, E. (2001) Exploratory Analyses of Structured Images: A Test on Different Coding Procedures and Analysis Methods. *Archeologia e Calcolatori*. 12. p. 7-45.

CAMIZ, S. and ROVA, E. (2003) Quantitative Study of Images in Archaeology: I. Textual Coding. In: Schader, M., Gaul, W. and Vichi, M. (eds.). *Between Data Science And Applied Data Analysis*. Berlin: Springer. p. 624-632.

CAMIZ, S., ROVA, E. and TULLI, V. (1998) Exploratory Analysis of Images engraved on Ancient Near-Eastern Seals based on a Distance among Strings. *Statistica*. 58. p. 670-689.

CAMIZ, S., ROVA, E. and TULLI, V. (2003) Quantitative Study of Images in Archaeology: II. Symbolic Coding. In: Schader, M., Gaul, W. and Vichi, M. (eds.). *Between Data Science And Applied Data Analysis*. Berlin: Springer. p. 633-641.

DIGARD, F. (ed.). (1975) *Répertoire analytique des cylindres orientaux*. Paris: Editions du CNRS.

DI LUDOVICO, A. (2005) Scene-in-frammenti: una proposta di analisi delle 'scene di presentazione' dei sigilli a cilindro mesopotamici orientata all'elaborazione statistica ed informatica dei dati. In: Di Ludovico, A. and Nadali, D. (eds.). *Studi in onore di Paolo Matthiae presentati in occasione del suo sessantacinquesimo compleanno*, Contributi e Materiali di Archeologia Orientale. X – Numero Speciale. Roma: Sapienza. p. 57-95.

DI LUDOVICO, A. (2011) Experimental Approaches to Glyptic Art Using Artificial Neural Networks. An Investigation into the Ur III Iconological Context. In: Jerem, E., Redő, F. and Szeverényi, V. (eds.). *On the Road to Reconstructing the Past, Proceedings of the 36th Annual Conference on Computer Applications and Quantitative Methods in Archaeology, Budapest, April 2-6 2008*, Budapest: Archaeolingua. p. 135-146.

DI LUDOVICO, A. (in press) A Look into Late Third Millennium Mesopotamian History. Supported by Cross-disciplinary Interpretive Tools. In: *Proceedings of the Workshop on Social Theory in Archaeology held at the University of Warsaw on May 1st 2012*.

DI LUDOVICO, A. and CAMIZ, S. (in press) A Quantitative Approach to Ur III Mesopotamian Figurative Languages: Reflections, Results, and New Proposals. *Archeologia e Calcolatori*. 25.

DI LUDOVICO, A., CAMIZ, S. and PIERI, G. (2013) Comparative Use of Mathematical Models in an Investigation on Mesopotamian Cylinder Seals. In Contreras, F., Farjas, M. and Melero, F.J. (eds.). *Fusion of Cultures. Proceedings of the 38th Annual Conference on Computer Applications and Quantitative Methods in Archaeology, Granada, April 6-9 2010*. British Archaeological Reports – International Series 2494. Oxford: Archaeopress. p. 495-498.

DI LUDOVICO, A. and PIERI, G. (2011a) Artificial Neural Networks and Ancient Artefacts: Justifications for a Multiform Integrated Approach Using PST and Auto-CM Models. *Archeologia e Calcolatori*. 22. p. 99-128.

DI LUDOVICO, A. and PIERI, G. (2011b) How to Facilitate Interpretation of Natural Computation Results by Converting Binary Codes of Images back to Images. *CiiT International Journal of Artificial Intelligent Systems and Machine Learning*. 3 (7). p. 437-446.

DI LUDOVICO, A. and RAMAZZOTTI, M. (2008) Reconstructing Lexicography in Glyptic Art: Structural Relations between the Akkadian age and the Ur III Period. In: Biggs, R.D., Myers, J. and Roth, M. (eds.).

Proceedings of the 51st Rencontre Assyriologique Internationale, Held at the Oriental Institute of the University of Chicago, July 18-22, 2005. Studies in Ancient Oriental Civilization 62. Chicago: Oriental Institute. p. 263-280.

GANSELL, A.R., TAMARU, I.K., JAKULIN, A. and WIGGINS, Ch.H. (2007) Predicting regional classification of Levantine ivory sculptures: a machine learning approach. In: Clark, J.T., Hagemeister, E.M. (eds.). *Digital Discovery: Exploring New Frontiers in Human Heritage, CAA 2006, Computer Applications and Quantitative Methods in Archaeology, Proceedings of the 34th Conference*, *Fargo, United States, April 2006*. Archaeolingua: Budapest. p. 369-378.

GANSELL, A.R., van de MEENT, J.-W., ZAIRIS, S. and WIGGINS, Ch.H. (2014) Stylistic Clusters and the Syrian/South Syrian Tradition of First-Millennium BCE Levantine Ivory Carving: A Machine Learning Approach. *Journal of Archaeological Science*. 44. p. 194-205.

GARDIN, J.-C. (1958) Four Codes for the Description of Artifacts: An Essay in Archeological Technique and Theory. *American Anthropologist*. 60 (2). p. 335-357.

GARDIN, J.-C. (1966) Éléments d'un modèle pour la description des lexiques documentaires. *Bulletin des bibliothèques de France*. 5. p. 171-182.

GARDIN, J.-C. (1967) Methods for the Descriptive Analysis of Archaeological Material. *American Antiquity*. 32. p. 13-30.

GARDIN, J.-C. (1997) Le questionnement logiciste et les conflits d'interprétation. *Enquête*. 5. p. 35-54.

GARDIN, J.-C. (2002) Les modèles logico-discursif en Archéologie. *Archeologia e Calcolatori*. 13. p. 19-30.

KELLY-BUCCELLATI, M. (1977) Towards the Use of Quantitative Analysis in Mesopotamian Sphragistics, *Mesopotamia* 12: 41-52.

KELLY-BUCCELLATI, M. (1979-1980) Iconographic and Formal Analysis of Mesopotamian Cylinder Seals: System Rescription, *Colloque International d'études ugaritiques à l'occasion du Cinquantenaire des fouilles archéologiques à Ras Shamra (UGARIT), Lattaquié 10-13 octobre 1979*. Annales Archéologiques Arabes Syriennes. Revue d'archéologie et d'histoire. 29-30 – Numéro special. Damascus. p. 273-284.

KELLY-BUCCELLATI, M. and ELSTER, E. (1973) Statistics in Archaeology and its Application to Ancient Near Eastern Data. In: Buccellati, G. (ed.). *Gelb Volume: Approaches to the Study of the Ancient Near East*. Orientalia. 42 – Special Issue. Rome: Pontificium Institutum Biblicum. p. 195-211.

MOSCATI, P. (2013) Jean-Claude Gardin (Parigi 1925-2013). Dalla meccanografica all'informatica archeologica. *Archeologia e Calcolatori*. 24. p. 7-24.

PITZALIS, D., CIGNONI, P., MENU, M. and AITKEN, G. (2008) 3D Enhanced Model from Multiple Data Sources for the Analysis of the Cylinder Seal of Ibni-Sharrum. In: Ashley, M., Hermon, S., Proença, A. and Rodriguez-Echavarria, K. (eds.). *VAST 2008: The 9th International Symposium on Virtual Reality, Archaeology and Intelligent Cultural Heritage, Braga, Portugal, 2008*. Aire-la-Ville: Eurographics Association. p. 79-84.

ROVA, E. (1994) *Ricerche sui sigilli a cilindro vicino-orientali del periodo di Uruk/Jemdet Nasr*. Orientis Antiqui Collectio 20. Roma: Istituto C. A. Nallino.

ROVA, E. (1995) Metodi di codifica e analisi per lo studio di immagini strutturate: un'applicazione ai sigilli cilindrici del periodo di Uruk/Jemdet Nasr. *Archeologia e Calcolatori*. 6. p. 7-34.

Undena Publications. Cybernetica Mesopotamica. DSC: Data Sets - Cuneiform Texts. [Online] http://128.97.6.202/up/dsc.html. [Accessed: 14th April 2014].

Undena Publications. Cybernetica Mesopotamica. GC: Graphemic Categorization. [Online] http://128.97.6.202/up/gc.html. [Accessed: 14th April 2014].

Archaeology and Computer Applications: The Automatic Cataloging of Italian Archaeological Heritage

Alessandra CARAVALE

Istituto di Studi sul Mediterraneo Antico (ISMA) - Consiglio Nazionale delle Ricerche (CNR), Italy
alessandra.caravale@isma.cnr.it

Abstract

This paper is centered on the birth and early development of the automatic cataloging of archaeological heritage, with particular reference at the Seventies and Eighties that marked two very important decades in this sector. In fact, in these years in Italy as well as in other European countries, a lively debate aroused on this theme and set the basis for the development of the instruments and for the procedures and the techniques to be adopted. In this paper we will refer in particular to the Italian situation, where the introduction of computers in documentation procedures mostly pivot around two pioneering scholars: Oreste Ferrari and Paola Barocchi. Some reference will be made also to the situation in France and Britain. Some recent examples of ISMA on-line projects, such as the database Sethlans. Bronzi del Museo Faina, will also be examined

Keywords: automatic cataloging, heritage, archaeology

The cataloging of an archaeological item means to study and know it appropriately, by determining its rightful place in space and time for the purpose of its proper conservation and enhancement. The cataloging of an archaeological item also means to give further information and details about its relationship with other items or disciplines. The use of computer science for cataloging has enabled us not only to have a rapid and effective means of research, but also to check and analyze data that would otherwise be difficult to achieve when using traditional paper cards (Barocchi & Fileti Mazza, 2000; Moscati, 2002; Corti, 2003; Gamba, 2007; Caravale, 2009). The first computer experimentations of cataloging cultural heritage date back to the Sixties in most countries of the world. These kind of experimentations developed along with other more important projects of inventorying national heritage which were organized by public institutions in charge of documentation and preservation. The introduction of computer technology brought about problems linked to automatic techniques of data processing and language standardization. In other words, how to describe similar objects homogeneusly, so as to be able to perform search in databases created to catalogue archaeological items or monuments and how to retrieve that information correctly.

The first solutions to ensure uniformity of description and standardize the language were oriented to the creation of analytical 'codes' designed to describe various antique objects. In France, Jean-Claude Gardin devoted his pioneering work to this purpose (Moscati, 2002; 2013). Controlled vocabularies or thesauri, containing lists of terms or 'keywords', were then created. Later on, the general trend was to define different files for the various categories of goods, as well as the so-called 'standard catalogue'. All these tools were intended to support and control the rules and methodological guidelines to be followed in order to acquire knowledge on the goods and produce documentation relating to them, according to homogeneous and shared criteria (D'Andrea, 2006).

1. The Seventies and the Eighties

The Seventies and the Eighties marked two very important decades in the development of automated cataloging. In these years in Italy as well as in other European countries, such as France and Britain, a lively debate aroused on this theme and set the basis for the development of the instruments and for the procedures and the techniques to be adopted. In this paper we will refer in particular to the Italian situation, where the introduction of computers in documentation procedures mostly pivot around two pioneering scholars: Oreste Ferrari and Paola Barocchi. Some reference will be made also to the situation in France and Britain (Caravale, 2009).

In Italy an important contribution in this sector was given by the Istituto Centrale per il Catalogo e la Documentazione established in 1975 together with the Ministry of Cultural Heritage (http://iccd.beniculturali.it/index.php?it/95/istituto-centrale-per-il-catalogo-e-la-documentazione/). This Institute was managed from its establishment until 1990 by Oreste Ferrari (Fig. 1), an art historian who brought forward the first phase of national cataloging with enthusiasm and great determination (Gamba, 2007). He stated that the major purpose of cataloging was to ensure an effective protection of the heritage by means of its detailed knowledge, in order to provide a valid cognitive instrument of the national territory for a proper territorial planning. The catalogue was therefore seen as an essential knowledge base to defend and better appreciate the true value of the Italian cultural heritage to be able to promote a sustainable development of national economy (Ferrari, 1972; 1975; 1979). Apart from promoting the catalogue on a large scale, Oreste Ferrari also sustained the computerized cataloging from its initial stages and followed the difficult transition from paper record cards to the automated ones. Since the early years of his leadership of the Institute, he strongly supported the need for a correct use of computers to achieve a more detailed cataloging of cultural items. The major problems were connected to the definition of the methods of a formalized cataloging, accounting for all the variety and complexity of information that each individual item might contain. This method was also intended to ensure the correctness and the terminological homogeneity in particular, thus allowing the Information Retrieval in very large archives (Ferrari, 1979; 1989; 1991; Papaldo & Ruggeri, 1993).

Figure 1: Oreste Ferrari (from Gamba, 2007).

Since the early Seventies, the Institute had established some models of type-written cards, which were organized as descriptive documents with information on items, also including photographic documentation. ‘In many cases, the approach taken to storing catalogue information was very similar to the one used by librarians. The basic idea was to describe objects with ‘cataloging cards’ where information was organized in several semantically consistent sections, describing, for example author, period, excavation data, subject, historical and critical notes. The first organisation of the Italian Catalogue was based on a manual approach, where each object was described by a typewritten card. The basic ideas were very valid and all subsequent work has been greatly influenced by the intellectual efforts that led to the definition of the fundamental principles of the cataloging rules. The most important issues were: identification of a reduced set of different cards, corresponding to different types of objects (art objects, archaeological objects, drawings, architecture, gardens, historical centres, etc.); grouping of the information in several very general categories, like author, location, material, historical info, etc.; topological arrangement of the catalogue cards’ (Signore, 2009, p. 112). For the purpose of adapting the paper record cards to allow computer processing of the data and in order to avoid vagueness and non homogeneity in the terminology, the Institute began to provide controlled vocabularies for some categories of items. The first one for the archeological area was related to the materials of the Final Bronze Age and the early Iron Age of 1980 (Parise Badoni, 1980). Then, the Institute started publishing new manuals providing guidance for card compilation. In particular a manual for the stratigraphic excavation was published in 1984, followed by those for the archaeological and art-historical goods and for archaeological buildings (Parise Badoni & Ruggeri Giove, 1984; 1988). In the same years, the Institute also created and distributed to the Italian Superintendences a system of guided and controlled data entry, called SAXA (SAXA, 1988). In the early Nineties it was replaced by a software called DESC (Lavecchia & Poggi, 1992). During the first XXI century ICCD concluded the General Cataloging Information System (SIGEC), a project focused on making possible the integrated management of the different information (alphanumeric, multimedia, geographic) available on the national heritage (Mancinelli, 2004); in 2009 it was released as SIGECweb (http://www.iccd.beniculturali.it/index.php?it/118/sistema-informativo-generale-del-catalogo-sigec/).

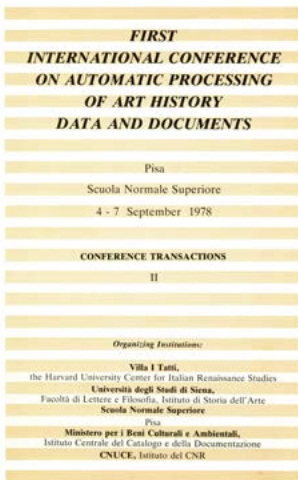


Figure 2: First International Conference on Automatic Processing of Art History Data and Documents, Pisa, 4-7 September 1978: cover.

In the Seventies, the Scuola Normale Superiore in Pisa was also interested in the creation of glossaries and thesauri for the art-historical sector, in collaboration with the Accademia della Crusca and the John Paul Getty Trust in Los Angeles (Parra, 1993; Vesentini, 2009). The research was also centered on cataloging, data management and automated processing of sources and documents. Another art historian, Paola Barocchi, professor at the Scuola Normale Superiore, led this work and also managed the Computer laboratory and its ‘Bollettino d’Informazioni del Centro di Ricerche Informatiche per i Beni Culturali’. In the Eighties the journal published articles on the various projects of archiving and cataloging produced in that period both in the field of archeology and in the art-historical sector. In 1978 the First International Conference on Automatic Processing of Art History Data and Documents was held in Pisa, with a section devoted to archaeology (Fig. 2) (Barocchi & Bisogni, 1978). Here some international programs were presented, some of them just experimental, which gave rise to a large discussion on computerized cataloging, favored by the widespread use of personal computers. In 1984 another meeting followed in Pisa, still focusing on the relationship between computing and cultural heritage (Corti, 1984).

In the late Eighties, the Italian government allocated approximately six hundred billion old liras to finance archaeological computing projects (Rapporto sui progetti, 1989). They aimed at enhancing cultural heritage through the use of advanced technologies as well as the employment of young people in this sector. These projects involved large IT companies, such as IBM Italia, that for the first time collaborated in the field of cultural heritage. Unfortunately, the evaluation of these initiatives was not positive since they came into conflict with the others brought forward simultaneously by the Ministry. First of all, problems concerned the types of projects undertaken, that often went on regardless of the state of progress of the cataloging in the various departments and they were deprived of a broad and general vision required for a correct management of cultural heritage. Furthermore, they were limited to just a few territories and certain types of items. As a consequence they disregarded the correlation and historical contextualization between the various items that distinguished the catalogue from the inventory. Nonetheless, this experience was useful to highlight some common issues related to computerized cataloging, which were later dealt with, like the one connected to the use of mutually compatible and consistent systems.

2. France and Great Britain

In Europe other countries had contributed in this sector before Italy. For example, in England, the roots of the English Heritage Archive date back to 1908 with the foundation of the Royal Commission on the Historical Monuments of England (RCHME). Its approach was topographical. The Commission was set up to compile and publish an inventory of the ancient and historical monuments, by county and by parish, constructed from ancient age until the Eighteenth century. The activities of the Royal Commission continued during the Twentieth century and included the National Buildings Record (NBR) and National Archaeological Record (NAR). In 1999 the Royal Commission merged with English Heritage to form a new body for the historic environment. Within a context transformed by the rise of digital media and the Internet, English Heritage (http://www.english-heritage.org.uk/) still carries on its historic role of creating a record of the historic environment; making that records available to all those interested in the history around them; and preserving the record for future generations (Aberg & Leech, 1992). Now the record of England's archaeological and architectural sites contains over 400,000 records. The data set provides basic information about each site together with sources, archive and activity details as appropriate. The data set is compiled to clearly defined standards and is compatible with MIDAS (the Manual and Data Standard for Monument Inventories) and uses nationally recognized terminology standards.

More recently, in the archaeological field, the Archaeology Data Service (ADS, http://archaeologydataservice.ac.uk/), that was established in September 1996, has strongly contributed to the creation and maintenance of digital archives. The ADS is led by the University of York. Its mission is to support 'research, learning and teaching with high quality and dependable digital resources', by maintaining 'digital data in the long term, and by promoting and disseminating a broad range of data in archaeology'. This service also 'promotes good practice in the use of digital data in archaeology, provides technical advice to the research community, and supports the deployment of digital technologies'. The Archaeology Data Service web-site offers some useful databases, like for instance the one on Roman amphorae.

The problems related to language standardization and data structuring were also dealt with in France, where the research by Jean-Claude Gardin and Rene Ginouvès gave a strong impetus to the development of the automated analysis of documentary character.

L'Inventaire général des monuments et richesses artistiques de la France (http://www.inventaire.culture.gouv.fr/) was set up in 1964. Its establishment was promoted by the minister of culture Andre Malraux and by André Chastel, with the objective of reviewing, studying and getting to know all the historical, artistic and archaeological works of art that are part of the national heritage. It responded to the need of better knowing the national heritage of the country through a systematic analysis of all the elements that witnessed the art history of a territory, from the smallest to largest ones. This work was carried out on a topographic bases and by regional commissions, but following principles developed at a national level (de Massary & Coste, 2007; Meyer, 2008). The need for the standardization of information emerged because of the huge number of monuments being recorded, and it proved to be the only instrument allowing correct statistical analysis and homogeneity of documentation. The first results of the normalized treatment of data were presented in the early Seventies, earlier than in Italy. Various controlled vocabularies for databases were prepared before the end of the decade, together with software specifics. In that period some of the data collected was already published. Gardin collaborated in the implementation of controlled vocabularies with his Center d'Analyse Documentaire pour l'Archeologie (CADA) (Moscati, 2013, p. 17-19). Since 1995 the French Inventory has been available on the Ministry's web-site.

3. New ISMA projects

Over time, the development of information technologies has had a significant influence on the methodologies of cataloging cultural heritage, causing the evolution from databases to multimedia systems and then enlarging the objective of these instruments from cataloging to disclosure. In particular since the second half of the Nineties, the Internet network has become an important environment for consulting and sharing knowledge, to facilitate the work and the integration of bodies in the documentation and the protection of cultural heritage, but also to be used by a wider audience. In this direction, focusing on computerized cataloging and open access is also guiding my own research at the Istituto di Studi sul Mediterraneo Antico (ISMA) of the Italian National Research Council (CNR), which is focused on an on-line project of automated cataloging. This project concerns the bronze collection of Museo Claudio Faina in Orvieto (Figs. 3-4).


Figure 3: Faina collection: a votive bronze (© Museo).
Figure 4: Faina collection: a bronze situla (© Museo).

The Faina collection was started in 1864 by to two important members of the family: the counts Mauro and Eugenio. It is believed that the initial collection was made up of 34 vases donated to the count Mauro by princess Maria Bonaparte Valentini, Napoleon's niece and daughter of Luciano Bonaparte, who discovered Vulci necropolis. Mauro Faina was responsible for the collection until 1868, when, after his death it was inherited by his brother Claudio and subsequently given to his nephew Eugenio. The collection was initially kept in the family residence in Perugia and then transferred to its current home in Orvieto. Eugenio started to become interested in antiquities that at that time were being found during the excavations in orvietan necropolis. Inherited by Claudio junior, the collection, at that point complete, was made open to the public in 1954.

There are around one thousand different bronze objects in the this rich collection, which date from the Bronze Age up the Roman period. The majority of these objects are votive bronzes and vases (Caravale, 2003; 2006). There are also mirrors, figured applique, candelabras, thymiateria, lamps, weapons and small tools for female care and ornament. As with other objects of the collection, there is no precise indications as to where these bronze objects originated. Some data, however, can be found in some bibliographic sources and archive documents analyzed by B. Klakowicz (1970). In the bronze collection, however, it is possible to distinguish between those collected by Mauro Faina from 1864 to 1868 from the areas of Orvieto, Chiusi, Perugia, Todi and Bolsena, and those collected by Eugenio, originated exclusively from orvietan excavations conducted between 1869 and 1881.

The bronzes Faina database makes use of the Content Management System open source Museo & Web (Natale & Saccoccio, 2010). This is a system that was created by the Technological Observatory for Heritage and Cultural Activities (OTEBAC) of Italian Ministry for Cultural Heritage and Activities and Tourism (MIBACT) in order to develop and manage high quality web-sites devoted to museums or cultural institutions. This system facilitates the creation of a database of the objects kept in museums and makes use of metadata for the retrieval and access management of digital resources. 'The importance of this CMS, characterised by modules especially planned for cultural institutions, is that it is not imposed by third parties, but is designed with the contributions of the cultural institutions which participate in the enrichment of the platform by expressing their needs. The diffusion of the kit (which also includes a series of guidelines on how to build the architecture of the web-sites of cultural institutions) contribute to increase awareness among the stakeholders of cultural institutions dealing with communication and web publishing on accessibility, usability and quality of cultural website in general' (Natale & Saccoccio, 2010, p. 47).

The Faina web-site is named 'Sethlans. Bronzi del Museo Faina' (http://bronzifaina.isma.cnr.it/) (Fig. 5). It is organized with some general pages devoted to the history of the Orvietan collection; it also includes some more detailed pages dedicated to the most important bronze items of the collection with links to other databases in the web (for example: the archaeological database of Soprintendenza per i Beni Archeologici dell'Umbria, http://www.archeopg.arti.beniculturali.it/index.php?it/157/banca-dati-beni-archeologici; or The Metropolitan Museum collection on-line, http://www.metmuseum.org/collection/the-collection-online/).

Recently, an exhibition at the Museo Claudio Faina ('Sethlans. I bronzi etruschi e romani nella collezione Faina') (Fig. 6) was developed around our database, in which objects are implemented with a brief description of the finds and a special attention to their origin and chronology. From April to July 2014 the web-site counter has detected more than 250 hits; by analyzing the accesses' areas of origin, we observe that there is a good interest from Europe (Italy, Spain, France, Germany, Russia) and United States.

The aim today is not only the cataloging of a single item, but also its relationship with the cultural context of reference. From an IT point of view there is a tendency not so much to define uniform standards, but rather to develop the interoperability between different systems, fundamental for the growth of information in a wider dimension. The key objective is to create on-line archives of data, in order to enable the scholars of art and archaeology to exploit new technologies for their own research and for the exchange of information at an international level.

FIGURE 5: SETHLANS. BRONZI DEL MUSEO FAINA WEB-SITE: HOMEPAGE.

FIGURE 6: 'SETHLANS. I BRONZI ETRUSCHI E ROMANI NELLA COLLEZIONE FAINA': EXHIBITION POSTER.

References

ABERG, F.A. & LEECH, R.H. (1992). The National Archaeological Record for England. Past, Present and Future. In: *Sites and Monuments. National Archaeological Records*. Kobenhavn: The National Museum of Denmark, pp. 157-170.

BAROCCHI, P. & BISOGNI, F. (eds.) (1978). *First International Conference on Automatic Processing of Art History Data and Documents, Pisa, 4-7 September 1978*, Pisa.

BAROCCHI, P. & FILETI MAZZA, M. (2000). Beni culturali e ambientali. Beni culturali e informatica. In: *Enciclopedia Italiana, Appendice 2000*. Volume 1. Roma: Istituto della Enciclopedia Italiana Treccani, pp. 177-181.

CARAVALE, A. (2003). *Museo Claudio Faina di Orvieto. Bronzetti votivi.* Electa: Editori Umbri Associati.

CARAVALE, A. (2006). *Museo Claudio Faina di Orvieto. Vasellame.* Electa: Editori Umbri Associati.

CARAVALE, A. (2009). La catalogazione informatica del patrimonio archeologico. In: MOSCATI, P. (ed.) (2009) *La nascita dell'informatica archeologica. Atti del Convegno Internazionale, Roma, Accademia Nazionale dei Lincei, 24 ottobre 2008. Archeologia e Calcolatori* 20, pp. 179-187.

CORTI, L. (ed.) (1984). *Second International Conference on Automatic Processing of Art History Data and Documents, Pisa, 24-27 september 1984*, Pisa, Scuola Normale Superiore.

CORTI, L. (2003). *I beni culturali e la loro catalogazione*. Milano: Bruno Mondadori.

D'ANDREA, A. (2006). *Documentazione archeologica, standard e trattamento informatico*. Budapest: Archaeolingua.

DE MASSARY, X. & COSTE, G. (2007). *Principes, méthode et conduite de l'Inventaire général du patrimoine culturel.* Paris: Ministère de la Culture et de la Communication, direction de l'architecture et du Patrimoine. Available from: http://www.culture.gouv.fr/culture/inventai/presenta/normes/livretPMC/livretPMC_2007.pdf. [Accessed: viewed 16 July 2014].

FERRARI, O. (1972). La catalogazione dei beni culturali. *Bollettino d'Arte* 57. pp. 224-229.

FERRARI, O. (1975). Esperienze e prospettive dell'automazione in rapporto all'attuale situazione del catalogo dei beni culturali. In: *L'automazione del catalogo del patrimonio storico-artistico. Roma, 14 gennaio 1975. Informatica e documentazione. Supplemento*, pp. 13-22.

FERRARI, O. (1979). Lessici come strumento della catalogazione. In: *Atti del Convegno nazionale sui lessici tecnici delle arti e dei mestieri, Cortona, 28-30 maggio 1979*. Firenze: Eurografica. p. 181-189.

FERRARI, O. (1986) Informatica, programmazione e giacimenti culturali. *Notiziario del Ministero per i Beni Culturali e Ambientali* 8-9. pp. 27-28.

FERRARI, O. (1989) Esperienza archeologica e catalogazione. *Bollettino di Archeologia* 1. p. 24-25.

FERRARI, O. (1991) La catalogazione dei beni archeologici e le tecnologie informatiche. *Archeologia e Calcolatori*. 2. pp. 13-17.

FERRARI, O. & PAPALDO, S. (1978) Progetto di automazione del catalogo dei beni culturali in Italia. In: BAROCCHI & BISOGNI, 1978, chapter V.

GAMBA, C. (ed.) (2007). *Oreste Ferrari. Catalogo documentazione e tutela dei beni culturali. Scritti scelti*. Roma: Iacobelli.

KLAKOWICZ, B. (1970). *La collezione dei conti Faina in Orvieto. La sua origine e le sue vicende*. Roma: L'Erma di Bretschneider.

LAVECCHIA, F. & POGGI, F. (eds.) (1992). *Schede catalogo territoriale. DESC. Programma di data entry, stampa e consultazione*. Roma: ICCD.

MANCINELLI, M.L. (2004). Sistema Informativo Generale del Catalogo: nuovi strumenti per la gestione integrata delle conoscenze sui beni archeologici. *Archeologia e Calcolatori* 15. pp. 115-128.

MEYER, J.-P. (2008). Le Service de l'Inventaire du Patrimoine culturel et sa documentation. *Revue d'Alsace.* 134. p. 389-417 Available from: http://alsace.revues.org/1430#tocto1n2/. [Accessed: 16 July 2014].

MOSCATI, P. (2002). L'informatica in archeologia. In: *Il mondo dell'archeologia*. Volume 1. Roma: Istituto della Enciclopedia Italiana Treccani, pp. 318-323.

MOSCATI, P. (2013). Jean-Claude Gardin (Parigi 1925-2013). Dalla meccanografia all'informatica archeologica. *Archeologia e Calcolatori*. 24. pp. 7-24.

NATALE, M.T. & SACCOCCIO, R. (2010) Museo & Web: un kit pratico per le istituzioni culturali che vogliono realizzare un sito web di qualità. *Archeologia e Calcolatori*. 21. pp. 27-47.

PAPALDO, S. & RUGGERI, M. (1993). La catalogazione automatizzata del patrimonio archeologico nazionale in Italia. In: *International Conference on Data and Image Processing in Classical Archaeology, Ravello, 3-4 April 1992*. *Archeologia e Calcolatori* 4. pp. 323-327.

PARISE BADONI, F. *et al*. (eds.) (1980). *Dizionario Terminologico Materiali dell'età del Bronzo finale e della prima età del Ferro*. Firenze: Centro Di.

PARISE BADONI, F. & RUGGERI GIOVE, M. (eds.) (1984). *Norme per la redazione della scheda del saggio stratigrafico*. Roma: Multigrafica.

PARISE BADONI, F. & RUGGERI GIOVE, M. (eds.) (1988) *Strutturazione dei dati delle schede di catalogo: beni archeologici immobili e territoriali*. Roma-Pisa: ICCD-CNUCE.

PARRA, M.C. (1993). Applicazioni informatiche nel campo dei beni culturali: le esperienze della Scuola Normale Superiore di Pisa. In: *International Conference on Data and Image Processing in Classical Archaeology, Ravello, 3-4 April 1992. Archeologia e Calcolatori*. 4. pp. 315-321.

Rapporto sui progetti ex art. 15 Legge 41/1986 (1989), Roma.

SAXA, (1988). *Sistema per l'acquisizione controllata delle schede dell'arte*. Roma: ICCD.

SIGNORE, O. (2009). Representing knowledge in archaeology: from cataloging cards to Semantic Web. In: MOSCATI, P. (ed.) (2009) *La nascita dell'informatica archeologica. Atti del Convegno Internazionale, Roma, Accademia Nazionale dei Lincei, 24 ottobre 2008, Archeologia e Calcolatori* 20. pp. 111-128.

SIGNORE, O. (2011). Un approccio 'sociale' e ontologico alla catalogazione. *SCIRES-IT. SCIentific RESearch and Information Technology. Ricerca Scientifica e Tecnologie dell'Informazione 1(2)*. Available from: http://www.caspur-ciberpublishing.it/index.php/scires-it/article/view/90/. [Accessed: 16 July 2014].

VESENTINI, E. (2009). Gli anni '70 e la Scuola Normale. In: MOSCATI, P. (ed.) (2009) *La nascita dell'informatica archeologica. Atti del Convegno Internazionale, Roma, Accademia Nazionale dei*

Chapter 2 Field and Laboratory Data Recording

Practical Assessment of a Multi-Frequency Slingram EMI for Archaeological Prospection

François-Xavier SIMON,[1] Alain TABBAGH,[2]
Apostolos SARRIS,[1] Julien THIESSON[2]

[1]Laboratory of Geophysical – Satellite Remote Sensing and Archaeoenvironment (Lab GeoSat ResArch), Institute for Mediterranean Studies (IMS), Foundation of Research and Technologies – Hellas (FORTH), Rethymnon, Greece
[2]UMR 7619, Milieux environnementaux, transferts et interactions dans les hydrosystèmes et les sols (METIS), University Sorbonne Universités, Paris, France
fxsimus@hotmail.com

Abstract
This paper is an original approach from a practical point of view of the use of a multi-frequency Slingram Electromagnetic instrument (EMI), namely GEM-2. His use in the field of archaeological prospection is not common. Regarding previous studies, it's mainly employed for the measurement of the electrical conductivity. For this one, the depth of investigation governed by coils spacing and coils geometry, give us too deep information for most archaeological sites. Nevertheless, if we hold the GEM-2 instrument in another way, that suggested by the manufacturer, we should first obtain a useful information on the magnetic susceptibility (which is usual with EMI). Afterwards, we can go further to discuss here the potentiality and the limit of this instrument to improve the measurement of other physical properties of the soil, both well-known: electrical conductivity, the range of which can be extended to small conductivities, and magnetic susceptibility but also new physical properties: magnetic viscosity and dielectric permittivity.

Keywords: Geophysical Prospection, EMI Multi-Frequency, Magnetic Susceptibility, Electrical Conductivity, Magnetic Viscosity

Introduction

The main purpose of this presentation is to take into consideration the interest of a multi-frequential Slingram EM instrument for archaeological survey. For the past forty years, electromagnetic methods were employed for archaeological prospection. Some recent studies, undertaken with commercial instruments now available, confirm that their use avoids disappointment when only electrical or magnetic survey could be used (Bonsall *et al.*, 2013) which has been well-known since the past 30 years (Scollar *et al.*, 1990; Tabbagh, 1986, 1984). Under the LIN approximation, e.g. for the most common field and with usual frequency (not far from10 kHz), we can simultaneously achieve measurements of the electrical conductivity and of the magnetic susceptibility. Moreover interest of the EM instrument is the rate of acquisition, because the measurement doesn't require, unlike electrical methods, any contact between the ground surface and the instrument, allowing a fast data collection.

Nevertheless EM instrument are less used than magnetic or electric survey for evident reasons of efficiency and robustness. The main limitation, regarding the electrical resistivity, is that the contrasts in apparent electrical conductivity are usually less than those of apparent electrical resistivity between archaeological features and soil sediment. Likewise, magnetic survey is sensitive to the sum of the different magnetizations unlike EM instruments which are only related to the magnetic susceptibility. Another explanation is the low spatial resolution of this instrument and some technical problems, like its calibration, the strong drift during the acquisition, current problem of offset, and a poor global estimation of the physical parameter due to the dependency on electronic operating.

In the last ten years however, new multi-coil/multi-spacing instruments have a strong impact on the development of EM studies. Their ability to map different volume of soils for both electrical conductivity and magnetic susceptibility are especially interesting for studies related to the depth of the target that is impossible with a single spacing. They start to be used for the estimation of the depth of soil, and try to figure out the paleaotopography in some specific area with well-known soils characteristics (De Smedt *et al.*, 2014). Recent studies also show some new datasets with high spatial resolution and correction of drift and acquisition effect (De Smedt *et al.*, 2013).

The use of multi-frequency EM Slingram instruments is anecdotic but it has been proposed 20 years ago (Bongiovanni *et al.*, 2008) and discussed in soil-science to map the conductivity at different depths. This was not been the case in the field of archaeological research where their efficiency was never approved for archaeological characterization (Bullas *et al.*, 2000; Schmidt and Bonsall, 2002). These first works show that the utility of multi-frequency EMI (GEM300) is limited by some technical trouble like drift etc., and that not any survey has shown new information related to the depth of investigation.

Until now, examples of use of multi-frequency EM instrument are quite limited and could be subjected to discussion. The main utility of this instrument is to do inversion of multi-frequency data for the electrical conductivity (Brosten et al., 2011). Even if the inversion is completely conceivable it is based on the same volume of soils, and the resulting inversion is in this way a solution of an extremely ill posed problem. Indeed, according to the theory of EM Slingram measurement and LIN limitation the use of multi-frequency could not affect the depth of investigation of this instrument, unlike the changing of the coils' spacing. This point could explain the main disappointment in its application in the archaeological domain.

Accepting this point, the multi-frequency EMI shows also new possibilities. Multi-frequency could allow to surpass the LIN approximation limitation and to characterize electrical conductivity in very conductive areas (Delefortrie et al., 2014) and it also allows to correct the effect of conductivity for the measurement of the magnetic susceptibility (Tabbagh, 1986). Furthermore, this correction, suggests the determination of the magnetic viscosity, which could be measure in different ways. The first one is the measurement of the frequency-dependent susceptibility, with multi-frequential instrument like the MS2B (Bartington ltd) does (Dearing et al., 1996). Another one is the measurement of the magnetic viscosity by TDEM instrument like the VC100 (Thiesson et al., 2007). Problem for both of them is the time of acquisition. Today TDEM is only able to take measurement point by point while measurement on soil sample requires a specific and heavy process: soil sampling is not able to cover a large area with a sufficiently narrow step.

We will discuss here the effect of frequency on the measurement, try to figure out the question of the depth of investigation by a visual presentation, and suggest a new procedure of acquisition to enhance the sensitivity of the instrument.

Instrumental specification

GEM-2 was specifically employed for the particular study (Fig. 1). It's a multi-frequency instrument developed by Geophex Ltd and used here for some practical reason. At first it presents a great flexibility and the possibility to choose up to 5 frequencies between 300 Hz and 90 kHz. The use of 5 frequencies limits the acquisition rate to 1 Hz, which is still well-adapted for a field walking acquisition. Transmitter-Receiver coils are in a co-planar configuration and the design of the instrument suggests HCP geometry. For VCP it will be useful in the future to improve the way to hold and fix the instrument in order to minimize the slight change of coil orientation. This effect could considerably affect the measurement of the magnetic susceptibility.

Figure 1: Acquisition with the GEM-2 with the new procedure on the site of Zerelia (IGEAN project- www.igean.ims.forth.gr) (photo by M. Manataki).

This is the reason that we will focus for the following part on HCP measurements.

The instrument doesn't have a common design, since it consists of three coils. A transmitter coil, which generates the electromagnetic field and a receiver coil which collects the secondary field and a bucking coil, used to buck the primary field. The two mains coils are separated by 1.66 m and the bucking coils are separated to the transmitter coils by 1.035 m.

Effect of the frequency on the physical parameters

One of the main interests of electromagnetic methods comparing to the other methods is to offer the possibility to measure simultaneously different physical parameter: electrical conductivity and magnetic susceptibility, but also magnetic viscosity. Until now, the measurement of this last physical property remains experimental. Further development needs to be focus on this behavior. These three physical properties have some link with the frequency, for intrinsic reasons in the case of the magnetic susceptibility, or related to the way of the measurement with a Slingram instrument.

Electrical resistivity is well known in archaeology and probably one of the most useful physical property helping to detect remains in soil. Differences between construction materials and sediment layers, with more or less clay content, allow the characterization of the targets. High contrasts of electrical resistivity in soil seem more common than high contrast of electrical conductivity in the case of archaeological survey. Detection of resistive features seems hazardous with EM instrument, even if some papers prove the interest of the EM methods in this case (Thiesson et al., 2009). However, measurement of electrical conductivity could be very useful in the case of soils with high clay content. In this case, this property could show a contrast in the filling of ditch or pit, that electrical resistivity sometime doesn't reveal so clearly. Regarding the dependency of the electrical conductivity to the frequency, in this range of measurement between 1 and 1000 Ω.m (and its equivalent 1 to 1000 mS/m), the resistivity is assumed independent on frequency.

Nevertheless changing the frequency will considerably change the values of the EM measurements. It's the reason why we need to take into account the dependency of the frequency for the conversion of the instrumental response in physical properties.

Magnetic properties of soil are, if not the first ones, among the main physical properties used to detect archaeological remains. This property affects the earth magnetic field distribution and allows detection of remains by magnetic prospection. Magnetic susceptibility is well-known, as a physical property useable to features' detection. It plays an important role in EM response, on the in-phase part of the signal, in respect of the Low Induction Number condition.

Natural enhancement of the magnetic susceptibility in the top soil, and increasing of magnetic grain in anthropogenic soil is well-known. It's also dependent on the substrate and its content in iron bearing particles. Some studies showed a strong correlation between the magnetic susceptibility and the organic matter content, and correlation with the phosphorus content that is of great interest in archaeological survey (Hulin *et al.*, 2014).

Actually magnetic susceptibility has a complex form, which means it presents a quadrature out-of-phase part. This part is called the magnetic viscosity and it is close to 6 % of the in-phase magnetic susceptibility, as it was proved experimentally by Tite and Mullins on soil samples of different sites (Mullins and Tite, 1973). It was also studied by Dabas and Skinner (Dabas and Skinner, 1993), who showed that for metallurgic wastes, the ratio between magnetic susceptibility and magnetic viscosity could decrease to 2%. The explanation of this complex behavior is related to the nature and the size of the magnetic grains involved in metallurgic activities.

As proved by these previous studies, magnetic susceptibility is frequency dependent. When the frequency is increasing, the in-phase magnetic susceptibility is decreasing. On the contrary, the magnetic viscosity is non-dependent of the frequency, whilst it is linked to the frequency dependency of the magnetic susceptibility. Knowing the value of the susceptibility for different frequencies one can deduce the magnetic viscosity. The quadrature out-of-phase part of the susceptibility, magnetic viscosity, logically affects the quadrature out-of-phase part of the EMI measurement. In this case knowing the value of the electrical conductivity allows extracting the value of magnetic viscosity. We will see later on, that this processing is not so easy due to the low values of the magnetic viscosity compared to the higher effect of the conductivity on EM response. Consequently uncertainty on the determination of the electrical conductivity could strongly affect the determination of the magnetic viscosity of soils.

The dielectric permittivity could also affect the electromagnetic measurement (Huang and Fraser, 2001). Usually, EM instrument use frequency around 10 or 30 kHz, and in most cases, dielectric permittivity has no effect on the measurement. By increasing the frequency to be close to 90 kHz, dielectric permittivity starts to affect considerably the response. Compared to the one of the electrical conductivity this effect is quite limited, but if one is interested in the measurement of the magnetic viscosity, then he needs to take it into account. The dielectric permittivity is completely frequency dependent, as it corresponds to different polarization effects. In the frequency range (<90 k Hz), it is correlated with the electrical conductivity, and thus it is quite different from the dielectric permittivity governing the measurement with GPR systems (from 50 to 1000 MHz). Any comparison of both dielectric permittivities is possible. But permittivity measurement could be helpful for archaeological survey because it could show some new information and be a solution in the case of low conductivity soils where the measurement of the electrical conductivity is very tricky.

Regarding these different behaviors of magnetic and electric physical properties, we suggest to use for the measurement different frequencies with a logarithmic progression. In the future it will be interesting to reach a frequency above 100 kHz and try to determine the dielectric permittivity. But this measurement needs also a thrust study of the behavior of the instrument. The use of a high frequency increases the conductivity response and thus facilitates its measurement. Even if the dielectric permittivity and magnetic susceptibility affect the measurements for highest frequencies, the dynamic of the measurement of electrical conductivity is extended and probably well adapted for more resistive areas.

Depth of investigation and volume of soils affected by the measurement

Two ways exist to perform an electromagnetic sounding. The first one is to change the coils' spacing or coil altitude (geometric sounding) and the second one is to change the frequency of the measurement (frequency sounding that can be Fourier transformed in time-domain sounding). Both ways allow changing the skin depth of the measurement, which will affect the depth of investigation. But this consideration is only available for deep targets as for geological characterization (Schamper *et al.*, 2014). In the case of an archaeological prospection, the target depth range and geometry limit the size of the instrument and the induction number remains low. Only coil spacing and coils' array will then be able to change the depth of investigation. Although the previous studies clearly showed that the depth of investigation is determined by the coils' spacing and not by the frequency (Mc Neill, 1996), some recent papers dispute the particular issue (Bongiovanni *et al.*, 2008; Brosten *et al.*, 2011; Minsley *et al.*, 2012)

Regarding all these specifications about multi-frequency measurement and dependency of the physical properties on the frequency, we suggest to use this type of instrument, practically the GEM-2, for another purpose: to extract and/or to correct the values of the different physical EM properties. However, to achieve this aim we need to check if the volume of soil affected by the measurements at the different frequencies is almost the same, in both 1D and

3D. So, we will do a 1D simulation for different frequencies and then extend this consideration to a 3 D simulation.

1D simulation

We do a simulation of the response of the instrument for a thin magnetic and resistive layer (thickness: 0.1 m, electrical resistivity: 200 Ω.m, magnetic susceptibility: 100 10-5 SI, magnetic viscosity: 6 10-5 SI) moving from the ground surface until 5 m in depth, in a low magnetic and conductive medium (electrical resistivity: 20 Ω.m, magnetic susceptibility: 10 10-5 SI magnetic viscosity: 6 10-5 SI) at two different frequencies, 5010 Hz and 40050 Hz. If the curve of sensitivity is not the same for the different frequencies this will affect the way to extract new information like magnetic viscosity.

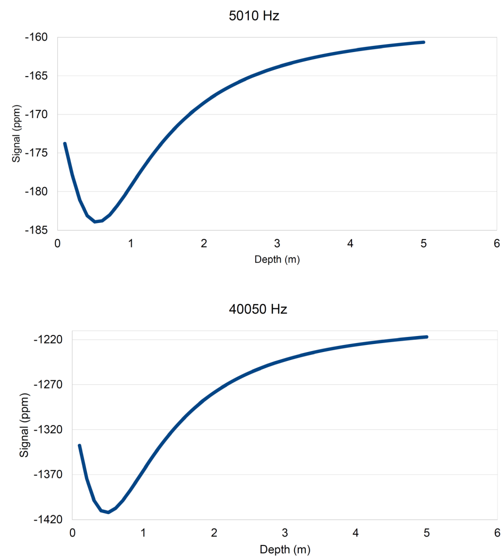


Figure 2: 1D simulation of the sensitivity to a thin layer for the quadrature part of the HCP signal

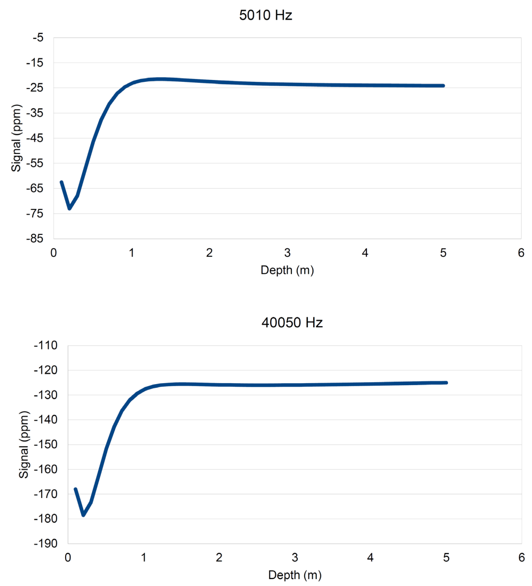


Figure 3: 1D simulation of the sensitivity to a thin layer for the in-phase part of the HCP signal.

To do this simulation we use the solution which takes into account the electrical conductivity, the complex magnetic susceptibility and the dielectric permittivity of the different layers, the frequency of the instrument and the altitude of its sensors. Mc Neill's (Mc Neill, 1980) solutions used to do such the simulations would be here obsolete because they don't take into account the altitude of the device, the susceptibility and the self and cross-induction coupling between media. Maxwell's equations leads here to a Hankel transform expression (Guptasarma and Singh, 1997).

Even if it was predictable, these results show that we cannot expect an electromagnetic sounding with the multi-frequency instrument for both electrical and magnetic properties of soil (Fig. 2 and 3). We can also conclude from this study that, unlike the sensitivity to the depth of the target, the quadrature out-of-phase part signal dynamically changes with the frequency. The experiment shows also the depths of investigation for this instrument: approximately 2.5 m for the conductivity and 1.5 m for the magnetic susceptibility.

3D simulation

The 1D simulation clearly shows a similarity despite using different frequencies. To confirm that we can use different frequencies to extract the properties, we need also to check if the volume of investigation is always the same. To do this, we proceed to a 3D simulation, taking into account the same model as for the 1D simulation (Fig. 4). We do it for a block with a size of 2*2*1 m in a homogeneous media with the same physical properties as for the 1D simulation

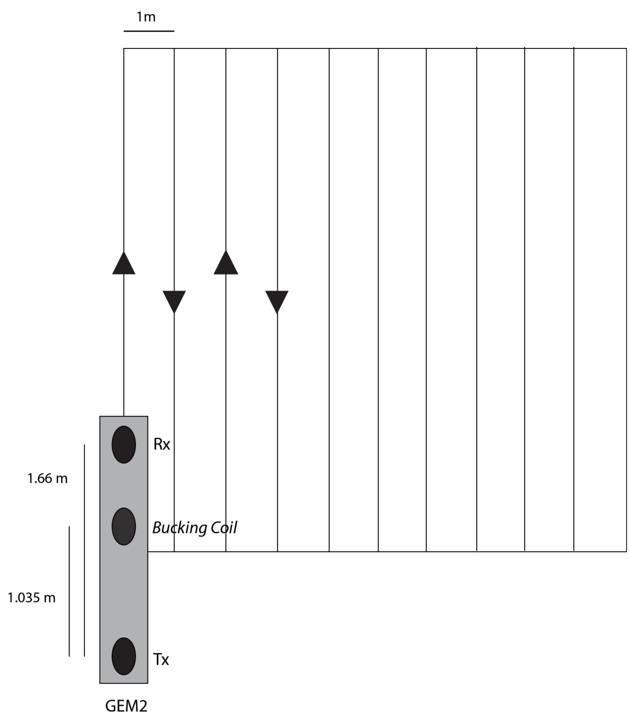


Figure 4: Schema of the 3D simulation.

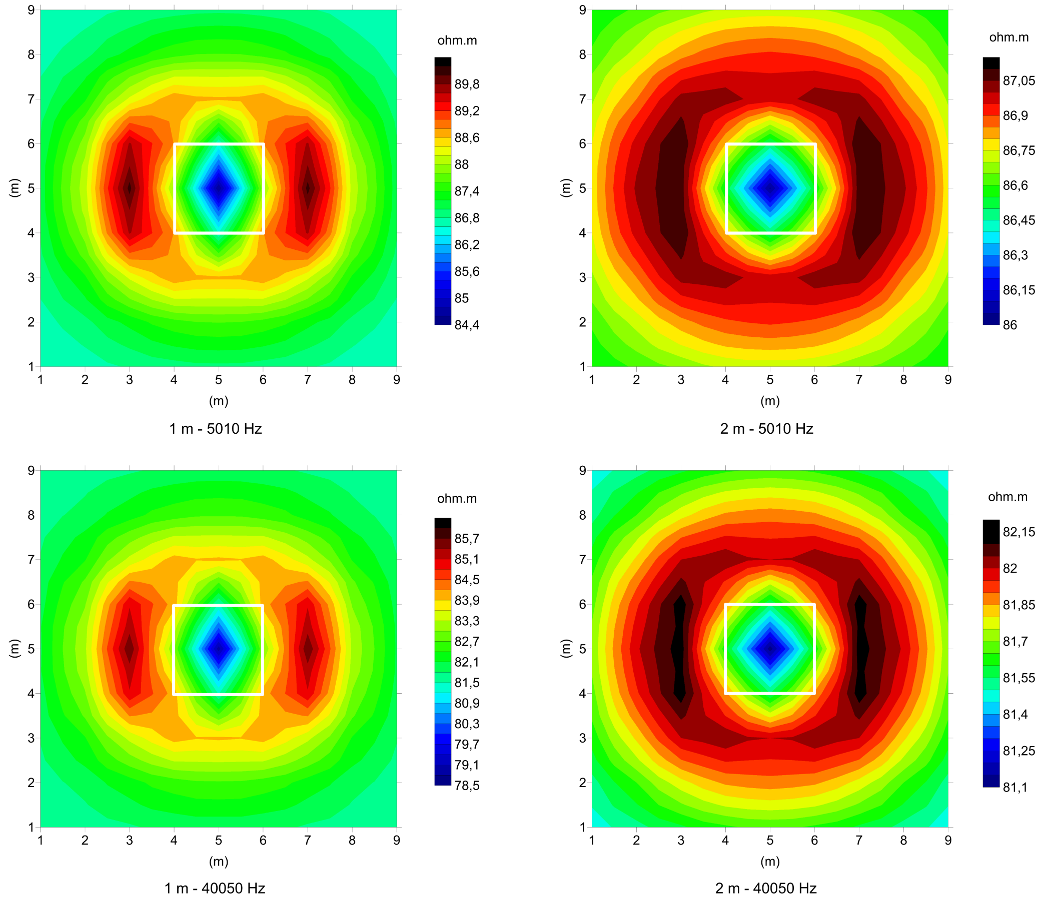


Figure 5: Results of the 3D simulation for the quadrature out-of-phase part of the HCP signal. Results are express in electrical resistivity units (Ω.m).

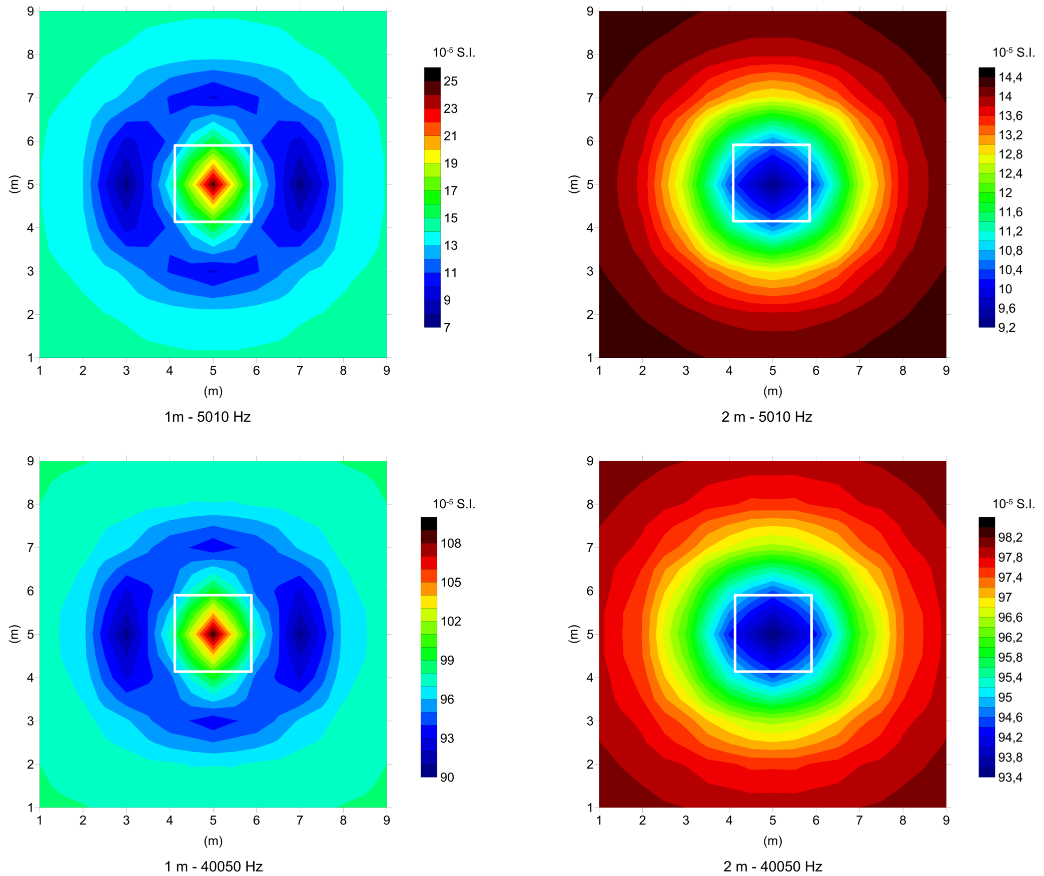


Figure 6: Results of the 3D simulation for the in-phase part of the HCP signal – Results are expressed apparent magnetic susceptibility units.

at two different depths (1 and 2 meter). We are here more interested by the shape and the intensity of the anomalies.

For both frequencies, 5010 and 40050 Hz, the shape of the anomaly is exactly the same for the electrical conductivity (Fig. 5). We can see a very small difference in amplitude for the anomaly (lower than 2 Ω.m) which could come from the effect of the magnetic viscosity which was not deleted before the conversion of the response in electrical conductivity.

For the in-phase part of the signal and conversion of magnetic susceptibility we obtain the same shape (Fig. 6). Magnitude of anomaly is exactly the same but with an offset. As we can see magnetic susceptibility increases with the frequency which is unexpected by theory but results from the geometrical complexity of the response in HCP measurement.

This result confirms that the use of different frequencies is not affected by the depth of investigation or the volume of soils investigated. In this case we can use the different frequencies to extract directly the physical properties.

Calibration of the instrument

The main default of EM instrument is its dependency from electronics and mechanics (for in-phase response). This means that for each survey we could have an offset for the measurement. Another problem is the determination of the gain coefficient linking in-field measurements and theoretical responses. We follow here the procedure described by Thiesson (Thiesson et al., 2014).

The first step is the determination of the coefficient of both part of the complex signal. For the quadrature part, we choose to neglect the influence of magnetic viscosity on the quadrature out-of-phase part of the signal and to take into account only the electrical conductivity when it is sufficiently high.

We do the calibration on a specific location and we compare the measurement at two different altitudes with a theoretical response deduced from resistivity sounding interpretation. We use then a small resistivity meter (Liepmann Ltd.) and we proceed to a sounding with a Wenner array up to 10 meter between electrodes. This spacing and the corresponding depth of investigation is large enough.

This comparison gives us the coefficient for the quadrature part of the signal and the offset of the instrument for the conductivity measurement. We repeat this calibration around 10 times during different fieldworks, and with some exception, we found always a coefficient around -1 digit/ppm. This means that the instrument is well-calibrated for electrical conductivity but the response is reverse.

To do an estimation of the offset of the in-phase part of the signal we do the same measurement at two different heights but in this case we consider that the measurement of the magnetic susceptibility response at the highest altitude (which needs to be at least 2 meter for the GEM2) is null and only affected by the electrical conductivity. Then we use this level to calibrate the offset of the instrument.

To do the estimation of the coefficient between the instrument and theoretical simulation for the in-phase part of the signal, we use an aluminum sphere (Thiesson et al., 2014). This one affects exclusively the in-phase part of the signal and not the quadrature part. This test allows also checking if we can use this signal to measure independently both electrical conductivity and magnetic susceptibility. We do this test with an aluminum sphere of 4 cm radius. The instrument is laid on the ground, in a stationary position and we move the aluminum ball every 10 cm to do the measurement. This experiment shows that the signals need to be divided by 4 to obtain the true values in ppm and then to be converted in magnetic susceptibility (4 digits/ppm and -1.9 10-5 ppm/u.S.I).

New archaeological oriented procedure

Taking into account the previous result and the theoretical background related to the multi-frequency instrument we suggest changing the procedure in order to enhance the reliability of the response for the magnetic susceptibility but also for electrical conductivity and magnetic viscosity.

Holding the instrument at a height of 1 m (as suggested by the manufacturer) implies a reduction of the sensitivity of the instrument to the magnetic susceptibility. Fig. 7 show that for the electrical conductivity, the difference between holding the instrument a 1 m or 0.1 m close to the ground don't change so much the volume of soil affected by the measurement. But at an altitude of 1 meter, the instrument is more affected by the first layer of soil, which could be a problem when the soil contains some metallic scraps. In the case of the in-phase response (Fig. 8), the difference is substantial. At 1 m we loose almost completely the response of the measurement while at 0.1 m we measure a large volume of soil that allows to have a good determination of the magnetic susceptibility. Regarding the cause of the response of the magnetic susceptibility and the response of the magnetic viscosity, we can deduce

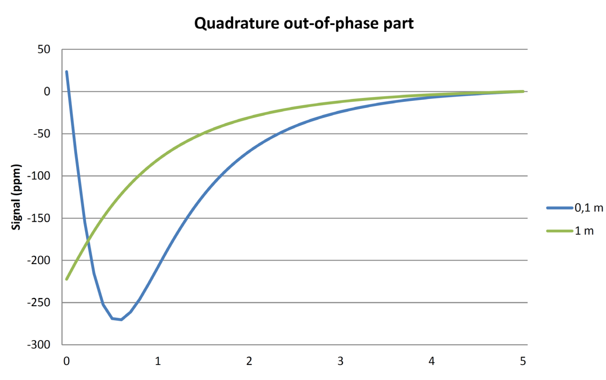


Figure 7: Difference between the sensitivity to a thin layer for altitudes of the sensor of 0.1 m and 1 m for the quadrature out-of-phase part of the HCP signal

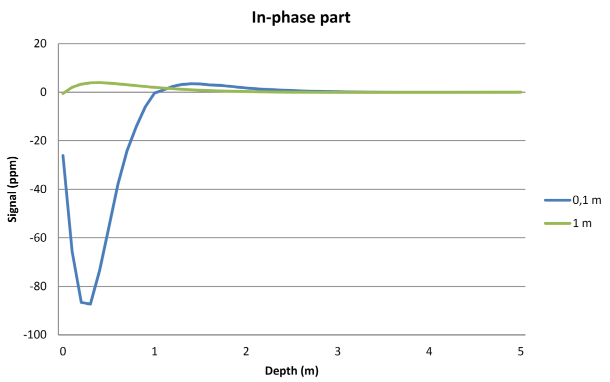


Figure 8: Difference between the sensitivity to a thin layer for altitudes of the sensor of 0.1 m and 1 m for the in-phase part of the HCP signal.

the same observation for the measurement of the magnetic viscosity.

The different simulations, in 1D and 3D indicate that the spacing between profiles doesn't need to be more than 1 m (regarding the sensitivity of the instrument for magnetic susceptibility). Actually, this spacing is given by the problematics of the measurements, but for archaeological purposes, which we are discussing here, the more interesting spacing seems to be 1 m. We can expect a more accurate map with 0.5 m for very near surface targets, but this accuracy is required only for few sites with a well-known stratigraphy. The calibration needs henceforth to be a crucial step of the EM measurement, to make an interpretation regarding the true value of apparent resistivity and not only variations in measurements.

To illustrating this specification we will present some results achieved with the GEM-2 taking into account what we said about the way to hold the instrument, the frequency currently used, etc. This presentation will allow checking what we can expect on sites from this kind of measurements in this configuration. The first one is a Neolithic site, studied as a part of a large project about Neolithic settlements in Thessaly, Greece and the second one a Hellenistic quarter in ancient Demetriada city, close to Volos. The main object of the study is not to speak about the archaeological results that need another kind of integrative approach but discussing the EM result regarding what we did with this kind of method.

In the first case of Demetriada, we present results for the magnetic susceptibility, the electrical conductivity (Fig 9) and the magnetic viscosity after having achieved the calibration process to transform the signal in these three physical properties; the last one, magnetic viscosity (Fig 11), exhibits an important high frequency noise. The way to extract at each point the three properties is the following: the difference between the quadrature responses at 5010 and 13370 Hz (theoretically free of magnetic viscosity effect) is used to define the apparent conductivity, and this value is used to calculate the in-phase conductivity response which is subtracted from the in-phase response at 5010 Hz and the quadrature response

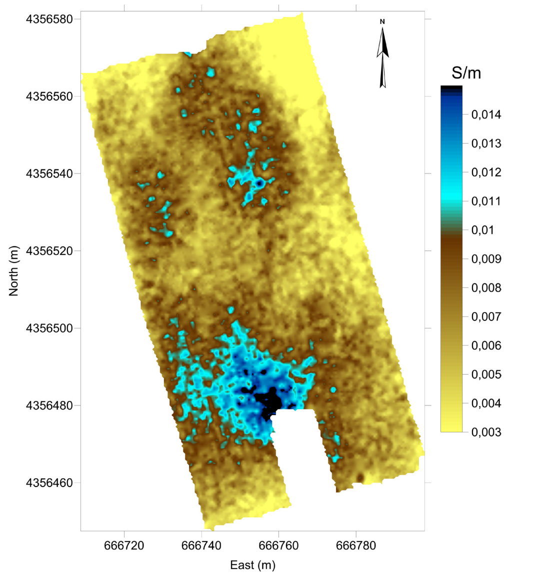


Figure 9: Apparent electrical conductivity on Demetriada site, Thessaly (Greece) (GEM-2, HCP).

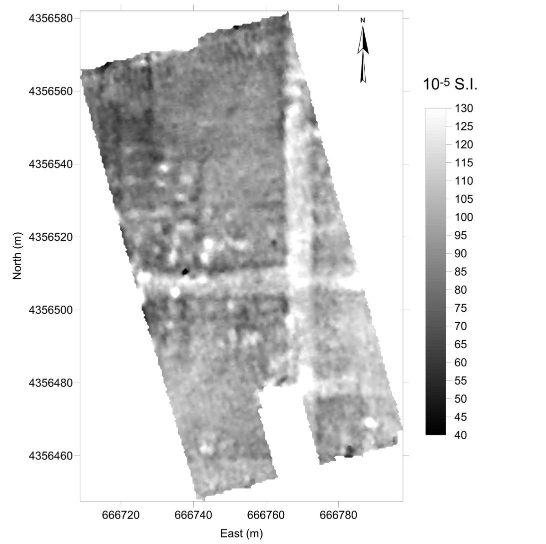


Figure 10: Apparent magnetic susceptibility on Demetriada site, Thessaly (Greece) (GEM-2,HCP).

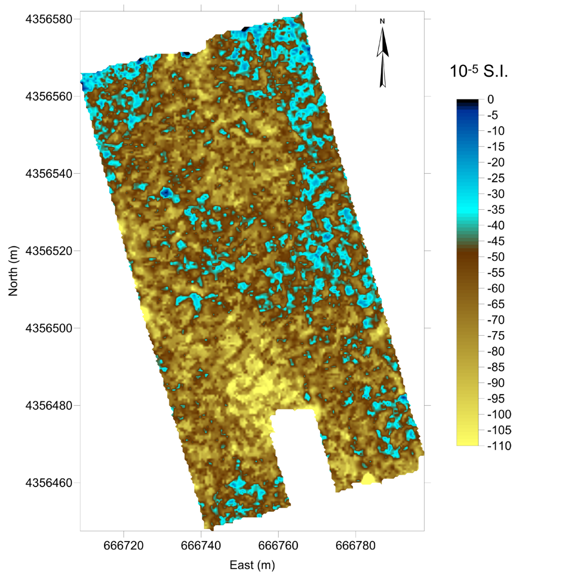


Figure 11: Apparent magnetic viscosity on Demetriada site, Thessaly (Greece) (GEM-2, HCP).

at 5010 Hz which is subtracted from the total quadrature response to get the magnetic viscosity. In the case of a high resistivity, the accuracy of EM decreases exponentially. In this case a large contrast of resistivity will produce a very weak contrast in term of electromagnetic response and may affect determination of the magnetic viscosity. Nevertheless, the noisy data could also come from a lack of sensitivity of the EM instrument on this field. Weak value of magnetic viscosity will produce a weak signal, less than the incertitude on the electrical conductivity determination.

In this case magnetic susceptibility (Fig 10) is more interesting. The high value of the sediment and the vicinity to the archaeological targets may produce clearer images. Other methods such as GPR in a high resistivity context could give very good results, whereas magnetic data are affected by metallic scraps. The main cross roads are characterized by a positive contrast unlike walls of building which are characterized by a negative contrast. In reality, both of them could be built with the same material and then to have the same contrast of physical properties but in the case of the HCP measurement, the depth of the target will affect the sign of the apparent anomaly. In this case, the road is probably closer to the ground surface and walls are deeper. This explanation is found by HCP theoretical responses and on the comparison with other datasets. For a single survey there is no way to say if the target is deep or if the contrast is inverted and complementary technics are required.

In the case of Almiriotiki site, we did the same processing following the same way of acquisition. We used the two lowest frequencies (5010 Hz and 13370 Hz) to extract the information and we did not consider the highest frequencies because they could be affected by dielectric permittivity.

For this area the conductivity is quite high (Fig 12). The vertical electrical sounding shows a geological bank which reaches 10 Ω.m. As usually, the depth of investigation for electrical conductivity is too much important and doesn't allow a characterization of archaeological targets. Nevertheless, this site looks like having formed by large deposits of anthropogenic sediment. The electrical conductivity clearly shows a surrounding ditch (also clearly visible on the magnetic survey), a difference of material between the site and the area around the hill and some spatial distribution of features which could reveal paths and enclosures. The magnetic susceptibility is clearer (Fig 13), probably because the depth of investigation is nearer and fits with the depth of the archaeological layer. In this case, we can clearly follow some walls, and magnetic anomalies, as ditches. Regarding the magnetic viscosity (Fig 14), we lost the spatial accuracy and this technic doesn't allow until now to extract good archaeological information for the magnetic viscosity. In terms of a large scale investigation this information is nevertheless very interesting because it could reveal variation in the ratio of the magnetic viscosity on magnetic susceptibility and

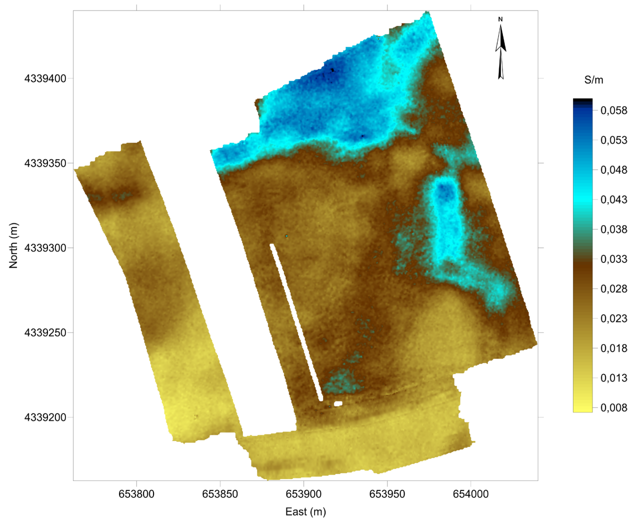


Figure 12: Apparent electrical conductivity map at Almiriotiki site, Thessaly (Greece) (GEM-2, HCP)

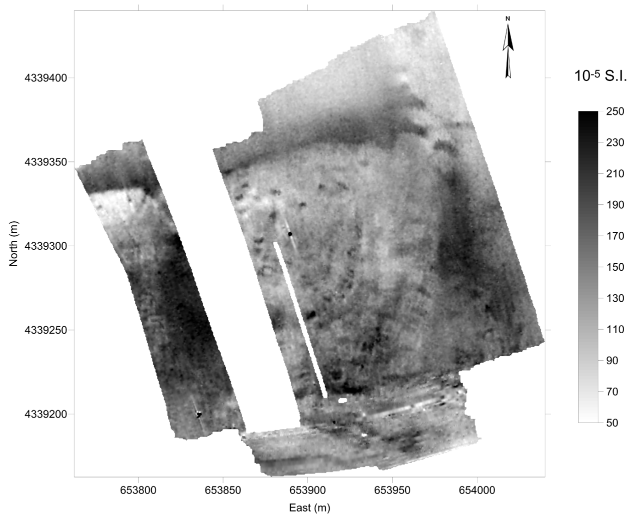


Figure 13: Apparent magnetic susceptibility map at Almiriotiki site, Thessaly (Greece) (GEM-2, HCP)

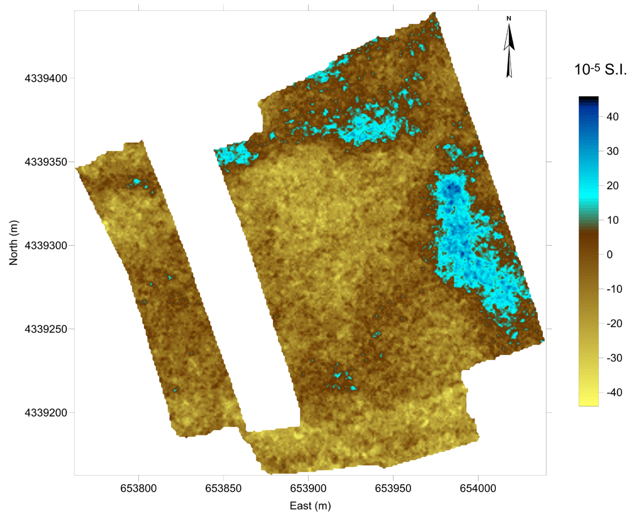


Figure 14: Apparent magnetic viscosity map at Almiriotiki site, Thessaly (Greece) (GEM-2, HCP)

also gives us new information on the grain size affected by the formation process of magnetic properties. But until now this process is not achieved, because the complexity of the HCP response and the poor calibration of the EM instrument make this estimation difficult to be achieved.

Conclusion

After some experiments, multi-frequency EM for archaeological investigations looks well adapted unlike the report given by previous studies. In the specific case of the GEM-2 with 1.66 meter coil spacing, electrical conductivity has a too big depth of investigation which explains some difficulty in the characterization of archaeological features. Nevertheless for ditches, with a large width and high depth, it could be very useful. On the other hand, the GEM-2 looks like one of the best instruments for the characterization of the magnetic susceptibility. The instrument has a huge sensitivity for the magnetic properties, especially for the in-phase part of the signal. Now we need to check more accurately and confirm the ratio digit/ppm to give the exact value of the magnetic susceptibility. Previous studies show a ratio of 4, but this part need to be improved. The interest of multi-frequency EM methods is also focused in the use of different frequencies to correct for the LIN and instrumental limitations: correction of the electrical conductivity (Benech and Marmet, 1999), correction of the effect of the electrical conductivity on the in-phase part of the signal, and estimation of the magnetic viscosity on the quadrature out-of-phase part of the signal. For the magnetic viscosity, these first results don't give information related to the identify features and, partly thanks to the noise rather than regional variations. Nevertheless, this instrument offers for the first time the possibility to map this property with a continuous acquisition, which was until now unfeasible. With the possibility to reach 100 kHz and above, we expect also to extract some information about dielectric permittivity in further studies. This one could be very useful for some areas, with a high resistivity. These studies indicate the promising future of the evolution of multi-frequency EM Slingram instruments and their impact in archaeological surveys.

Acknowledgement

This work was performed in the framework of the IGEAN ('INNOVATIVE GEOPHYSICAL APPROACHES FOR THE STUDY OF EARLY AGRICULTURAL VILLAGES OF NEOLITHIC THESSALY') project which is implemented under the 'ARISTEIA' Action of the 'OPERATIONAL PROGRAMME EDUCATION AND LIFELONG LEARNING' and is co-funded by the European Social Fund (ESF) and National Resources. The authors would like to thank all the team which participated to the dataset acquisition during November 2013 and March 2014 fieldwork campaigns.

References

BENECH, C., MARMET, E., (1999). Optimum depth of investigation and conductivity response rejection of the different electromagnetic devices measuring apparent magnetic susceptibility. *Archaeological Prospection,* 6, 31–45. doi:10.1002/(SICI)1099-0763(199903)6:1<31::AID-ARP112>3.0.CO;2-N

BONGIOVANNI, M.V., BONOMO, N., DE LA VEGA, M., MARTINO, L., OSELLA, A., (2008). Rapid evaluation of multifrequency EMI data to characterize buried structures at a historical Jesuit Mission in Argentina. *J. Appl. Geophys*. 64, 37–46. doi:10.1016/j.jappgeo.2007.11.006

BONSALL, J., FRY, R., GAFFNEY, C., ARMIT, I., BECK, A., GAFFNEY, V., (2013). Assessment of the CMD Mini-Explorer, a New Low-frequency Multi-coil Electromagnetic Device, for Archaeological Investigations. *Archaeological Prospection* 20, 219–231. doi:10.1002/arp.1458

BROSTEN, T.R., DAY-LEWIS, F.D., SCHULTZ, G.M., CURTIS, G.P., LANE JR., J.W., (2011). Inversion of multi-frequency electromagnetic induction data for 3D characterization of hydraulic conductivity. *J. Appl. Geophys.* 73, 323–335. doi:10.1016/j.jappgeo.2011.02.004

BULLAS, S., BARRAT, G., GRASSI, R., (2000). Paradox or panacea?-archaeological field trials with the GEM-300 Multi Frequency Electromagnetic Profiler, in: *Making the Connection to the Past. CAA99 : Proceedings of the 27th Conference. Presented at the Computer Applications and Quantitative Methods in Archaeology*, Faculty of Archaeology, Leiden University, Leiden, Dublin, pp. 75–84.

DABAS, M., SKINNER, J., (1993). Time/domain magnetization of soils (VRM), experimental relationship to quadrature susceptibility. *GEOPHYSICS* 58, 326–333.

DE SMEDT, P., SAEY, T., LEHOUCK, A., STICHELBAUT, B., MEERSCHMAN, E., ISLAM, M.M., VAN DE VIJVER, E., VAN MEIRVENNE, M., (2013). Exploring the potential of multi-receiver EMI survey for geoarchaeological prospection: A 90 ha dataset. Prox. Soil Sens. Pap. Second Glob. Workshop Prox. Soil Sens. Second Glob. Workshop Prox. Soil Sens. 199, 30–36. doi:10.1016/j.geoderma.2012.07.019

DE SMEDT, P., VAN MEIRVENNE, M., SAEY, T., BALDWIN, E., GAFFNEY, C., GAFFNEY, V., (2014). Unveiling the prehistoric landscape at Stonehenge through multi-receiver EMI. *Journal of Archaeological Science* 50, 16–23. doi:10.1016/j.jas.2014.06.020

DEARING, J.A., DANN, R.J.L., HAY, K., LEES, J.A., LOVELAND, P.J., MAHER, B.A., O'GRADY, K., (1996). Frequency-dependent susceptibility measurements of environmental materials. *Geophys.*

J. Int. 124, 228–240. doi:10.1111/j.1365-246X.1996.tb06366.x

DELEFORTRIE, S., SAEY, T., VAN DE VIJVER, E., DE SMEDT, P., MISSIAEN, T., DEMERRE, I., VAN MEIRVENNE, M., (2014). Frequency domain electromagnetic induction survey in the intertidal zone: Limitations of low-induction-number and depth of exploration. *J. Appl. Geophys.* 100, 14–22. doi:10.1016/j.jappgeo.2013.10.005

GUPTASARMA, D., SINGH, B., (1997). New digital linear filters for Hankel J0 and J1 transforms. *Geophys. Prospect.* 45, 745–762. doi:10.1046/j.1365-2478.1997.500292.x

HUANG, H., FRASER, D., (2001). Mapping of the resistivity, susceptibility, and permittivity of the earth using a helicopter borne electromagnetic system. *GEOPHYSICS* 66, 148–157.

HULIN, G., BROES, F., FECHNER, K., (2014). Caractérisation de phénomènes anthropiques par la mesure de paramètres magnétiques sur surface décapée : Premiers résultats sur le projet Canal Seine-Nord Europe. *ArchéoSciences Rev. Archéom.* 61–70.

MC NEILL, J.D., (1980). Applications of transient electromagnetic techniques.

MC NEILL, J.D., (1996). Why doesn't Geonics limited build a multi-frequency EM31 or EM38?

MINSLEY, B.J., SMITH, B.D., HAMMACK, R., SAMS, J.I., VELOSKI, G., (2012). Calibration and filtering strategies for frequency domain electromagnetic data. *J. Appl. Geophys.* 80, 56–66. doi:10.1016/j.jappgeo.2012.01.008

MULLINS, C.E., TITE, M.S., 1973. Magnetic viscosity, quadrature susceptibility, and frequency dependence of susceptibility in single-domain assemblies of magnetite and maghemite. *J. Geophys. Res.* 78, 804–809. doi:10.1029/JB078i005p00804

SCHAMPER, C., AUKEN, E., SØRENSEN, K., (2014). Coil response inversion for very early time modelling of helicopter-borne time-domain electromagnetic data and mapping of near-surface geological layers. *Geophys. Prospect.* 62, 658–674. doi:10.1111/1365-2478.12104

SCHMIDT, A., BONSALL, J., (2002). Archaeological Evaluation of The Multi-frequency Electromagnetic Slingram Device Gem 300. Presented at the EGS XXVII General Assembly, Nice.

SCOLLAR, I., TABBAGH, A., HERZOG, I., HESSE, A., (1990). *Archaeological prospecting and remote sensing*, Cambridge: Cambridge University Press.

TABBAGH, A., (1984). On the comparison between magnetic and electromagnetic prospection methods for magnetic features detection. *Archaeometry* 26, 171–182. doi:10.1111/j.1475-4754.1984.tb00331.x

TABBAGH, A., 1986. Applications and advantages of the Slingram electromagnetic method for archaeological prospecting. *GEOPHYSICS* 51, 576–584.

THIESSON, J., DABAS, M., FLAGEUL, S., 2009. Detection of resistive features using towed slingram electromagnetic induction instruments. *Archaeological Prospection* 16, 103–109. doi:10.1002/arp.350

THIESSON, J., KESSOURI, P., SCHAMPER, C., TABBAGH, A., 2014. Calibration of frequency-domain electromagnetic devices used in near-surface surveying. Near Surface Geophysics. doi:10.3997/1873-0604.2014012

THIESSON, J., TABBAGH, A., FLAGEUL, S., 2007. TDEM magnetic viscosity prospecting using a Slingram coil configuration. Surf. Geophys., *Archeogeophysics* 5, 363–374. doi:10.3997/1873-0604.2007018

Utilizing Magnetic Prospection and GIS to Examine Settlement Organization in Neolithic Southeastern Europe

Alexis NIEKAMP,[1] Apostolos SARRIS[2]

[1]Ohio State University, Anthropology, USA
[2]Laboratory of Geophysical – Satellite Remote Sensing and Archaeo-environment, Institute for Mediterranean Studies, Foundation for Research and Technology, Greece (I.M.S.-F.O.R.T.H.)
niekamp.29@osu.edu, asaris@ret.forthnet.gr

Abstract
On the Great Hungarian Plain and surrounding region, archaeological research on the Late Neolithic period has focused mainly on the excavation of tell settlements, while many adjacent flat-lying settlements are unrecognized or under-studied. The large horizontal extent of the flat settlements imposes unique difficulties not incurred while excavating tells; time and cost constraints often limit the amount of features that can be uncovered and other methods, such as surface collection and soil coring, offer limited information about the site's overall layout and organization. The use of magnetic survey offers a unique view of the settlement as a whole, while GIS analyses of magnetic survey data allows us to distinguish trends and anomalies in community-level organization. To explore the use of this methodology, samples of house features from six Late Neolithic to Early Copper Age sites were analyzed and compared in terms of house size, orientation, and spatial distribution.

Keywords: Magnetometry, GIS, Settlement Organization, Neolithic

1. Background

The Late Neolithic period in much of Southeastern Europe witnessed the increasing nucleation of permanent settlements, changes in the organization of households and communities, and changes in associated ideas concerning individual and group identities; these trends have roots in the Early and Middle Neolithic, beginning around 6500 BC and underwent a new sequence of socioeconomic reorganization in the mid-5th millennium BC with the advent of the Early Copper Age (Bailey, 2000; Link 2006; Yerkes *et al.*, 2007; Parkinson *et al.*, 2010). As sedentism and nucleation increased throughout the Neolithic, new social, economic, and environmental stresses emerged; the built environment, through practices in layout, architecture, and continuity took on new importance, allowing people to 'ground interpersonal relationships in explicit ways that had not be utilized before' (Bailey, 2000:75) and to some degree, mitigate the stresses of large permanent settlements (Byrd, 1994; Kienlin, 2012b). Site types range from tell, to tell-like, to single-layer, or flat,, and there are examples of these types occurring in isolation and in combinations (Kalicz and Raczky, 1987; Whittle, 1996; Kienlin, 2012b).This is congruent with an overarching theme of variation in the built environment of Neolithic Southeastern Europe, which is present at various scales, from household to community, to networks of sites (Bailey *et al.*, 1998; Bailey, 2000; Hofmann *et al.*, 2006; Kienlin, 2012b).

Past research on this period has been devoted predominantly to understanding the importance of tells, their beginnings, variability, and decline; however, this tell-centric viewpoint often neglects the importance of and relationship between fluid settlement types (Kienlin, 2012b). Because of this, many single-layer settlements remain unstudied (Kienlin, 2012b). Similarly, focusing on tell stratigraphy often results in a limited amount of excavated horizontal exposure and consequently, much more is known about household level changes over time than intrasite community-wide trends at any given time (Byrd, 1994; Whittle, 1996; Kienlin, 2012b). Any effort to gain insight into community-wide organizational trends through excavation presents disadvantages in terms of investment of time and funding not incurred when focusing on vertical stratigraphic exposure. Some studies have incorporated soil coring and systematic surface survey and collection, but these methods offer limited information about the site's layout and organization. More recently, geophysical prospection has been employed as a quick, cost-effective and non-destructive way to collect information on a large area (Chapman *et al.*, 2014; Hofmann *et al.*, 2006; Sarris, 2011; Schier and Draşovean, 2004).

Geophysical prospection utilizes active and passive remote-sensing devices to measure physical properties of the soil, detect subsurface features, and generate high resolution imagery of the findings (Sarris, 2008). Magnetic surveying techniques in particular are useful in detecting anomalies, or 'subsurface targets with magnetic properties different from those of the surrounding soil matrix', such as pits, wall trenches, hearths, kilns, ditches, burned deposits, and habitation structures (Figure 1) (Sarris *et al.*, 2004:930). Conducting a high-resolution magnetic survey over a large horizontal extent provides a unique lens with which to view the built environment at the scale of the community.

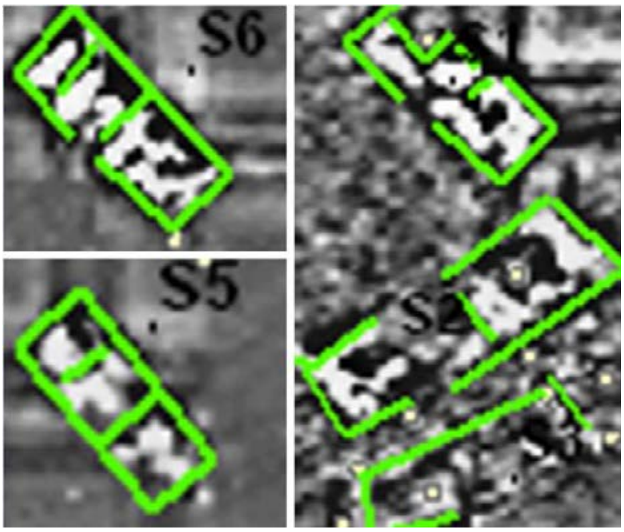


Figure 1: Magnetic anomalies.

2. Methodology

The current study stems from research being done at the Late Neolithic site complex of Szeghalom-Kovácshalom by the Körös Regional Archaeological Project; extensive geophysical prospection, excavation, and systematic surface collection of the site over the last four years has identified a surprisingly large settlement complex exhibiting a wide range of variation in terms of house size, orientation, and distribution (Parkinson and Gyucha, 2012; Duffy *et al.*, 2013). The variety in settlement organization at the Szeghalom-Kovácshalom complex is in accordance with the increased variation we might expect to coincide with the 'increased economic options of early farmers' (Flannery, 2002b:422). The goal of this study is to run similar GIS analyses on data from magnetic survey at other, roughly contemporary sites in the region and determine whether the high amount of variation in the built environment at Szeghalom-Kovácshalom, possibly evidence of household autonomy, exists elsewhere. We also explore how magnetic prospection and Geographic Information Systems can contribute to the discussion on community-wide, intrasite organization and architectural patterns of early village societies. Ideally, all of the sites examined in this study would have been Late Neolithic and/or Early Copper Age sites from the Great Hungarian Plain to limit the effect that different cultural trajectories and environments might have on settlement organization; however, magnetic survey has not yet been practiced widely enough to allow for such a fine resolution analysis. As geophysical prospection continues to increase in use, this will be more possible in coming years. The study sample for this project, then, draws on data from roughly contemporary sites in Southeastern Europe for which enough of the site's horizontal extent is accessible through either magnetic survey or excavation to allow for examination at the level of the community. These sites include Iclod, Turdaş-Luncă, Okolište, Pietrele, Polgár-Csőszhalom, and Szeghalom-Kovácshalom. (Fig. 2.1).

In order to conduct the core of the methodology (see Fig. 2.2 for flowchart), the selected magnetograms and site plans underwent preprocessing to enable their use in a Geographic Information System (GIS). Raster images generated through magnetic survey or large-scale excavation were imported into ESRI's ArcMap version 10.1. Using the scale or grid system included with each image, the raster's extent was defined relative to scale and the map document displayed in the appropriate unit (Fig. 3). In cases where appropriate spatial reference information was available, rasters were georeferenced and projected into the appropriate UTM zone. Rectilinear anomalies interpreted as house structures were traced to create polygon shapefiles which would serve as the basis for GIS analyses. Groundtruthing rectilinear magnetic anomalies through excavation to confirm the presence of Neolithic house structures has been successful in several studies (Sarris et al., 2004; Hofmann et al., 2006; Schier, 2008). During the digitization process, incomplete features, such as those lying on the survey boundary, were removed to prevent skew in data on house orientation and size.

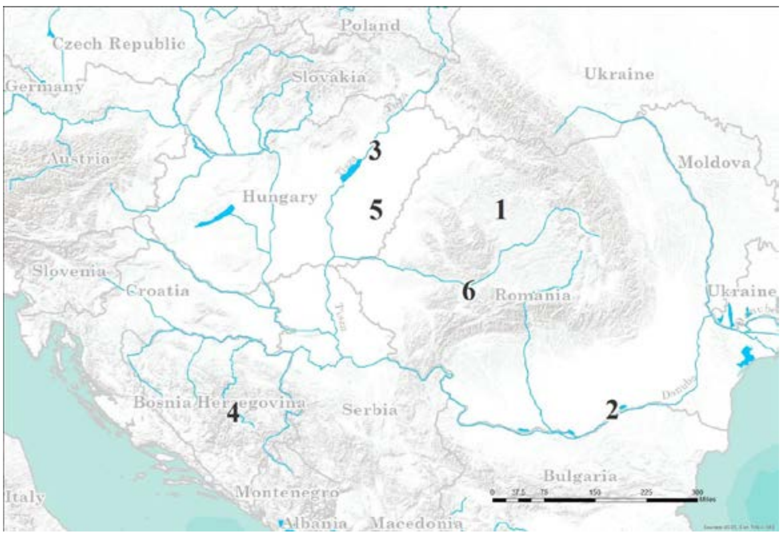


Figure 2.1: Site index map.

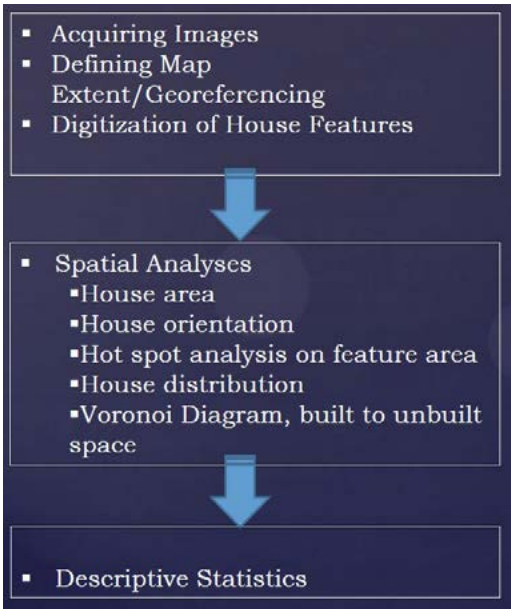


FIGURE 2.2: FLOW CHART OF METHODOLOGY.

The different core analyses conducted include: 1) calculating house size in meters squared, 2) determining house orientation in degrees from true north, 3) calculating a nearest neighbor index to classify the sample's spatial distribution as clustered, dispersed, or random, 4) running a hot spot analysis to detect clustering based on house size, and 5) generating Thiessen polygons to examine amount of unbuilt space per house feature, or packing of houses. These analyses were chosen because of their ability to generate data about the built environment based solely on the magnetogram or site plan. For some of the sites in the study, additional sources of data were limited or unavailable.

The size of houses, which is one element of what Bánffy terms the architectural canon, may be adhered to strictly or more loosely, depending on the culture, and it also may vary from group to group (Bánffy, 2013). The household is the basic unit of production and consumption in communities during the Neolithic and some factors that influence house size include the number of occupants, the social status of occupants, types of activities carried out in the home, and amount of space available to be built upon (Flannery, 2002b; Whittle, 1996). The distribution of house sizes within the community also is informative; the range and standard deviation in particular give an indication of the degree to which any norm of sizing was adhered to (see Fig. 4). To examine house sizes, the area of each house polygon was calculated in meters squared using the Calculate Areas tool from the Spatial Statistics toolbox.

Trends in building alignment and house orientation are another element of the architectural canon that has been noted to vary in terms of adherence and focus (Bánffy, 2013). A low degree of variation in feature orientation is usually considered coordinated arrangement, that

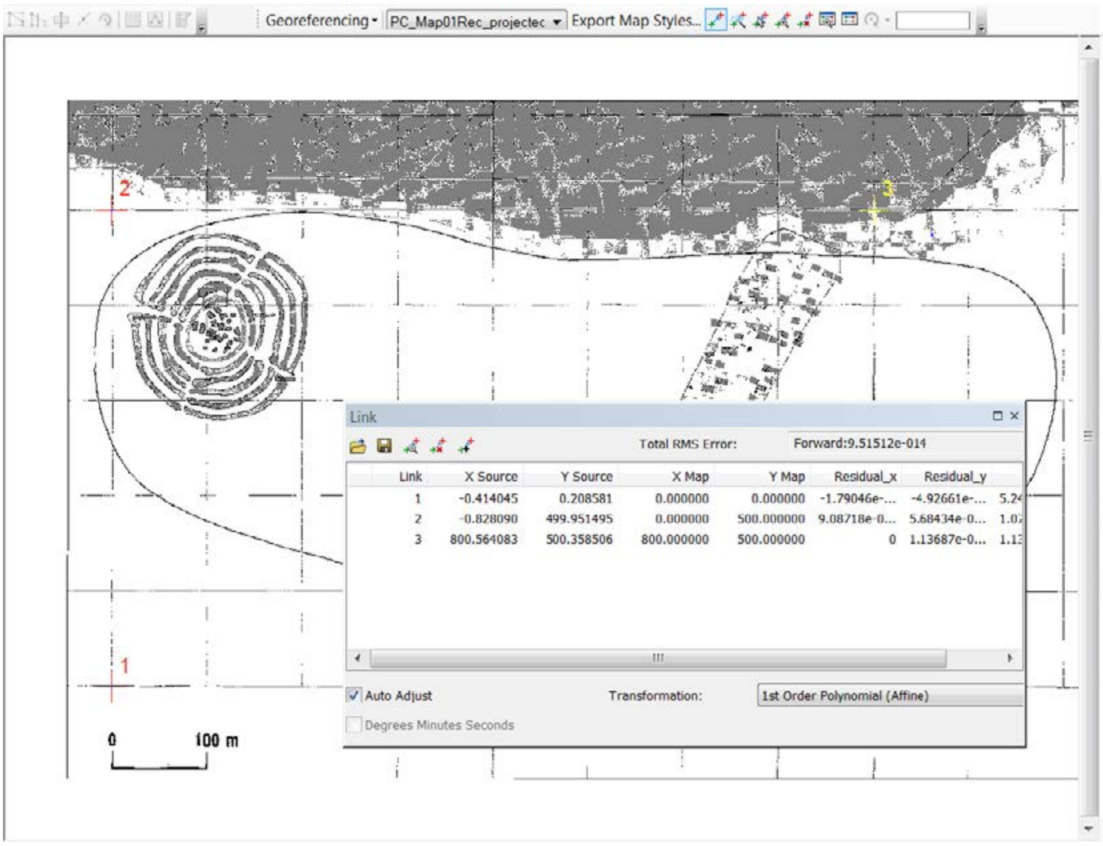


FIGURE 3: GEOREFERENCING/DEFINING IMAGE EXTENT.

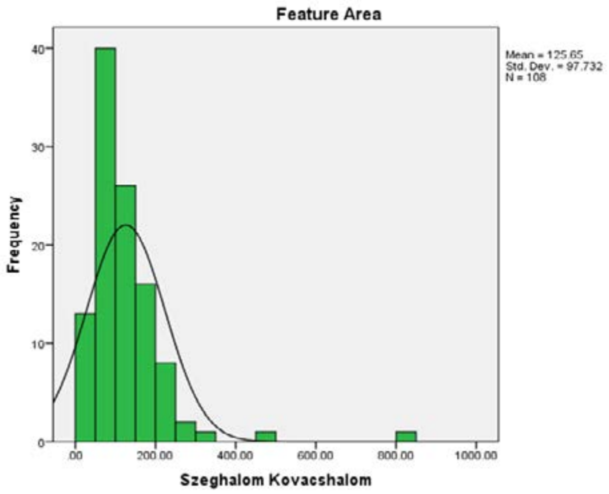


FIGURE 4: HOUSE SIZE HISTOGRAM FOR SZEGHALOM.

buildings were 'constructed with reference to one another' (Smith, 2007:8). Houses that adhere to a collective alignment to the cardinal directions, large man-made features, or landmarks in the landscape may reflect a shared communal identity or ideology (Bánffy, 2013). It is also possible that building orientation was determined by more 'practical considerations' such as wind direction or topography (Bradley, 2001:53; Marshall, 1981). To determine the dominant angle of house orientation for all features in a sample, we used the Calculate Polygon Main Angle tool from the Cartography tool set. For rotation method, we used the default, graphic, which calculates the angle of orientation in degrees from true North, with North being zero and rotation moving counterclockwise.

The next three analyses involve examining the set of houses as a whole and an attribute or location of each house in relation to its neighbors. When interpreting data from these analyses, certain limitations must be kept in mind. Magnetic survey data does not provide temporal information and cannot verify whether houses are contemporary. In studies where test excavations and soil corings have been conducted in conjunction with magnetic survey, such as the KRAP project at Szeghalom-Kovácshalom, it is possible to assess whether houses are more or less contemporaneous. Also, the built environment 'organizes, regulates, and delimits contact between individuals and households' and by doing so may inform on these social relations; however, social organization and spatial organization are not a 'one-to-one' correlation (Byrd, 1994:643).

Distance and spacing between buildings is one element that Hall (1972) suggests is socially meaningful and regulated (Byrd, 1994). Examples of socially influenced patterns in the spatial distribution of houses may include neighborhood or kinship-based clustering, autonomous random placement, or communally planned regularity and standardization (Byrd, 1994; Bánffy, 2013; Raczky 2009). On the other hand, the spatial distribution of houses may also be influenced by more practical concerns, such as the availability of space. The Average Nearest Neighbor tool from the Spatial Statistics toolbox, using the polygon shapefile of houses as input, produces a nearest neighbor index which indicates whether the sample exhibits a clustered, random, or dispersed distribution (see Fig. 5). The null hypothesis of the Average Nearest Neighbor Statistic states that the features are randomly distributed. The z-score and p-values generated with this tool are measures of statistical significance which indicate whether or not to reject the null hypothesis. The nearest neighbor index is expressed as the ratio of the Observed Mean

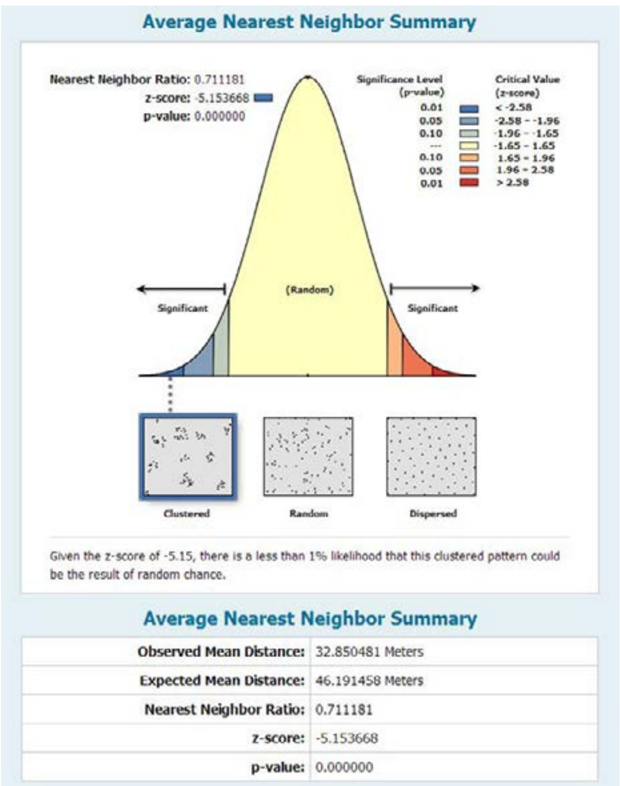


FIGURE 5: AVERAGE NEAREST NEIGHBOR RESULTS

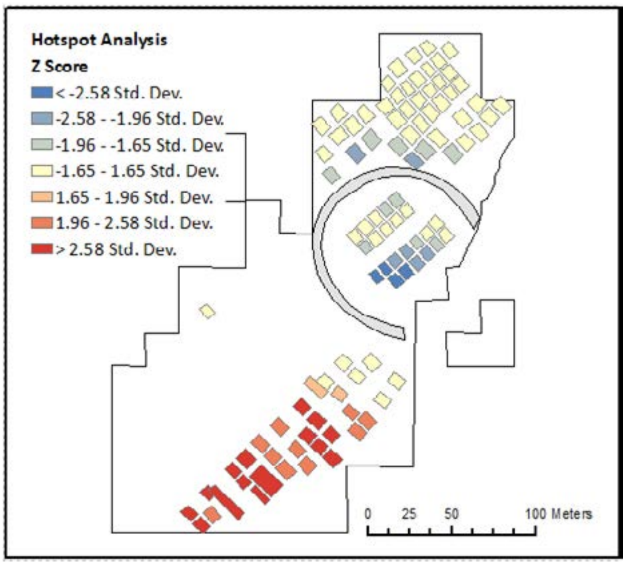


FIGURE 6: PIETRELE HOTSPOT ANALYSIS

Distance to the Expected Mean Distance; if the index value is less than one, the sample exhibits clustering. If the index value is greater than one, the sample exhibits dispersion. The study area for each is the minimum enclosing rectangle around the house features. In the case of some sites, like Szeghalom-Kovácshalom, the survey boundaries are complex and the minimum enclosing rectangle used for the average nearest neighbor statistic may not correspond well. Results in these cases should be interpreted with caution.

Theory on the creation of social space, which suggests that people can create and maintain social integration and cohesion through shared aspects of the built environment, can be applied to subgroups within a community as well (Rapoport, 1982). These subgroupings may present in clusters formed by more than simple distance-based associations; aspects of the built environment that may delineate social subgroups include orientation, size,

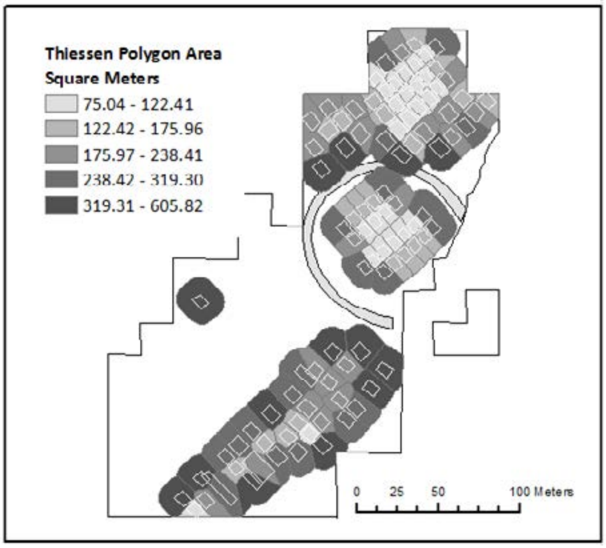


Figure 7: Iclod Thiessen polygons symbolized by area.

and ornamentation among others. In order to determine whether significant clusters based on house size were present in our samples, a Hot Spot analysis was employed using the Getis-Ord Gi* statistic, a tool located in the Spatial Statistics toolbox. This tool utilizes a weighted set of data, in this case, house area in square meters, to detect geographic clusters of high values and low values within a sample. An example of Hot Spot analysis results are featured in Fig. 6. The Getis-Ord Gi* statistic is a distance-based measure of spatial association which examines each feature within the context of a neighborhood of features; the feature is assigned a statistically significant z-score result when values for the feature and its neighborhood deviate enough from the values that would be expected if the values were a result of random chance. Use of this tool is most reliable with a sample of 30 features or larger.

The third analysis that examines houses in reference to its neighbors provides a more nuanced view of the density, or packing, of houses. By utilizing Voronoi diagrams, we were able to approximate the amount of unbuilt space

Figure 8: Iclod magnetogram from Mischka, 2012:157.

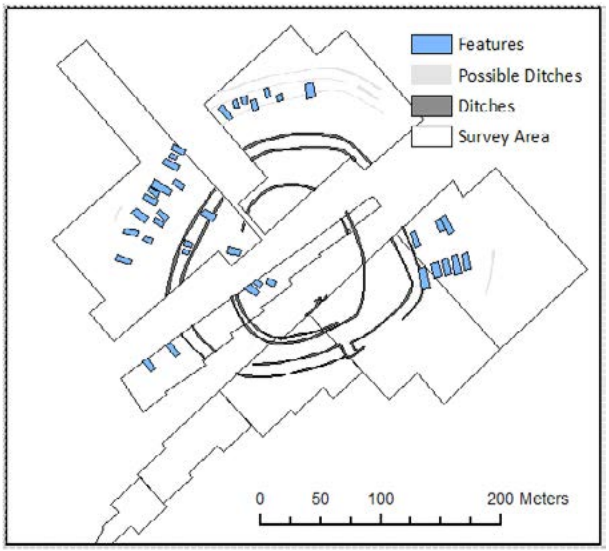


Figure 9: Iclod digitized features.

available to each house, and in doing so, obtain an idea of how densely house features were situated and how this varied throughout the settlement. A Voronoi diagram, also called Thiessen polygons, is a set of adjoining polygons, created from a series of points, whose edges are exactly half the distance to the nearest feature in which all space within the polygon is closer to the polygon's associated point than any other point (Kvamme, 1999). This study used the Create Thiessen Polygons tool from the Analysis toolbox and the centroid of each house as the input set of points (see Fig. 7). Because the magnetic survey boundaries sometimes extend beyond the settlement, buffers were created around each individual house and dissolved into one large buffer zone. To standardize how we obtained the radius of the buffer zone for each site, we used the Expected Mean Distance value derived from the Average Nearest Neighbor tool. This helps eliminate skew in the data, however, it also imposes a limit to the polygons which may not have any realistic value. Afterwards, the area in meters squared was calculated for each polygon in the same way it was calculated for house features.

All of the data generated through these analyses were exported into IBM's SPSS Statistics software, version 20. Using SPSS, descriptive statistics were compiled for each feature attribute of each site, and histograms and boxplots were created to examine intra- and inter-site trends for each variable. All images included in this paper that pertain to the GIS analyses run for the project were created using ArcGIS 10.1.

3. Case Studies

Iclod is a Late Neolithic site of the Zau culture located on a low terrace near the Someşul Mic river in Cluj county, Transylvania. Iclod was occupied by groups of the Iclod phase (approximately 4800 BC – 4500 BC), part of the late Zau culture, and later on the newer Petreşti culture (Mischka, 2012). Excavations first began at the site in the 1970's but have focused mainly on the graveyard and site center (Mischka, 2012; Lazarovici 1991). In 2008 and 2011, a German team led by Carsten Mischka explored other areas of the site with magnetic survey and test excavations (Mischka, 2012). The geophysical prospection has surveyed nearly the entire site which covers more than 11 hectares (see Fig 8). The survey detected a three-part ditch system surrounding the inner part of the site, with a possible gate construction in the South, 37 house structures, and traces in the North of what might be smaller ringed ditches surrounding the entire Neolithic settlement (Mischka, 2012). House features overlap with the ditch system in the West and the possible most outer ditch system in the North, indicating that not all of these features are contemporary. Little is known about the chronology of the site, but Mischka believes that the inner ditch and house features within the ditch are likely the oldest features (Mischka, 2012). See Fig. 9 for the digitized version of the magnetogram.

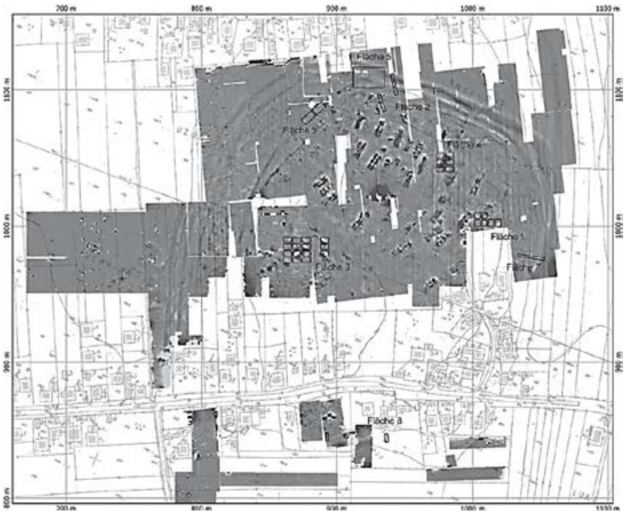


Figure 10: Okolište magnetogram from Müller et al. 2011:83.

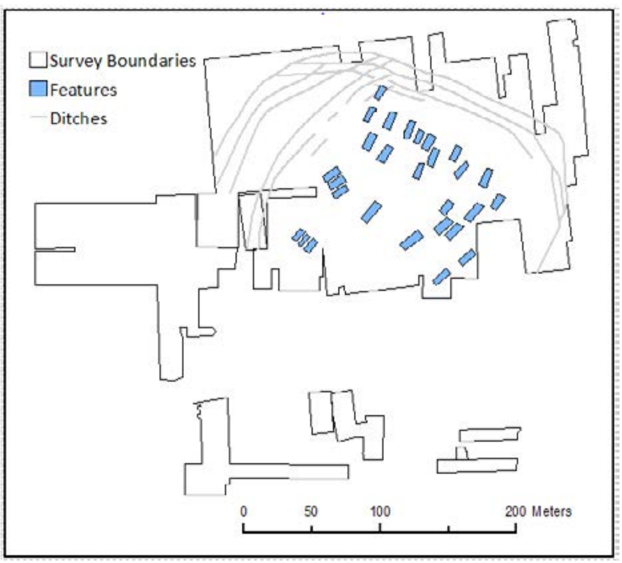


Figure 11: Okolište digitized features.

Okolište is a Late Neolithic site located in the northern part of the Visoko basin in Central Bosnia. Mountains are located farther north of the settlement which circle the Visoko Basin and have elevations of up to 1,000 meters in some places. Okolište is located on a river terrace adjacent to the Bosna river and measures approximately 7.5 ha, much larger than nearby contemporary sites (Hofmann *et al.*, 2008). A Bosnian-German research team began studying the site in 2002 (Hofmann *et al.*, 2008). Magnetic survey was carried out by T. Schüler in 2003 (Fig. 10). The survey detected a complex system of overlapping ditches as well as 28 longhouse features within the ditch system. The ditch system begins as one strand in the east, nearest the river, and gradually separates into three strands in the north, and then into five, possibly six strands, in the Northwest and West. Hofmann and colleagues believe that at least two of these strands in the North were contemporary. Ramparts and a palisade were discovered during excavation (Hofmann *et al.*, 2008). Okolište is thought to be occupied from approximately 5200 to 4500 BC and to have begun declining fairly early on, with the outermost strands being the oldest fortifications (Hofmann et al., 2008). The fortifications gradually moved towards the river through processes of filling in and renewal up until

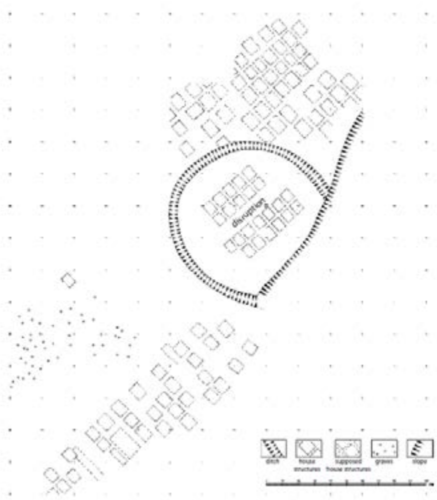


Figure 12: Pietrele magnetogram interpretations from Reingruber et al.2010:174.

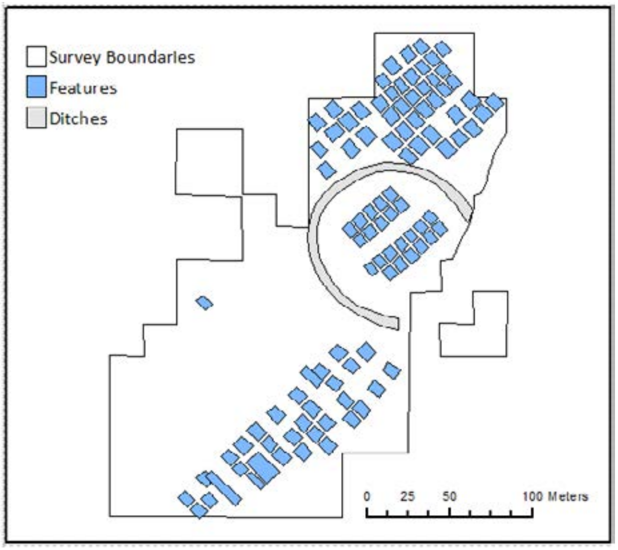


Figure 13: Pietrele digitized features.

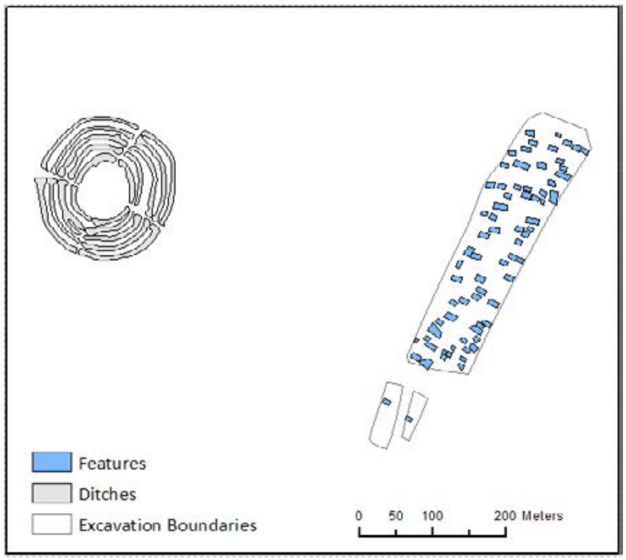


Figure 15: Polgár-Csőszhalom digitized features.

around 4900 BC (Hofmann *et al.*, 2008). Stratigraphy of the site core has proven through excavation to be complex and house features discovered by the magnetic survey are from various phases as well as interspersed with unburnt houses not detected by the magnetometer (Hofmann et al., 2008). The process of decline visible at Okolište is considered to be part of the wider trend of abandonment of large, well-established sites and subsequent dispersal, signaling socioeconomic reorganization (Hofmann *et al.*, 2008). See Fig. 11 for the digitized version of the magnetogram.

Pietrele - Măgura Gorgana, located near the modern village of Pietrele in southern Romania, in the Lower Danube region, is an Eneolithic site assigned to the Kodžadermen-Gumelniţa-Karanovo VI complex and dates from approximately 4600-4250 cal BC, a period when most other tells in Southeastern Europe, with the exception of tells in Bulgaria, were already abandoned (Hansen and Toderaş, 2012; Reingruber *et al.*, 2010). The settlement is located on a low river terrace, 8 km from the Danube River in an area that was meadow in prehistoric times but has since gained approximately 6 meters in elevation through sediment deposition (Hansen *et al.*, 2009:289). Magnetic survey was conducted on the tell in 2004 by B. Song and detected four parallel rows of houses (Hansen *et al.*, 2009). Areas surrounding the tell were the focus of magnetic survey in 2005, which detected a flat settlement as well as a graveyard to the tell's Southwest (Hansen et al., 2009). Boundaries of the site have not yet been reached, although distinct spaces of dense occupation and culturally sterile areas were noted in the magnetogram (Fig. 12) (Hansen *et al.*, 2009). Through remote sensing, the team was able to detect a system of ditches on the highest terrace a short distance from the tell (Hansen *et al.*, 2009). Radiocarbon dating and excavation have indicated that the house features depicted in the magnetogram may not all be contemporary; superimposed houses and different occupations have been uncovered (Hansen and Toderaş, 2012). See Fig. 13 for the digitized version of the magnetogram.

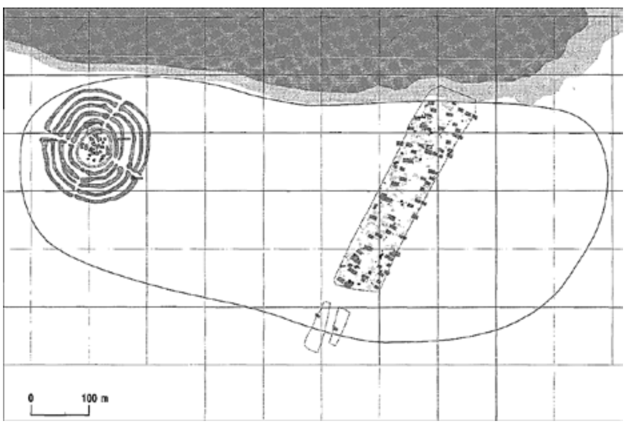


Figure 14: Polgár-Csőszhalom site plan from Raczky and Anders, 2008:40.

The Late Neolithic site Polgár-Csőszhalom is 'the northernmost tell in the Carpathian Basin' and is located next to a former meander of the Tisza River (Raczky and Anders, 2008). The site is comprised of a Tisza-Herpály type tell, Lengyel type five-ring concentric ditch and an associated flat settlement discovered during rescue excavations conducted approximately 500 meters from

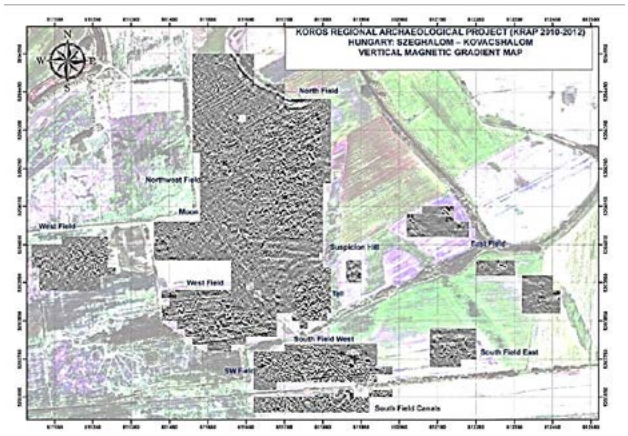


Figure 16: Szeghalom-Kovácshalom magnetogram.

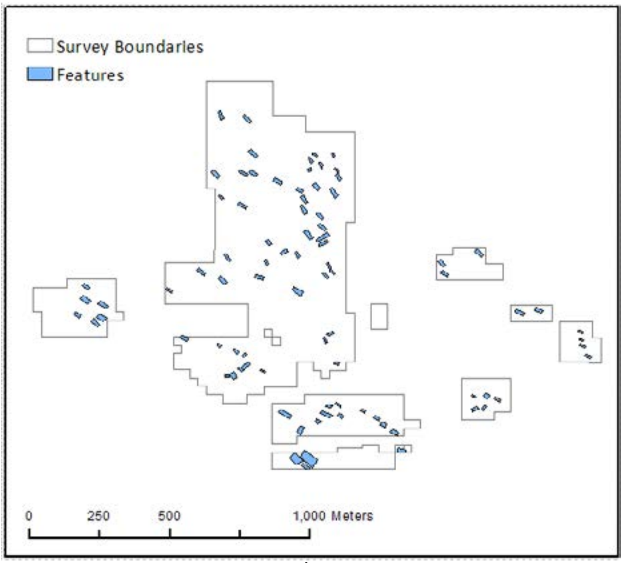


FIGURE 17: SZEGHALOM-KOVÁCSHALOM DIGITIZED FEATURES.

the tell (Raczky *et al.*, 2007). Based on excavation, geophysical prospection and soil coring, the entire site is estimated to cover 28 ha, with the flat settlement accounting for 24 ha (Raczky and Anders, 2010). The tell itself is thought to have been occupied from 4820 to 4530 BC and the associated flat settlement from 4830 and 4600 BC (Raczky and Anders, 2010). The results of the excavation done prior to the building of the M3 motorway were combined with prior excavations and the 1989's geophysical prospection of the tell and ditches to form a site plan. This site is unique in being the only site included in the project with data on house features derived from large-scale excavation rather than magnetometry (Fig. 14). Using this site plan, 73 house structures were digitized, which compares well with the excavation results in which 79 houses were excavated (Raczky and Anders, 2010). Some interpretations had to be made as house structures were not always clearly delineated in the image. At least thirteen houses were detected on the tell through magnetometry and excavation with an artifact assemblage that suggested a ritual function, and additional buildings, wells, and pits were discovered in both sections but are not included in this analysis (Raczky and Anders, 2010). See Fig. 15 for the digitized version of the site plan.

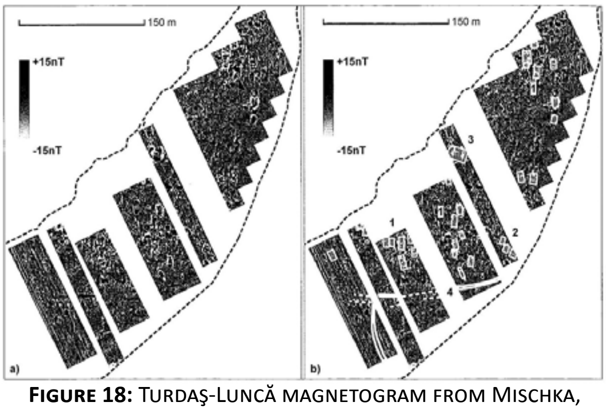


FIGURE 18: TURDAŞ-LUNCĂ MAGNETOGRAM FROM MISCHKA, 2012:165.

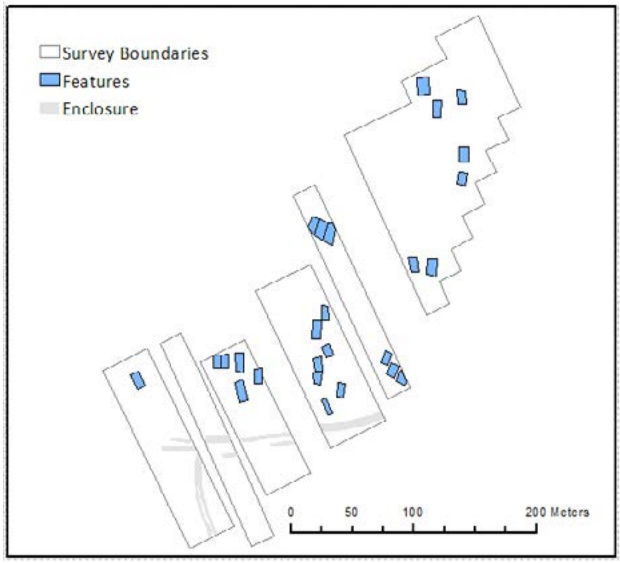


FIGURE 19: TURDAŞ-LUNCĂ DIGITIZED FEATURES.

Szeghalom-Kovácshalom is a settlement complex covering nearly 90 hectares; the complex is comprised of a relatively small tell surrounded by an extensive flat site and isolated clusters of houses. The tell and some parts of the surrounding flat site have origins during the Szakálhát phase of the Middle Neolithic (5200-5000 B.C., cal.) and underwent expansion during the Late Neolithic Tisza phase (5000-4500 B.C., cal.). Small-scale reoccupation of the tell occurred after a hiatus, during the Early Copper Age, as well as minor occupation episodes of the flat settlement during the Middle Neolithic, Late Copper Age and Bronze Age. The Szeghalom-Kovácshalom site is located near the modern village of Szeghalom and just North of the modern Sebes-Körös River channel. The Körös Regional Archaeological Project has been systematically investigating the Szeghalom-Kovácshalom complex since 2010. This project has included reconnaissance, systematic survey, systematic topographic study, geochemical analysis, magnetic susceptibility, geophysical prospection, and test excavations on the tell and surrounding flat settlement and isolated houses. The magnetic survey, directed by A. Sarris, has covered over 51 hectares (Fig. 16) (Sarris and Papadopoulos, 2012). Unlike at Early Copper Age sites studied by the Körös Regional Archaeological Project, the magnetic survey did not detect any traces of a large surrounding ditch system, although it is likely that the presence of a paleomeander surrounding the tell made the construction of a ditch system unnecessary (Sarris, 2012). Features were digitized according to a line shapefile consisting of house features identified by Sarris and the KRAP team. Eighty-seven house features total were digitized. The amount of features digitized on the tell is misleading, indicating a lower density than reality, due to the difficulty of delineating anomalies in an area with so much occupational overlap. The same problem with distinguishing features occurs in the site's South but to a lesser degree. See Fig. 17 for the digitized version of the magnetogram.

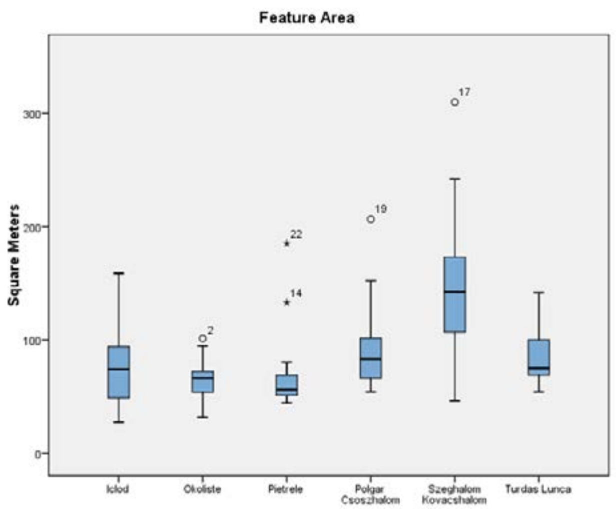


Figure 20: Boxplots displaying house sizes across sites

Turdaş-Luncă is a site in southern Transylvania with occupations dating to the Neolithic, Eneolithic and Bronze Age and is the eponymous site for the Turdaş culture (5000-4800 B.C. cal.) (Mischka, 2012). The settlement is located on a low terrace pinned between the Mureş River in the North and a small hill in the South, a setting which is typical for Transylvanian Late Neolithic settlements (Mischka, 2012). The site encompasses the entire terrace which is 14 hectares and has been subjected to severe erosion in the north, as much as 80 meters, due to its proximity to the river (Mischka, 2012). A large motorway was constructed to the South of the site and prompted rescue excavation in 2011 (Mischka, 2012). The site is referred to as a tell site; however, the vertical growth and multiple layers are due to repeated fluvial inundations (Mischka, 2012). No significant ditch system was detected through magnetic survey; however, Mischka interpreted traces of an enclosure in the south end of the terrace from the magnetogram (Mischka, 2012). In the past, excavation has had a northern focus because of the continuing erosion. In 2007, Carsten Mischka led a geomagnetic survey of the site center as a part of a larger research project (Fig. 18). According to Mischka, it is possible clusters of houses detected in the survey belonged to different periods or that the clusters are contemporary and connected by unburned houses (Mischka, 2012). Mischka was able to detect 27 house structures through magnetic survey; however, three structures were partial and left out of digitization. After consulting A. Sarris on the original magnetogram, the anomaly Mischka interpreted as one central anomalous building was broken down into three northeast to southwest oriented adjacent features (Fig. 19). This interpretation is more congruent with orientation and dimensions of all other structures in the sample.

4. Results

The analysis conducted to calculate house size in meters squared revealed that the range of variation present in the Szeghalom-Kovácshalom distribution is much larger than that of every other site in the study (Fig. 20). Szeghalom-Kovácshalom has not only the largest range, but the largest average house size and largest structure overall of all the sites. Iclod has the next largest range of values that does not include outliers, but has a mean that is relatively consistent with feature area means for Okolište, Polgár-Csőszhalom, and Turdaş-Luncă. Pietrele has the smallest range of house sizes with a few significant outliers and a mean that is slightly lower than that of the other sites. Interestingly, the interquartile range of house sizes for all sites but Szeghalom-Kovácshalom fit within a span of 50 to 100 meters squared.

The results from the analysis conducted to determine house orientation in degrees from true north, indicated that all possible orientations are represented within the sites examined (Fig. 21). This includes, in general, north to south at Turdaş-Luncă, west to east at Polgár-Csőszhalom, northeast to southwest at Okolište, and variations of northwest to southeast for Iclod, Pietrele, and Szeghalom-Kovácshalom. The analysis also verified differences between sites in how strictly building alignments were adhered to. Pietrele has the smallest range of observed

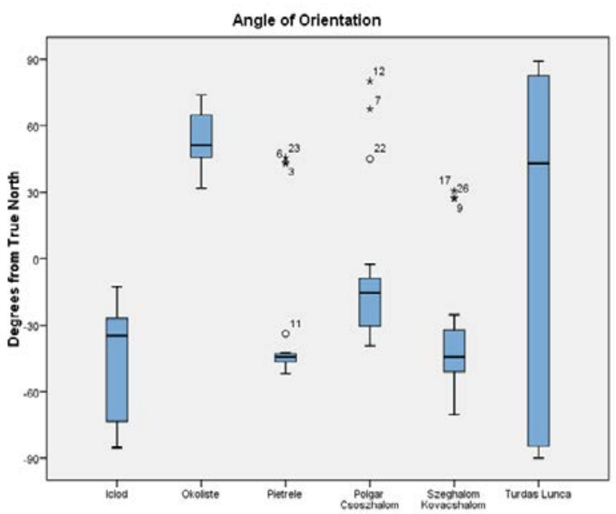


Figure 21: Boxplots displaying house orientations across sites

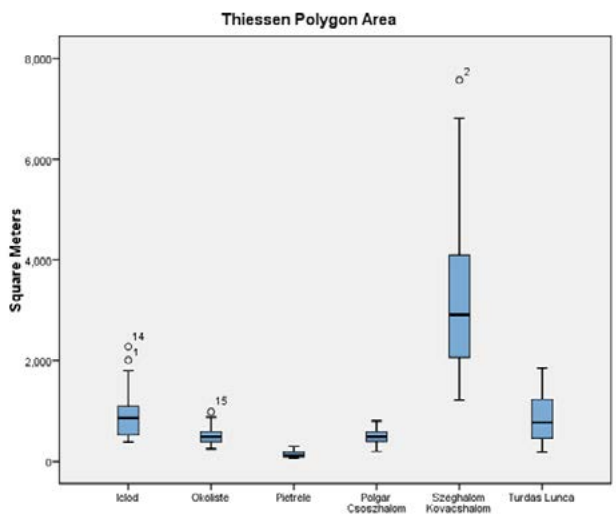


Figure 22: Boxplots displaying Thiessen polygon area across sites

orientation angles with a couple significant outliers, while the site with the largest range in variation is Iclod, due to its interesting semi-radial alignment pattern. The range of values for Turdaş-Luncă is misleading because values hover at either end of the spectrum, 90 and -90. In comparison, Okolište, Szeghalom-Kovácshalom, and Polgár-Csőszhalom are very similar in range when not taking outliers into account.

According to the Average Nearest Neighbor analysis, Iclod, Turdaş-Luncă, and Szeghalom-Kovácshalom all exhibit a clustered distribution of houses. The analyses for Okolište and Polgár-Csőszhalom classified these distributions as dispersed and the analysis for Pietrele as a random distribution. The results of this analysis may be affected by several factors and these need to be taken into account during interpretation. For some of the sites, excavation has indicated that the houses detected in the magnetogram may be an under-representation or over-representation of the settlement at a given point in time; this is possible if the site has multiple occupation horizons or if some houses were not detected by the magnetometer. Lastly, the analysis can only offer one classification per site when in reality a site may be occupied more variably and present more than one type of distribution.

The hotspot analyses were able to detect significant clusters of either statistically large or small houses for every site in the study. The analysis for Iclod highlighted a group of structures in the East as a cluster of high values and a few structures in the North for extreme low values. The analysis for Okolište highlighted a group of three structures in the Southeast as a cluster of low values; however, the sample size for Okolište does not meet the recommended sample size of 30. The Pietrele hot spot analysis illustrated the most distinct clustering based on size, with larger structures located South of the tell, smaller structures on or just North of the tell and average sized structures to the North. The analysis for Polgár-Csőszhalom highlighted only small grouping of structures in the excavation zone's South. The Szeghalom-Kovácshalom analysis featured only one distinct cluster of large houses in the Southern portion of the site. Turdaş-Luncă, another site that has a sample size below the appropriate 30, exhibited a cluster of smaller houses in the middle zones of the survey.

The analysis which calculated the area in square meters of Thiessen polygons indicated that Szeghalom-Kovácshalom, again, has the largest range, largest mean and largest maximum values (Fig. 22). In other words, Szeghalom-Kovácshalom ranges the most in amount of spacing between houses and houses tend to be spaced apart more on average. This is caused by the presence of both dense clusters of houses and dispersed isolated structures within the site and the very large geophysical survey area, relative to other sites examined. Pietrele, again, has the smallest range in variation, and smallest average amount of space between houses. Iclod and Turdaş-Luncă appear to have similar distributions and were both calculated to exhibit clustered distributions; in both of these sites, the largest Thiessen polygon areas are located on the edge of clusters. Okolište and Polgár-Csőszhalom also have very similar distributions with very little variation in the amount of space between houses; this is congruent with the dispersed classification that each received in the Average Nearest Neighbor analysis.

5. Interpretations and Conclusions

In summary, we found that in all but house orientation, the variation at Szeghalom-Kovácshalom was unmatched by all other sites. The houses at Szeghalom-Kovácshalom vary greatly in terms of size and amount of space between neighbors, exhibit clustering with and without house size being taken into account, and exhibit a relatively average range of orientation. The houses at Pietrele, on the opposite end of the spectrum, exhibit very little variation in terms of size, orientation, and space between neighbors, although there are statistically significant clusters based on house size. Patterns at the rest of the settlements fall in between these two extremes, exhibiting some variation in one aspect and more standardization in another.

While each of the patterns we examined is interesting and unique in terms of community organization, it might be most useful for the purposes of this article to focus on the extremes of the spectrum, conceptualized as household autonomy on the one end and community standardization on the other. The large range of variation detected at Szeghalom-Kovácshalom, suggests that households were autonomous in the decisions they made about their built environment. Although buildings generally align NW to SE, there is no consistency in spacing or size, let alone standardization, even though these houses are all believed to be roughly contemporary (Parkinson and Gyucha, 2012). The pattern at Pietrele on the other hand suggests standardization in size and an emphasis on regularity and planning in orientation and spacing. There is a greater degree of variation in size and spacing in the south, but it is still less than might be expected when considering that houses represent a mix of various phases of occupation (Hansen and Toderaş, 2012). The community exhibits very strong norms in the built environment and evidence for continuity, or maintaining this pattern of organization over time (Hansen and Toderaş, 2012). It may be that a strong community identity or standardization of the built environment helped to mitigate the stress of a large permanent settlement, in turn, allowing it to be successful for several generations. At Szeghalom-Kovácshalom, the large degree of variation suggests a lack of strong community identity, which is interesting in light of the site's relatively short span of occupation.

In the end, it is through the fusion of multiple lines of evidence, i.e. surface collection, topographic survey, geophysical prospection, and excavation, that these patterns can best be understood. What magnetometry and GIS offer is a uniquely valuable starting point for examining the built environment at a scale that until recently, has been difficult to achieve. While magnetometry provides a window at the level of the site and GIS allows us zero in on trends and anomalies, together they offer a means to target for excavation houses or other structures that will be most useful in informing on intra-site organization in early village societies.

References

BAILEY, D. W. (2000). *Balkan Prehistory: Exclusion, Incorporation and Identity*. London: Routledge.

BRADLEY, R. (2001). Orientations and origins: a symbolic dimension of the long house in Neolithic Europe. *Antiquity* 75, pp. 50–56.

BÁNFFY, E. (2013). Tracing the Beginning of Sedentary Life in the Carpathian Basin. In: Hofmann, D., Smyth, J. (Eds.). *Tracking the Neolithic House in Europe: Sedentism, Architecture, and Practice*. pp. 117-150.

BYRD, B. 1994. Public and private, domestic and corporate: the emergence of the Southwest Asian village. *American Antiquity* 59, pp. 639–666.

DUFFY, P. R., PARKINSON, W. A., GYUCHA, A. and YERKES, R. W., (2013). Coming Together, Falling Apart: A Multiscalar Approach to Prehistoric Aggregation and Interaction on the Great Hungarian Plain. In: Birch, J. (ed.). *From Prehistoric Villages to Cities: Settlement Aggregation and Community Transformation*. New York: Routledge. pp. 44-62.

FLANNERY, KENT V., (1972). The Origins of the Village as a Settlement Type in Mesoamerica and the Near East: A comparative study. In P. J. Ucko, R. Tringham and Dimbleby (Eds.). *Man, Settlement and Urbanism*, London: Duckworth, pp. 23-53.

FLANNERY, KENT V., (2002a). Prehistoric Social Evolution. In: Peter N. Peregrine, Carol R. Ember, and Melvin Ember (eds.). *Archaeology: Original Readings in Method and Practice,* Upper Saddle River, NJ: Prentice-Hal, pp. 225-244. l.

FLANNERY, KENT V., (2002b). The Origins of the Village Revisited: From Nuclear to Extended Households. *American Antiquity* 67(3), pp. 417-433.

HALL, E. T. (1972). Silent Assumptions in Social Communication. In: R. Gutman (ed.). *People and Buildings*, New York: Basic Books, pp. 135-151.

HANSEN, S., and M. TODERAS (2012). The Copper Age Settlement Pietrele on the Lower Danube River (Romania). Landscapes and Human Development: The Contribution of European Archaeology. Socio-Environmental Dynamics over the Last 12,000 Years: The Creation of Landscapes (1st - 4th April 2009), pp. 127-138.

HANSEN,S., M. TODERAS, and A. REINGRUBER (2009). Pietrele. A Chalcolithic Settlement on the Lower Danube River. In: F. Draºovean, D. Leopold Ciobotaru and M. Maddison (eds.) *Ten Years After: The Neolithic of the Balkans, as Uncovered by the Last Decade of Research*, Timisoara: Museum of Banat Publications pp. 289-310.

HOFMANN, R., Z. KUJUNDZIC-VEJZAGIC, J. MUELLER, N. MUELLER-SCHEESSEL, and K. RASSMANN. (2006). Prospektionen and Ausgrabungen in Okolište (Bosnien-Gerzegowina). Siedlungsarchaeologische Studien zum zentral bosnischen Spaetneolithikum (5200-4500 v. Chr.). *Bericht der Romisch-Germanischen Kommission* 87, 2006, pp. 41-212

HOFMANN, R., Z. KUJUNDZIC-VEJZAGIC, J. MUELLER, N. MUELLER-SCHEESSEL, and K. RASSMANN. (2008). Excavations in Okolište and the Reconstruction of Late Neolithic Settlement Processes in the Visoko Basin in Central Bosnia (5200-4500 B.C.). Electronic documnt, http://www.aegeobalkanprehistory.net/article.php?id_art=12.

HOFMANN, R. (2012). Style and Function of Pottery in Relation to the Development of Late Neolithic Settlement Patterns in Central Bosnia. In: R. Hofmann, F.-K. Moetz, J. Müller (Eds.). *Tells: Social and Environmental Space. Proceedings of the International Workshop 'Socio-Environmental Dynamics over the Last 12,000 Years: The Creation of Landscapes II* (14th–18th March 2011)' in Kiel. Volume 3. Universitätsforschungen zur prähistorischen Archäologie 207, Bonn: Habelt, pp. 181-201.

KIENLIN, TOBIAS L., (2012). Patterns of Change, Or: Perceptions Deceived? Comments on the Interpretation of Late Neolithic and Bronze Age Tell Settlement in the Carpathian Basin. In: Kienlin T. L. and Zimmermann A. (eds.), *Beyond Elites: Alternatives to Hierarchical Systems in Modeling Social Formations*, Universitätsforschungen zur prähistorischen Archäologie 215, Bonn: Habelt , pp. 251-310.

KVAMME, KENNETH L., (1999). Recent Directions and Developments in Geographical Information Systems. *Journal of Archaeological Research* 7(2), pp. 153-201.

MARSHALL, A. (1981). Environrnental adaptation and structural design in axially-pitched long houses from Neolithic Europe, *World Archaeology* 13, pp. 101-121.

MERLINI, M. (2009). *An Inquiri into the Danube Script.* Sibiu: Alba Iulia – Editura Altip

MISCHKA, C., (2010). Beispiele für Änlichkeit Und Diversität Neolithischer Und Kupferzeitlicher Regionaler Siedlungsmuster in Rumänien Anhand Von Geomagnetischen Prospektionen. In: Hansen Sv. (ed.). *Leben auf dem Tell als soziale Praxis, Beiträge des Internationalen Symposiums in Berlin vom 26.–27. Februar 2007, Kolloquien Zur Vor- Und Frühgeschichte 14,* Bonn: Habelt, pp. 71-84.

MISCHKA, C., (2012). Late Neolithic Multiphased Settlements in Central and Southern Transilvania: A Geophysical Survey and Test Excavation. In: R. Hofmann, F.-K. Moetz, J. Müller (Eds.). *Tells:*

Social and Environmental Space. Proceedings of the International Workshop 'Socio-Environmental Dynamics over the Last 12,000 Years: The Creation of Landscapes II (14th–18th March 2011)' in Kiel. Volume 3. Universitätsforschungen zur prähistorischen Archäologie 207, Bonn: Habelt, pp.153–166.

PARKINSON, W. A., (2006). *The Social Organization of Early Copper Age Tribes on the Great Hungarian Plain.* British Archaeological Reports B.A.R. International Series 1573. B.A.R., Archaeopress, Oxford.

PARKINSON, W. A., (2006). Tribal Boundaries: Stylistic Variability and Social Boundary Maintenance during the Transition to the Copper Age on the Great Hungarian Plain. *Journal of Anthropological Archaeology* 25(1), pp. 33-58.

PARKINSON, W. A., and GYUCHA, A., (2012). Tells in Perspective: Long-Term Patterns of Settlement Nucleation and Dispersal in Central and Southeast Europe. Tn: R. Hofmann, F.-K. Moetz, J. Müller (Eds.). *Tells: Social and Environmental Space. Proceedings of the International Workshop 'Socio-Environmental Dynamics over the Last 12,000 Years: The Creation of Landscapes II* (14th–18th March 2011)' in Kiel. Volume 3. Universitätsforschungen zur prähistorischen Archäologie 207, Bonn: Habelt, pp.105-116.

PARKINSON, W. A., YERKES, R.. W., GYUCHA, A., SARRIS, A., MORRIS, M. and SALISBURY, R., (2010). Early Copper Age Settlements in the Körös Region of the Great Hungarian Plain. *Journal of Field Archaeology* 35, pp. 164-183.

RACZKY, P., and A. ANDERS. (2010). Activity Loci and Data for Spatial Division at a Late Neolthic Site-Complex (Polgár-Csöszhalom: A Case Study). Kolloquien Zur Vor- Und Frühgeschichte 14(Leben auf dem Tell als soziale Praxis):143-164.

RACZKY, P. and A. ANDERS (2008). Late Neolithic Spatial Differentiation at Polgár-Csöszhalom, Eastern Hungary. In: Bailey, D. W.–Whittle, A.–Hofmann, D. (eds) *Living Well Together? Settlement and materiality in the Neolithic of south-east and central Europe.* Oxford: Oxbow Books, pp. 35-53.

RACZKY, P., L. DOMBORÓCZKI, and Z. HAJDÚ (2007). The Site of Polgár-Csöszhalom and its Cultural and Chronological Connections with the Lengyel Culture. In Janusz K. Kozlowski and Pal Raczky (eds.). *The Lengyel, Polgár and Related Cultures in the Middle/Late Neolithic in Central Europe*, Kraków: Jagiellonian University and ELTE, pp. 49-70.

REINGRUBER, A., S. HANSEN, and M. TODERAS (2010). Monumental Living: Pietrele Near the Lower Danube River in the 5th Millenium BC. Landscapes and Human Development: The Contribution of European Archaeology Socio-Environmental Dynamics over the Last 12,000 Years: The Creation of Landscapes (1st-4th April 2009), pp. 171-180.

SARRIS, A., GALATY , , M. L., YERKES , R. W. , PARKINSON , W. A., GYUCHA , A. , BILLINGSLEY, D. M. and TATE, R., (2004). Geophysical prospection and soil chemistry at the Early Copper Age settlement of Vésztó-Bikeri, Southeastern Hungary, *Journal of Archaeological Science*, 31(7), pp. 927-939.

SARRIS, A., (2008). Remote Sensing Approaches/ Geophysical, In: D. M. Rearsall (ed.). *Encyclopedia of Archaeology*, New York: Academic Press, pp. 1912-1921.

SARRIS, A., (2012). Multi+ or Manifold Geophysical Prospection? *Computer applications and Quantitative methods in Archaeology International Conference - CAA2012*, University of Southampton.

SMITH, M. E. (2007). Form and Meaning in the Earliest Cities: A New Approach to Ancient Urban Planning. *Journal of Planning History* 6(1), pp. 3-47.

WHITTLE, A. (1996). *Europe in the Neolithic. The creation of new worlds.* Cambridge: Cambridge University Press.

YERKES, RICHARD W., ATTILA GYUCHA, and WILLIAM A. PARKINSON, (2009). A Multi-Scalar Approach to Modeling the End of the Neolithic on the Great Hungarian Plain Using Calibrated Radiocarbon Dates. *Radiocarbon* 51, pp. 890-932.

YERKES, R., W., SARRIS, A., FROLKING, T., PARKINSON, W. A., GYUCHA, A., HARDY, M., and CATANOSO, L., (2007). Geophysical and Chemical Investigations at Two Early Copper Age Settlements in the Körös River Valley, Southeastern Hungary. *Geoarchaeology: An International Journal* 22(8), pp. 845-871.

Historic Forest Change: New Approaches to Land Use Land Cover

Charlotte E. STANCIOFF,[1] Robert G. PONTIUS Jr,[2]
Scott MADRY,[2] Elizabeth JONES[2]

[1]Faculty of Archaeology, Leiden University, The Netherlands
[2]School of Geography, Clark University, USA
[3]University of North Carolina- Chapel Hill, USA
stancioff@arch.leidenuniv.nl, rpontius@clarku.edu,
smadry@email.unc.edu, elizabeth.anne.jones@unc.ed

Abstract

Using historic maps as significant data in environment and urban planning has provided key details of land practices and planning of the past. (Baily, 2007; Eremiasova and Skokanova, 2009; Gasperi, 2007; Grosso, 2009) While there has been growing research into the use and accuracy of historic maps, there has been little research into how to incorporate real data from historic maps into modern land cover land use change analysis. Using historic maps to create a time series of changes allows for research into the visualization and quantification landscapes through time. In this study, the land cover change of the surrounding Uxeau Commune of Burgundy France, a site rich in archaeology and history, is explored through a time series created from historic maps, aerial photos, and recent topographic maps. Covering 163 years, this time series is constructed to measure the change in forest pattern land change of forest to non-forest by using GIS methods of georeferencing, digitizing, as well as Intensity analysis, as a platform to investigate the land use/ land cover change found in the region from 1840-2003.

Consideration into preprocessing of data layers to enable comparison must be followed. This in turn raises questions of methodology to best extract the data of a historic map. As such, this study investigates the conversion of historic attributes and features to modern counterparts, building on the existing research of georeferencing, digitalization and eventual extraction of historic maps for further analyze. Following these steps, attention is turned to how best analyze the changing landscape patterns. Intensity analysis, developed by S.Z. Aldwaik and Robert Pontius, is used to understand the interval change between forest loss, gain and persistence annually throughout the study site.

The integration of historic and modern data to create the observed pattern of land cover change could aid in subsequent studies on extended temporal land studies. As it questions the validity of historic map use in current day studies, it also provides a methodology that integrates historic data and modern data through the combination of historic maps and land cover data using freely available software programs. With an extended time series available provided by historic maps we are able distinguish the change in land cover through the time span, leading to future research in assessing the drivers of anthropogenic and natural processes of land cover.,

Keywords*:* GIS, Historic Data, France

1 Introduction

1.1 Research question

In this land use/land cover change analysis we make use of historic maps as a data source for the basis of performing an Intensity Analysis (Aldwaik and Pontius, 2011). This application in the current study analyzes seven different time points of forest change through types of change (loss, gain, exchange) through a temporal span of 163 years.

Using the historic forest cover found in maps dating back to 1840, a time series of land cover/land change is constructed for the Uxeau commune in Burgundy, France we explore the utility of using Intensity Analysis in an historical cartographic study of landcover change. The research considers the implications of historic data that must be manipulated in order to be appropriate for use in current analytical programs specifically Intensity Analysis. Our time series of 163 years consists of a variety of sources including historical maps and aerial photos of the study region. With the extracted forest data, we analyze the change in forest cover patterns.

Through the forest cover data, we see visible changes in the land patterns in our study site. Despite these changes throughout the 1840-2003 time period, we believe that the structure and form has remained the same. The landscape remains relatively undisturbed with the presence of many of the same populated and forested areas, as well as high agricultural productivity. This demonstration of this research method shows the potential for the integration of historical maps and modern analytical methods for the reconstruction of complex landcover change over time. (Madry *et al.,* 2012)

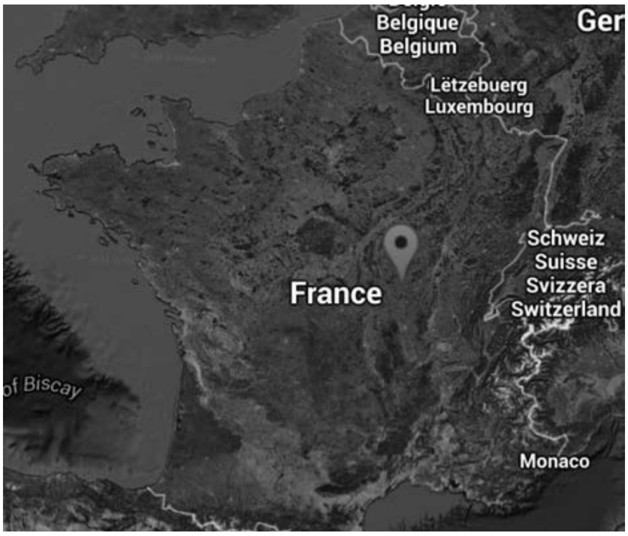
1.2 Study area

Located in the French department of Saone-Loire, the study site is an area of 136km^2 in the Arroux River valley, a tributary of the Loire River in Burgundy. The hilly landscape lies in the foothills of the Morvan Mountains to the north. The mixed agro pastoral economy is based largely on cattle production, with a lesser emphasis on sheep, goats, and cereal and fodder crops. This pattern of landscape use has remained consistent throughout the study period of the past 160 years. Forests in the area are naturally oak and beech, with hornbeam replacing beech at the highest elevations. In the last century, pine plantations have also joined the landscape. Forests are generally placed on portions of the landscape that are unsuitable for other uses- areas of steep slopes or poor soils. Forests in the area remain important for the value of the timber, the aesthetics of the tourist landscape, and for grazing and wildlife (Madry *et al.,* 2009).

2 Methods

2.1 Data and Pre-Processing

Our data consists of seven map which include the Carte d'Etat Major of 1840, and topographic maps of 1895, 1912, 1951, 1971, 1983 and 2003.



To conduct the needed analysis, it required a standardized methodology to convert historic maps into a usable and compatible format with the modern datum and projection systems of the latter half of our chronological series. The first step was the georeferencing of the topographic maps to ensure that the entire time series had the same projection system to overlay one over the other properly. The process consisted of defining points in a base map and choosing the same points for the X, Y value of that location in the map being georeferenced. The software, QGIS (Quantum GIS Development Team, 2009), open source software, was used. It easily allows a user to identify points on the base map and the georeferenced map through the plugin ' Georeferencer '.

In a study involving historic maps, such as this one, the process of georeferencing can prove difficult. First, the representation of objects or attributes or features used to identify one map with another may differ depending on the graphic representation of an attribute in one map versus that same attribute with a different graphic representation in a subsequent map. To remediate this issue in our study, the choice of points was crucial to correctly georeferencing a map from the time series. As such, churches and unusual landmarks, which were most likely triangulated by the cartographers of the era, were chosen (Grosso, 2009).

Additionally, a Root Mean Square (RMS) was calculated between the control points from the base map that correspond with the map being georeferenced. The RMS measures the average of the residual error between the points (Dammalage, 2005).Using QGIS, the RMS calculations are presented easily for each point and then a percentage based on the total is presented. With different

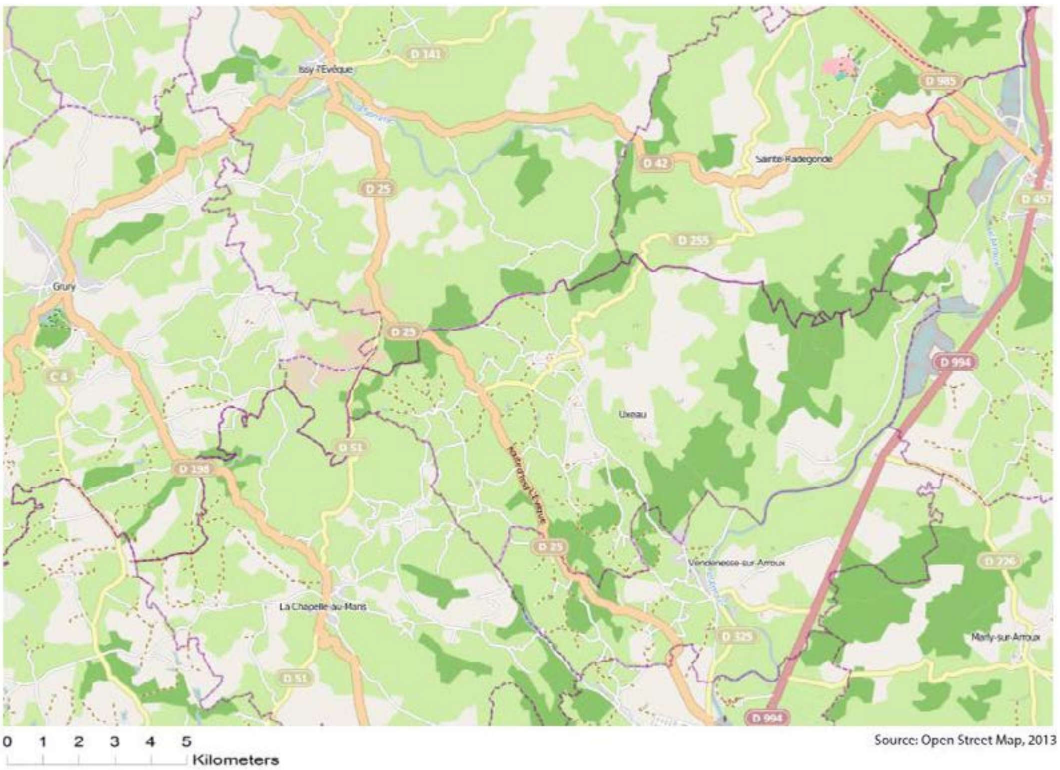


Figure 1: Map of location of the study site in France.

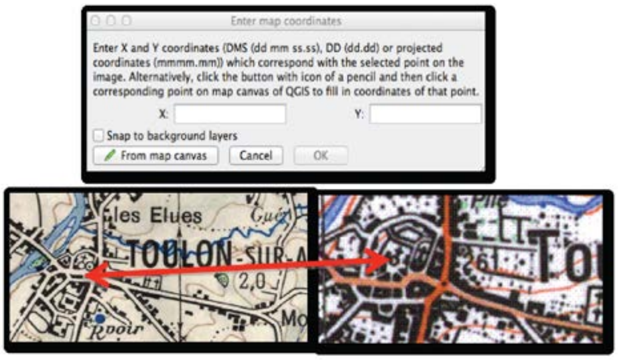


Figure 2: Comparison of Georeferenced Geographic Marker in the 1971 and the 1983 maps.

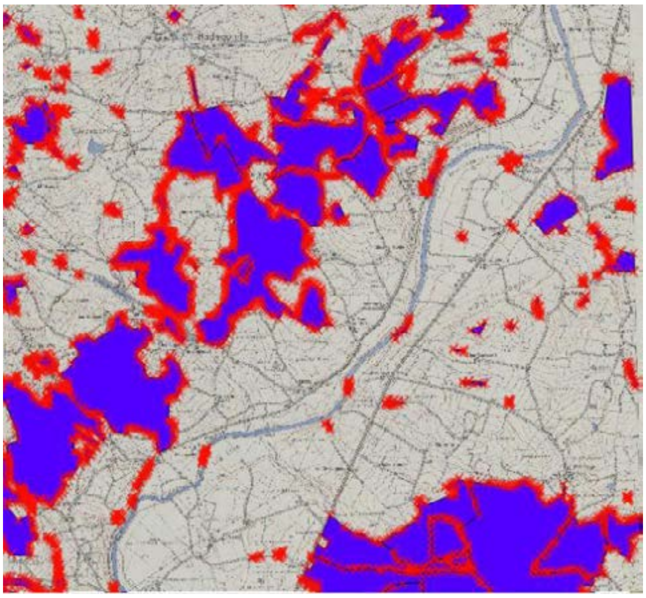


Figure 3: Map of Digitalization of Forested Regions in 1971.

scales present in the 10 dates of the time series of the research study, the total average RMS chosen was less than 20% for every singe date apart of our time series.

Second, the type of transformation used in the process of georeferencing is important to consider minimizing the amount of distortions in the final georeferenced map. As the data in the used in this research study consisted of mainly maps and aerial photos produced during the 20th century, it was decided to use polynomial 1 transformation. This transformation preserves 'collinarity and allows scaling and rotation only'(Quantum GIS Development Team, 2009).

2.2 Extraction of Forested Patches: Digitizing the Attributes from the Maps

After georeferencing the historical maps, it is now possible to extract the desired attributes in order to then analyze the changing nature of the forested regions. The most time consuming of steps, the extraction of forested regions and other attributes for the study site used the digitizing tool apart of QGIS. It was decided to use manual digitizing, specifically 'Heads Up Digitizing,' instead of automatic extraction of attributes because it is more precise in distinguishing between borders and outlines of a feature. While automatic extraction is faster, results can be filled with 'clutter' (Baily, 215) that require further processing. Additionally, as all our maps and aerial photos were not the same type of data, or color, it would be difficult to produce the same detailed results from automatic extraction for all maps. (Baily) Following the process of digitizing the first map, the process of 'Backward Editing'(Podobnikar and Kokalj, 2006) was followed to limit the amount of possible human error found in the digitization process. Additionally, as all our maps and aerial photos were not the same type of data, it would be difficult to produce the same detailed results for all maps (Baily, 2007).

At the completion of the digitizing method of all forest cover for the nine dates, each vector forest region was converted to a raster. This was done to manipulate the data for future analysis. With raster rather than vector data, the values of pixels were used to analyze of forest change. Using ArcGIS 9.3 and 10.0 (ESRI) because unlike QGIS, ArcGIS allows the user to choose the cell size in the creation of a raster. Specifying the cell size in the creation of the raster is an important step to diminish the human error from the digitizing technique. Depending on the size of the cell created, we can smooth the borders of forested regions produced from an incorrect digitized border. Using the function, ' Polygon to Raster ' and the center defined as specification for pixel formation, a cell size of 10 meters, or 100 square meters was chosen. It is very important that every map or aerial photo used from our time series had the same spatial extent to analyze the forests. After each new raster representing the ten previously vector shape files of digitized forested regions was clipped with a defined spatial extent, each new raster had 132376 cells. To analyze further, we then re-classified the values of cells as 1- forested, 0-other.

To understand the changing and evolving land cover patterns present in the study site, Intensity Analysis, developed in 2011, by Aldwaik and Pontius at the University of Rutgers, New Jersey, provided the detailed explanation of pixel loss, gain and persistence in the forested regions. Intensity Analysis is a new profound approach in research of the land use/ land cover change. This analysis aids in the 'calculation of the size and nature of hypothesized errors in the data that can explain deviations from non-uniform land change intensities at each level'(Aldwaik, 2011). Different from other types of analysis in land cover change research, Intensity Analysis allows the user to see the losses, the gains and the persistence between categories of soil occupation. The Excel program works for whatever period of time needed and with however many number of land use categories. The only specification is that the categories need to have the same spatial extent for the same area. There exist already numerous examples of land use/land cover research that use Intensity Analysis to explain the processes of an evolving landscape (Ponitus, 2004; Pontius, 2003; Aldwaik, 2011).

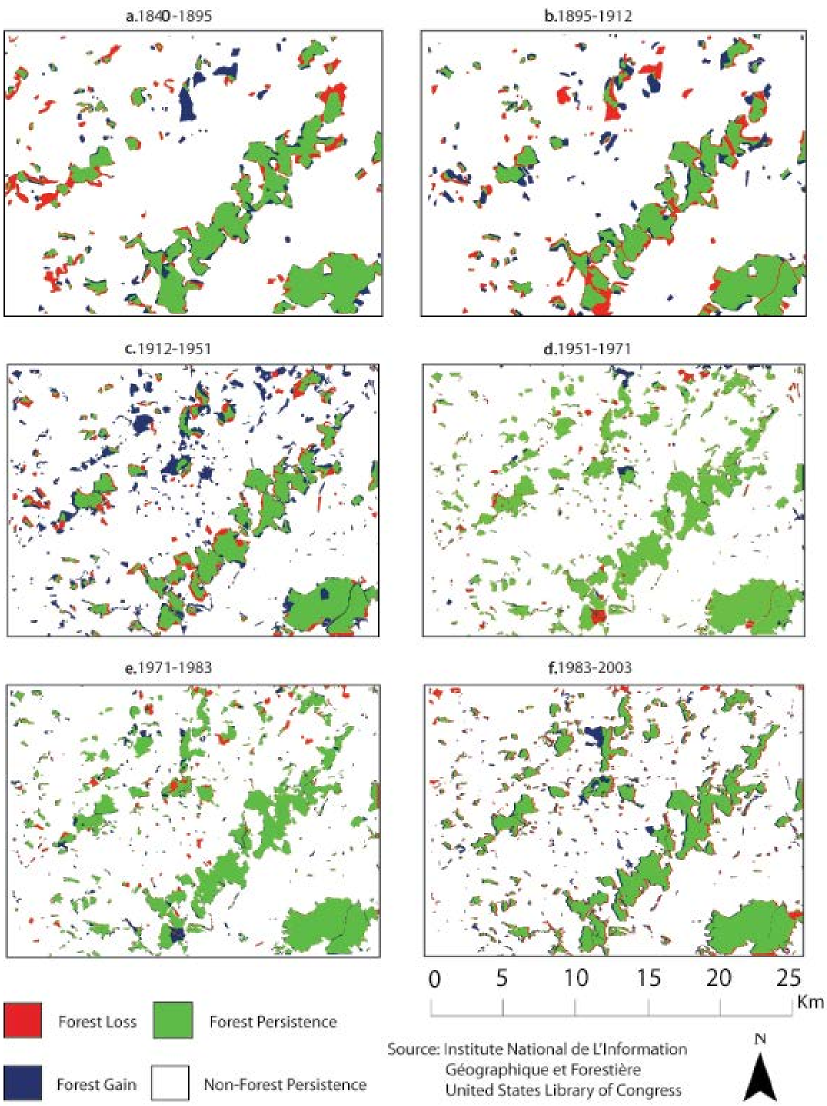


Figure 4: Six maps, where each map shows one time interval. Blue represents forest gain, green represents forest persistence, red represents forest loss non-forest persistence, white represents non-forest persistance. There is a black vector outline of the study area.

2.3 Quantity and Allocation Change

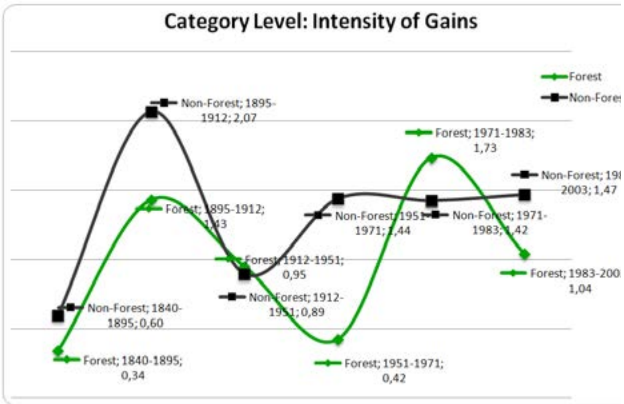
To analyze the changing evolution of the forested regions in the study site, we chose to focus on the type of change. This is defined by the quantity of pixels changed in our two categories and the allocated type of either loss, gains and persistence of pixels in our two categories. The pixels in these two categories were reclassed using the tool, ‘ Reclass ‘ in Arc GIS 10.0. For example, for the first interval of time, forest in 1840 is given the value 1; forest in 1895 is given the value 2, and 0 for non-forest in both 1840 and 1895. Using ‘ Raster Calculator ‘, a simple addition is made with the first map of 1840 and the second map of 1895. This process is repeated for each time interval (fig. 4).

2.4 Single Interval

From the results of the pixel addition of each time interval, a cross-tabulation matrix is created with four categories, Forest Persistence, Forest Gains, Forest Losses, and Other Persistence for the 6 intervals of time from the 9 dates of our map/aerial photo series (1840, 1895, 1912, 1945, 1951, 1971, 1983, 2003). These values of the cross-tabulation matrix are entered into the Intensity Analysis excel program (table 1).

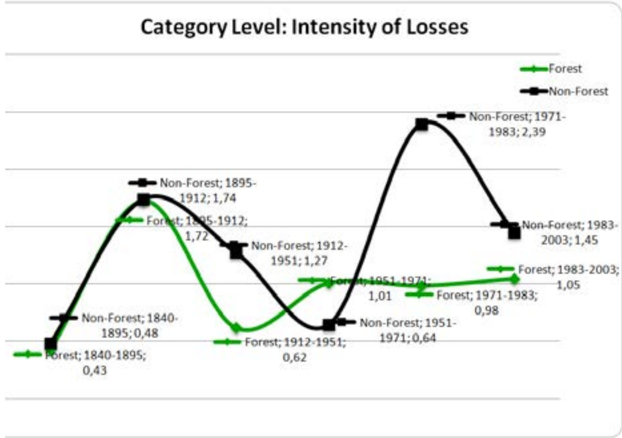
		1895 Forest	1895 Non-Forest
1840	Forest	173617	53995
1840	Non-Forest	39623	114713
		1912 Forest	1912 Non-Forest
1895	Forest	153351	59889
1895	Non-Forest	49582	115155
		1951 Forest	1951 Non-Forest
1912	Forest	153985	49579
1912	Non-Forest	190825	113108
		1971 Forest	1971 Non- Forest
1951	Forest	212905	31905
1951	Non-Forest	19002	116166
		1983 Forest	1983 Non-Forest
1971	Forest	184813	45939
1971	Non-Forest	53186	113043
		2003 Forest	2003 Non-Forest
1983	Forest	217329	21081
1983	Non-Forest	20670	115529

Table 1: Cross Tabulation Matrix of pixel values (non forest persistence, forest persistence, forest gain and forest loss) or time series (1840-2003) Pixel values.



Forest	0.34	1.43	0.95	0.42	1.73	1.04
Non Forest	0.60	2.07	0.89	1.44	1.42	1.47

Table 2a : Category Level: Intensity Gains in each interval. Red cell indicates category is dormant in gross gains. Green Cell indicates category is active in gross gains.



Forest	0.43	1.72	0.62	1.01	0.98	1.05
on-Forest	0.48	1.74	1.27	0.64	2.39	1.45

Table 2b : Category level: Intensity of Losses for each interval. Red cell indicates category is dormant in gross losses. Green Cell indicates category is active in gross losses.

2.5 Multiple Intervals

Finally, we use the cross-tabulation matrix again to analyze not only the change within intervals but also the change across all intervals in our time series. This analysis evaluates the quantity of instability of change throughout the time series in a comparable manner. In this research, this is extremely helpful when analyzing time intervals that are not consistently the same amount of time.

3. Results

With the entry of the pixel values for the two categories and six intervals into both the Intensity Analysis and the multiple time interval analysis, the results are easily tabulated and then compared in Excel. With this, we are able to discern information about three levels of change: interval (between dates), category transition gains and losses (between type of land soil cover), and finally, the type of change in comparison with the entire time series. Each of these types of results is displayed in the table 2a and 2b below.

From the results, we see that the changes across the six intervals of the time series display ebbs and flows of intensity of change. Represented by percent of the map, we see that the two peaks of interval change are from the 1895-1912 interval (1.69%) and the 1971-1983 interval (2.13%). This implies that these two intervals experienced the most change within the land cover. The first and last intervals represent the lowest interval level change with values of .45% and .53% respectively. These two intervals would be classified as having the least amount of change between intervals.

From table 2a, we see the side-by-side comparison of gains between forest and non-forest through the time series. We see that again the two high value peaks are found at 1895-1912 and 1971-1983. For second interval (1894-1912), forest has a value of 1.44 and non-forest has a value of 2.01. The fifth interval (1971-1983) gives a value for forest of 1.86 and a value of non-forest of 2.41. The value of forest gains for the first interval (1840-1895), the fourth interval (1951-1971) and the sixth interval (1983-2003) have a similar low value. 1840-1895 has the lowest forest gain value of .34. The fourth and sixth interval have values of .41 and .43 respectively. The only interval in which forest gains is more than non forest gains is interval three or 1912-1951. The value of forest is .91 compared with non-forest gains of .78. The results indicate an overall decline in forest as we see that only the third interval has forest higher forest gains. In Interval 5, 1971-1983, the intensity of loss seen by non forest is at its highest of 2.67.

From the results of the Intensity of Losses, table 2b, we see again the high peaks mimicking the interval level change results. Again, 1895-1912 and 1971-1983 have the highest values in both categories of forest and non-forest for losses.

To draw historical connections in our time series, we relied on literature and previous studies of the study site. In assessing our historical map data, we surmise that the interval of change between 1895 and 1912 may be explained by the shift in map scale and quality. The scale of the 1895 map is 1:100,000 and it may be the case that many of the patches at this scale have been concatenated for ease of representation. The 1912 map scale is 1:50,000, revealing a much greater level of detail, and would therefore, display more patches of smaller size. The interval of change between the 1971 and 1983 maps which are both at the same scale (1:25,000), and therefore represent a more valid comparison, is most likely due to the subsidies offered by the French government for establishing pine plantations (Clout, 1993; Oudin, 1956).

4 Discussion and Conclusions

4.1 Data Quality

In this study, the integration of multiple formats and dates of data as well as the use of land cover/land change statistical analysis intended for remote sensing data brings a variety of questions to consider. First, the historic maps that were included lack a modern projection system, which can create slight discrepancies in the georeferencing system. Second, digitizing manually allows for a range of small errors. This is especially true if the symbology and representation of the layer to be digitized varies. In the case of this research study, forest representation for each date was relatively similar. . Significantly the scale of the historical maps differs greatly affecting the level of detail shown which is critical for determining patch size. The maps prior to 1951 range in scale from 1:40,000 to 1:100,000. The maps in the after World War II range in scale from 1:20,000 to 1:25,000 and represent greatest consistency and level of detail in the map series. For the time series in this study, the first map of 1840 limits the possibility for inaccurate representation of a feature. The maps used were all maps produced for French official use, which also limits the artistic variability seen in some historic maps. Additionally, it remained important to include data that was valuable to the study and not only because it existed. There were three previous dates of maps included in the time series, which offered a longer time period of study. After the analysis was completed for a first time, it was decided that the data extracted from the three maps was questionable, which influenced the results of the land cover analysis. To limit erroneous results, we repeated the data extraction process.

While the quality of the data was not doubtful for the chosen maps, as scanning the maps was detailed enough for forest extraction, it remained important to establish quality measures in our methodology. This was established through the commitment to specific digitizing procedures. Two different quality controls, completed by another member of the team, were conducted to maintain the accuracy of the digitized data. Through this process, it was assured that all areas of forest were extracted. Finally, the same standards were used throughout the entire research project in order to ensure the least possible amount of variance within the analysis.

4.2 Conclusions

Through the forest cover data, we see visible changes in the land patterns in our study site. The purpose of this study was to explore the utility of this combined data methodology in understanding land cover change over time. The results of this research showed that while the total acreage in forest remained consistent throughout the research period, there were losses and gains in forest extent Despite these changes throughout the 1840-2003 time period, the forest structure and form has remained the same. The landscape remains relatively undisturbed with the presence of many of the same populated and forested areas, as well as high agricultural productivity (Madry *et al,* 2012).

This demonstration of this research method shows the potential for the integration of historical maps and modern analytical methods for the reconstruction of complex land cover change over time. In order to verify this trend indicated by the Intensity analysis, detailed historical studies will have to be conducted to show whether some of the change is due to real changes in land cover or to differences in map scale and mapping techniques. The utility of this method of Intensity Analysis in providing a starting point for further investigation has been demonstrated by this preliminary study. Important to this study, or any study involving historic data and maps, is the maintenance of strict standards and a research methodology to limit possible error from processing and analysis.

References

ALDWAIK, S. Z., & PONTIUS Jr., R. P. (2012). Intensity Analysis to Unify Measurements of Size and Stationarity of Land Changes by Interval, Category, and Transition. *Landscape and Urban Planning* 106: pp. 103-114.

ALDWAIK, S. Z., & PONTIUS Jr., R. P. (under revision). How Much Error Can Explain Non-Uniform Results from Intensity Analysis? *Journal of Geographical Information Science*.

BAILY, B. (2007). The Extraction of Digital Vector Data from Historic Land Use Maps of Great Britain Using Image Processing Techniques. *E-Perimetron*, 2, pp. 209-223.

CLOUT, H. D. (1993). The land of France 1815-1914. *Research Series in Geography*, 1, p. 124.

DAMMALAGE, T. L., SAMARAKOON, L., & SUSAKI, J. (2005). Geometric Correction of High Resolution Satelitte Data-Effect of the Distribution, Accuracy and Number of GCP. 26th Asian Conference on Remote Sensing, Hanoi, Vietnam, 7- 11 November 2005, viewed 11 July 2012, Avalaible from: http://a-a-r-s.org/aars/proceeding/ACRS2005/Papers/DIP-3.pdf

EREMIASOVA, R., & SKOKANOVA, H. (2009). Land Use Changes (Recorded in Old Maps) and Delimitation of the Most Stable Areas from the Perspective of Land Use in the Kasperske Hory Region. *Journal of Landscape Ecology,* 2 (1), pp.20-34.

ESRI. (2010). ArcGIS 10.0. Redlands, USA.

GASPERI, P. (2007). Cartographic Images of Slovenia through Time. *Acta Geographica Slovenica*, 47 (2), pp. 245-273.

GAVRILOVA, M. (2012). Data Quality Issues. Calgary, Canada. Computer Science Course, university of Calgary. [Online] Available from: http:// pages.cpsc.ucalgary.ca/~marina/695/Lecture_errors_quality.ppt

GRASS DEVELOPMENT TEAM. (2011). Geographic Resources Analysis Support System (GRASS) Software. (O. S. Foundation, Producer) Retrieved 5 10, 2012, [Online] Available from: http://grass.osgeo.org [Accessed: 23th February 2015].

GROSSO, E. (2009). Proposition pour une gestion unifiée des données anciennes. In *Colloque International de Géomatique et d'Analyse Spatiale, Spatial Analysis and GEOmatics (SAGEO 2009)*, pp.1-16.

JONES, E., MADRY, S., & TICKNER, A. (2009). An Improved Method for Extraction of Historical Cartographic Features into GIS: a French Study. In *Proceedings of Computer Applications in Archaeology 2009*. [Online] Available from: https://burgundylandscapes.files.wordpress.com/2012/04/madry-jones-tickner-2009-an-improved-method-for-extraction-of-historical-cartographic-features.pdf.

JONES, E., MADRY, S., & TICKNER A. (n.d.). *GIS and Remote Sensing in Burgundy France.* [Online] Available from: http://www.informatics.org/france/france.html. [Accessed 23/02/2015]

MADRY, S. (2012). GIS and Remote Sensing for Archaeology: Burgundy, France. Retrieved Juillet 9, 2012, [Online] Available from: http://www.informatics.org. [Accessed 23/02/2015]

MADRY, S., JONES, E., MURRAY S. and TICKNER A. (2011). Une Micro-histoire de la Terre et l'Utilisation des ressources. *Le Monde des Cartes* (208), pp. 75-92

MADRY, S., JONES, E., & TICKNER, A. (2009). An improved extraction of historical cartographic features into GIS: a French case study. In B. Frischer, J. Webb Crawford and D. Koller (eds.). *Making History Interactive. Computer Applications and Quantitative Methods in Archaeology (CAA). Proceedings of the 37th International Conference, Williamsburg, Virginia, USA, March 22-26, 2009*, viewed 5 March 2012, Available from: http://www.caa2009.org/articles/Madry_Contribution186_a.pdf.

OUDIN, A. (1956) Le Fonds Forestier National en France. *European Journal of Forest Research*. [Online] Available from: Retrieved 9 10, 2013, http://www.senat.fr/rap/a00-190/a00-1902.html. [Accessed 23/02/2015]

PODOBNIKAR, T., & KOKALJ, Z. (2006). *Triglav National Park Historical Maps Analysis. presentation.* [Online] Available from: http://www.mountaincartography.org/publications/papers/papers_bohinj_06/17_Podobnikar_Kokalj.pdf. [Accessed 23/02/2015]

PONTIUS, R. G., SUSAS, J. E., & MCEACHERN, M. (2004). Detecting Important Categorical Land Changes while Accounting for Peristence. *Agriculture Ecosystems and Environment*, 101, pp. 251-268.

PONTIUS, R. (2003). Using the Relative Operating Characteristic to Quantify Certainty in Prediction of Location of Land Cover Change in India. *Transactions in GIS*, 7 (4), pp. 467-484.

QUANTUM GIS DEVELOPMENT TEAM. (2011). Quantum GIS Geographic Information System. (O. S. Foundation, Producer). [Online] Available from: http://qgis.osgeo.org. [Accessed 23/02/2015]

Challenges and Perspectives of Woodland Archaeology Across Europe

Hauke KENZLER,[1] Karsten LAMBERS[2]

[1]Bavarian State Department of Monuments and Sites Munich, Germany
[2]Institute of Archaeology, Heritage Sciences and Art History University of Bamberg, Germany
hauke.kenzler@blfd.bayern.de, karsten.lambers@uni-bamberg.de

Abstract

This paper reviews the challenges and prospects of woodland archaeology across Europe and proposes a European network to safeguard archaeological heritage in woodlands. Woodlands and forests cover important parts of the European landmass but are often uncharted territory on the archaeological map since traditional methods of archaeological survey do not work well here. Many forests have grown on formerly open lands used for farming or settlement, and some forests have been used for charcoal burning and wood pasture. As a consequence they contain important archaeological remains that are often well preserved but little known and protected. Recent developments in the field of remote sensing have opened up new avenues for important archaeological research in woodlands. However, the legal and administrative framework to protect archaeological sites is of equal importance. While the economical, recreational, and ecological dimensions of forests are commonly known and accepted, their archaeological dimensions are rarely recognized.

Keywords: Woodland Archaeology, Cultural Landscapes, Heritage Conservation, Forestry, Ecology

Introduction

Forests and woodlands tend to be distinguished by the density of tree coverage and tree height, with 'woodlands' often being used as a more general term that also covers less densely overgrown areas. However, this distinction is gradual and often ambiguous (see for example, Forestry Commission, 2011, 4), and it is even less clear in other European languages, e.g. in German. In the following, both terms are used interchangeably. Following a recent official European Union publication on forest ecosystems, about 33% of Europe and 42% of the 27 EU countries are today covered by forests and woodlands (European Environment Agency, 2010, 4 f.). This percentage is even higher in the mountain ranges of Central, Eastern, and Northern Europe (fig. 1). Forests and woodlands are thus a formative part of the European cultural landscape, and are an important resource in economic, ecological, social, and cultural terms.

Less well known is the fact that forests in Europe have continuously expanded over the last 60 years and are still increasing today. Just in the EU countries, the forested area increased by 17 million hectares in the last two decades. Only a small part of this increase is due to natural regrowth in former agricultural or industrial areas. Most forest expansions are artificially planted and serving purposes such as timber production, promotion of renewable energy, biodiversity conservation, carbon storage, or prevention of erosion and soil degradation (Forest Europe *et al.,* 2011). Generally, there is little knowledge about tree planting having started already in the late Medieval Period, e.g. as firewood and a source for building material or as a means for preparing hunting ground for noblemen. Already in 1368, pine and fir trees were artificially seeded for reforestation in the Nuremberg Reichswald in the southeast of Germany (Stromer 1968, 26). Consequently, today only 26 percent of European forests, located primarily in remote and inaccessible areas in eastern and northern Europe, are regarded as more or less undisturbed (Forest Europe *et al.* 2011, 8). This means that most areas covered by forests today were once wholly or partially cleared and cultivated, as was the case during the Bronze and Iron Ages, in the Roman Period or during the Middle Ages. These areas were reclaimed by natural or artificially planted forest regrowth in later periods. As a result, many of these secondary forests cover today important vestiges of prehistoric and historic settlement and land use. Forests can thus be regarded as archives (Hamberger *et al.,* 2012) or time-machines (Angelstam *et al.,* 2011). This archaeology 'in the forest', dealing with archaeological sites and features in once open landscapes now covered by secondary forest, is one aspect of cultural heritage research in, and management of woodlands. Archaeology 'in the forest' can take on different forms, as accounts in classical sources show. For example, according to Caesar and Tacitus, woodlands served as buffer zones between neighbouring tribes. Thus, documenting no finds from a specific period in forests can also be interpreted as valuable historical information. Another important aspect is the archaeology 'of the forest', e.g., the study of the use of the forest itself as a natural resource for people in the past, e.g. as fire wood, charcoal and source of building material, as wood pasture, or as a hunting ground. Even forests that persist until the present day were used extensively in the past, especially those forests close to inhabited areas. Those forests often had a different, more open appearance than forests have today. Industrial sites such as mining, iron production or glass manufacture required a steady supply of wood fuel. Their remains are often located in or close to wooded environments, and they are therefore forming a link between the overall distinction of an archaeology 'in the forest' and 'of the forest' (e.g. Forestry Commission, 2011, 3 f.; Hamberger *et al.,* 2012; Ritchie and Wordsworth, 2010).

1. Challenge

On the one hand, the tree cover protected and preserved many archaeological monuments and landscapes in woodlands for a long time. Intrusion into, and disturbance and erosion of soils as well as sediments covering or containing archaeological remains was minimal as compared to open areas. More so than in arable land, many archaeological sites and features in forests are therefore still visible on the ground surface. This facilitates their study by non-invasive or minimally invasive methods causing less destruction than excavations. Thus, archaeological sites in woodlands are often better preserved than open-area sites, a fact that further increases their importance as historic monuments. However, some forest-specific disturbances, e.g. windfall and bioturbation, often affect woodland sites as well (Crow, 2004; Sippel and Stiehl, 2005, 48 ff.). On the other hand, tree cover has also prevented many of these sites and monuments from being detected, properly recorded, and preserved by archaeologists and cultural heritage managers. One of the reasons is that excavations and common methods of archaeological survey, such as aerial archaeology, field walking, and geophysical prospection are more difficult in forests than in open farmland since vegetation acts as a barrier. Furthermore, archaeological prospection is often focused on or even limited to open areas, the reason being that commonly used techniques are more easily applied there. Archaeological sites in open areas are also at a higher risk to be exposed to disturbances, such as in farmland or in areas under development. In addition, forests tend to be considered by the public as a natural landscape rather than a cultural landscape, such as e.g. open farmland, often preventing archaeological research in forests to receive appropriate public support and funding.

In turn, large woodland areas in different parts of Europe could not be explored in the past, and are literally uncharted territory on the archaeological map (fig. 2). Hence, important parts of the existing archaeological record remain inaccessible for analysis and interpretation by archaeological and historical research. Important geographical parts of the archaeological record are simply non-existent for those periods of European cultural history, such as the High Middle Ages, during which forest coverage in certain regions was more limited than today. The large gap in archaeological research affects our

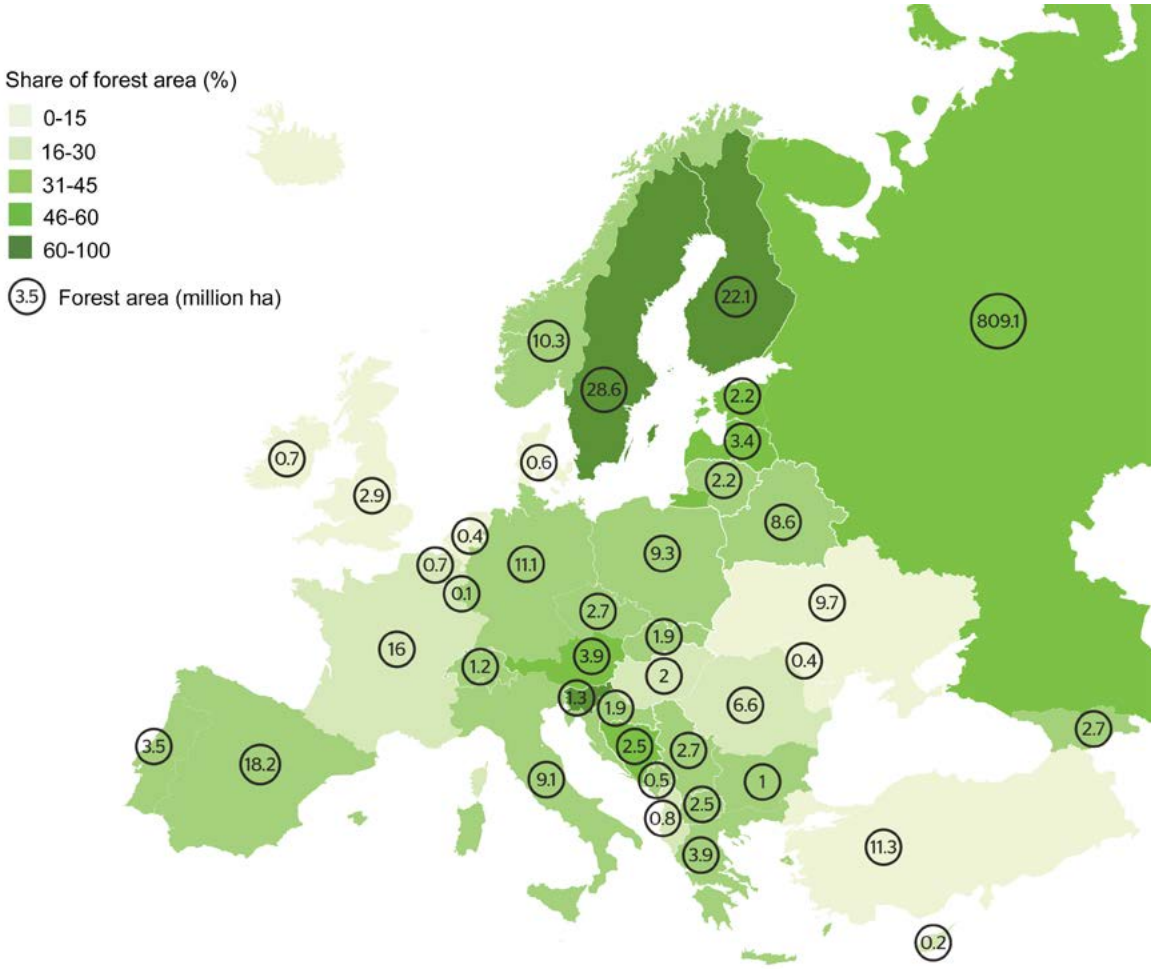


Figure 1: Forest area (million ha) and share (percent) of land area by country (after Forest Europe *et al.* 2011, 19).

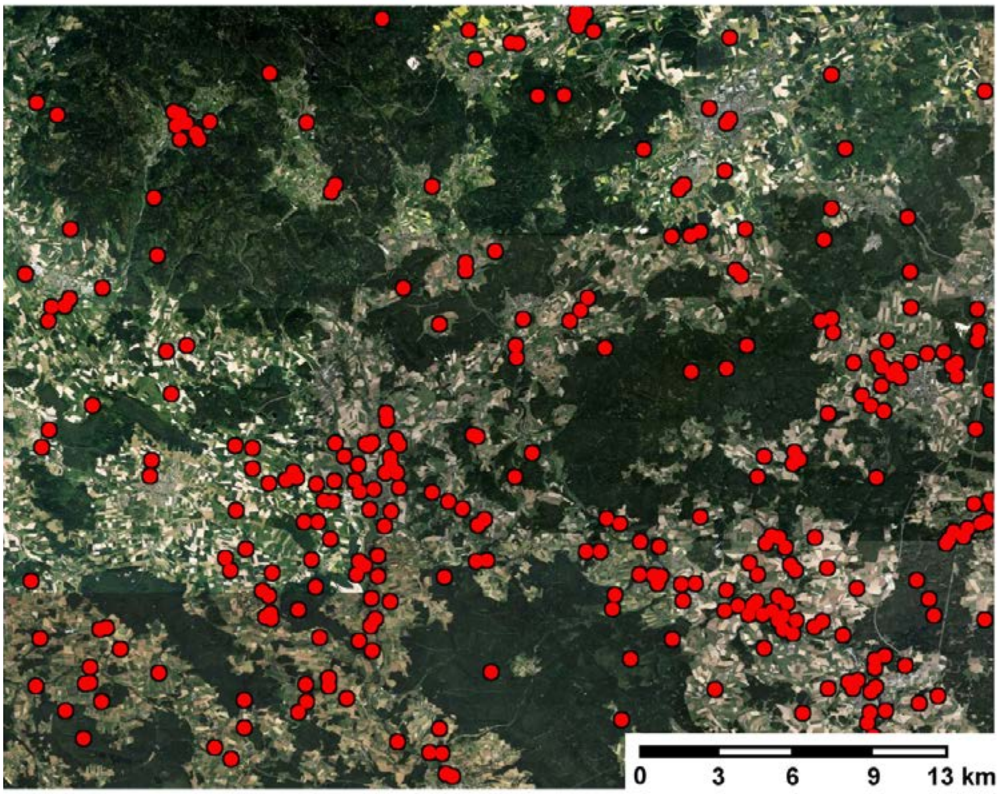


Figure 2: The archaeological map around Kemnath in Bavaria reveals forested areas as uncharted territories. (Authors, Bavarian State Department of Monuments and Sites).

understanding of the cultural history of these epochs and regions significantly. For cultural heritage management, the problem is even more pronounced since limited knowledge about existing archaeological sites and features makes it difficult to devise and implement effective strategies for their protection and preservation.

2. Technological and methodological aspects

The situation has markedly improved over the past 10 to 15 years. New tools have become available to archaeologists and cultural heritage managers opening up new and promising avenues of archaeological prospection in forests and woodlands. Where applied, these tools have significantly improved our knowledge about the rich archaeological heritage contained in forests. Many of these new tools have been developed from recent technological and methodological innovations in remote sensing. During the last two decades, the introduction of Airborne Laser Scanning (ALS, also known as LiDAR: Light Detection and Ranging) into cartography has proven the most important step in this development. Using an airborne laser sensor along with highly accurate positioning devices, ALS records ground elevation data in very high resolution (Crutchley and Crow, 2009). The nature of the laser signal allows differentiating between signals returned to the sensor from the ground, and signals returned from vegetation and other objects on or above the ground surface. Computationally removing these non-ground signals virtually strips the ground surface from vegetation cover, thus allowing an unimpeded and high-resolution view on the topography of the ground even beneath dense vegetation cover, such as forests. As demonstrated by numerous recent case studies, this technology has proven to be a major breakthrough for archaeological research (see examples in Cowley, 2011; Opitz and Cowley, 2013). In regions where ALS-derived digital terrain models (DTMs) have been analysed for the purpose of archaeological prospection, the number of known archaeological sites has increased greatly. The increased application of ALS over large areas also broadens our perspective of single archaeological sites and features as well as entire archaeological landscapes that hold not just settlements, but also field systems, roads and pathways, workshops, quarries. The wide variety of other features from different periods considerably improve our picture of historic and prehistoric landscape use (e.g. Doneus and Briese, 2011; Hesse, 2013; Risbøl, 2013). While ALS is a major technological breakthrough, it is by far not the only step forward in archaeological prospection. Other remote sensing technologies such as radar or multispectral and hyperspectral sensing are now being tested over forests (Beck, 2011; Wiseman and El-Baz, 2007). Terrestrial laser scanning (TLS), conventional pedestrian surveys, geoarchaeological investigations, analyses of phosphate contents and increasingly geophysical prospection such as geomagnetic or ground penetrating radar (GPR, fig. 3) have all contributed in recent years to a better knowledge and understanding of the archaeological record in woodlands. In some countries, conventional geodetic-topographical and geophysical survey methods have been significantly improved and adjusted to the special requirements of woodlands. The same applies to the interpretation of surface features that can reveal important information about a certain archaeological monument without the need for expensive and destructive excavations (e.g. Večeřa, 2004). Due to appropriate adaptation of the equipment and the used measurement method, geophysical prospection has been successfully applied in woodlands despite obstructions caused by trees and roots. To achieve the best results, there has been a strong trend in these recent developments to combine different techniques and methods, as well as the combined analysis of the resulting data using computational approaches, e.g., digital image analysis and computer vision (Doneus, 2013; see also recent examples in Neubauer *et al.,* 2013).

These advancements show that archaeology and cultural heritage management to date have more powerful tools at their disposal for archaeological prospection in woodlands than was the case only 15 years ago. In countries where Airborne Laser Scanning (ALS) has already been used for several years to explore monuments in forested

Figure 3: Ground Penetrating Radar with 400 mHz antenna in Ruppertshütten, 'Kloster Einsiedel', Germany. (Harald Rosmanitz, Spessart Archaeological Project).

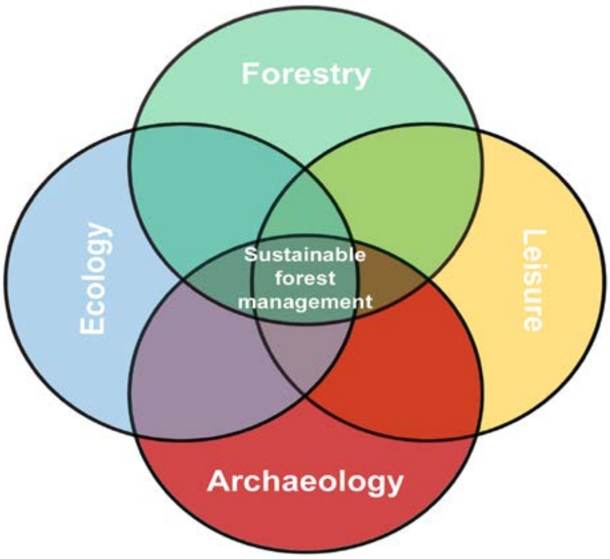


Figure 4: Different stakeholder interests are linked by a sustainable forest management. (Authors).

areas, the techniques have led to an enormous increase of known sites. Additionally, these new techniques allow for a better protection of important monuments. Even central European landscapes that have been explored archaeologically for well over a hundred years, observe due to the more effective techniques a 40% growth of archaeological monuments in certain areas. Thus, the amount and importance of knowledge gained is comparable to what has been learned after the introduction of aerial archaeology to arable land many decades ago. While these developments are encouraging, their downsides should not be overlooked. The application of new approaches and methods often requires heavy investments in terms of finances, time, training, and infrastructure for data processing, analysis and storage. These requirements often face limited, or even decreased resources available to archaeological and heritage professionals and institutions. Thus, access to these new approaches is not only unevenly distributed across Europe but often also limited, impeding the application of effective methods. In addition, these methodological and technical approaches are often not sufficiently available for the teaching and training of young professionals, which is likely to slow down their prospective application.

3. Legal and administrative aspects

The technological and methodological progress described above is yet to be matched by similar progress in the legal and administrative domain. There is little knowledge among the public as well as the individual stakeholders and decision makers involved in forest management about the fact that forests often hold well-preserved but understudied and underprotected archaeological remains. Features and sites forming the historic environment provide the key for a better understanding of our past. These sites are a finite, non-renewable resource, and once lost, they cannot be replaced. This historical archive of our society (Hamberger *et al.,* 2012, 47) is endangered by forestry, building activities, recreational use, tourism, and other activities. In many parts of Europe, forests continue to play an important economic role. In some industrial countries and regions, wood was and still is mainly used as building material or as raw material for paper. This type of economic use has led to fast-growing monocultures that have replaced traditional, more diverse forests. Additionally, in recent years forests have to a certain extent regained their traditional role as energy source through the production of bioenergy. This is especially true in countries with a policy to promote the percentage of renewable energy. As a result, large or straight timbers are just as much in demand today as is the entire available biomass, including coppice, hedgerows etc. The appreciation of the commodity wood has led to developments in agro-forestry that have damaged and destroyed ancient monuments in previously unknown ways. One example is the increased use of heavy machinery, such as bulldozers and harvesters that cause lasting damage to the ground. Another example is the preference of fast-growing monocultures that necessitate area-based forest establishment and planting as well as harvesting and extraction practices causing extensive damage and ground disturbance. Although, the use of wood chips requires more frequent transport related intrusions to the forests, the whole-tree harvesting practiced increasingly across Europe is the most devastating form of forestry since the roots also have to be removed. (Crow, 2004, 31 ff.). Over the last decades, these developments have affected forest soils as well as archaeological sites and features contained within them more severely than was the case during the entire history of forestry (Crow, 2004; Hamberger *et al.,* 2012, 42 ff.; Möllers, 2004; Sippel and Stiehl, 2005, 48 ff.). Sites of which private or public landowners, forest industry and authorities are unaware are obviously the least protected. For example, damage due to access roads cutting through well-preserved archaeological monuments like walls or grave mounds could have easily been avoided if the significance of these archaeological sites had been known. In essence, damage to archaeological monuments is mostly caused by ignorance, rather than intent (Hamberger *et al.* 2012, 42 f.; Sippel and Stiehl, 2005, 51). Since 1989, a special threat has been affecting forests along the former Iron Curtain across Europe. While in some regions new natural reserves were established on former border fortifications, in other regions it was the re-opening of closed or little frequented traffic during the Cold War. The economic and demographic development of formerly marginal regions, led to intensive road and railroad building and other developmental measures that required forests to be opened up.

Apart from their economic role, forests are nowadays also an important resource for tourism, recreation and leisure. Archaeology shares many interests with these fields. Some forest authorities have utilised the presence of important archaeological vestiges and created heritage walks with information signs to raise people's interest in the forest. Such trails attract more tourists and locals to the forests and thus provide both health and educational benefits. Archaeological sites and landscapes thus become part of the recreational value of forests. At the same time, a sprawling tourism and recreational and leisure-time activities can cause damage to valuable monuments, e.g. damage caused by the wear and tear due to mountain biking. However, destruction of historical features is mostly caused by hiking or biking trails, access roads, car parks or service facilities. Although, archaeology and tourism respectively leisure industry have many common interests, these have so far rarely been discussed between archaeologists and ecologists (fig. 4). A changing perception of the public constitutes a chance for an effective archaeological heritage management in woodlands. In some parts of Europe, where the industrial use and economic value of forests are slowly decreasing, forests are increasingly appreciated for their ecological and/or recreational value, especially among the urban population. This has led in some areas to a change of woodland resources management enforcing more sustainable practices. For example, industrial monocultures are replaced by more diversified vegetation closer to their natural composition. The result is a higher recreational, environmental and aesthetic value of forests, as variedly composed forests are more resilient to soil degradation and extreme weather events. Another motivation for the change of woodland resources management is an increasing biodiversity and reduced CO2 levels, to strengthen the importance of forests in the mitigation of the adverse effects in climate change. While slow, this process of forest conversion is often profound and brings about important changes of forest composition, perception, and management. A more sustainable management of forests and their different kinds of resources is clearly a chance for archaeological heritage management. Subsurface and near-surface archaeological remains benefit from decreased soil degradation, limited use of heavy machinery, and a general appreciation of forest variety and diversity. The creation of near-natural forests as compensation areas for landscape losses elsewhere is welcome specifically on archaeological sites, which are protected by permanent exclusion from intensive forest management. Conflicts may arise however, when measures to investigate and protect archaeological sites require the removal of trees.

4. Pathways toward the safeguarding of archaeological heritage in woodlands

The examples above show that woodlands today are subject to widely differing and often conflicting demands imposed upon them by a variety of stakeholders. These stakeholders include, but are not limited to, private and public landowners and forestry industry, public authorities in charge of administration, transportation, and the management of economic, cultural, and environmental resources, local communities, researchers across different disciplines, and people, companies and institutions engaged in recreational and leisure activities. While there is some interaction and exchange between these stakeholders to reconcile conflicting interests, cultural heritage management is usually underrepresented, among others because some stakeholders are often not even aware of cultural heritage being present in many wooded areas. In fact, the important function of forests as time capsules for cultural history is generally little known or acknowledged. For example, the archaeological dimension of forests is not mentioned in the recent EU Green Paper on forests (European Commission, 2010), in which a wide variety of other socio-economic and environmental functions of forests are listed and discussed in detail. The current State of Europe's forests report (Forest Europe *et al.,* 2011, 108) only states that there is a 'growing recognition of the importance of the cultural and spiritual values associated with forests', resulting in an increased number of countries being able to provide data on these categories. Thus, effective communication, interaction, and engagement between the different stakeholders are urgently needed to raise awareness of the cultural dimension of forests, and of the necessity of designing and implementing a strategy to preserve and protect the archaeological heritage in forests (fig. 4). This is a common challenge across most European countries, and joint international efforts are needed to define a baseline for future action. However, interaction between different stakeholders is organised in quite different ways from country to country, and sometimes from region to region, requiring individual solutions for each country and region.

4.1. Public outreach

In order to put woodland archaeology on the agenda it is necessary to raise awareness of the cultural dimension of forests and woodlands among the public, the broad research community, stakeholders, and decision makers. Scientific results need to be presented to the public in order to achieve appreciation and approval. On a local level, this may mean to prepare dissemination material such as leaflets, brochures, information boards etc (Forestry Commission, 2011; Hamberger *et al.,* 2012). Guided tours to archaeological monuments or special events are an approved instrument to directly reach specific groups of people. Furthermore, it has become highly important to promote ideas via Social Media due to their wide distribution. The outreach to the research community is the easiest task. The increased interest of archaeologists and heritage professionals in woodland archaeology has been demonstrated at recent pan-European conferences, for example through thematic sessions at the annual conference of the European Association of Archaeologists (EAA) in September 2013 in Pilzen (CZ) and the annual Computer Applications and Quantitative Methods in Archaeology (CAA) conference in April 2014 in Paris. Similar conferences have also taken place in several European countries on a smaller scale. In larger networks such as the European Forest Institute (www.efi.int), Archaeolandscapes Europe (ArcLand, www.archaeolandscapes.eu), or the EAA and EAC Joint Working Group on Farming, Forestry and Rural Land Management (http://www.e-a-a.org/wg2.htm), the subject treated here is also represented, although it is not their main focus. The best approach to achieve an impact on policies affecting the archaeological heritage in forests will be by exchanging experiences, identifying best practice and elaborating guidelines. Ultimately, such undertakings have always been promoted through the initiative of individual researchers and institutions. So far there has been little coordination, cooperation, or even exchange. Thus, it is necessary to bring these attempts from a regional to a national and maybe even European level.

4.2. Legal framework

While the fundamental problem of archaeological and cultural heritage resources in woodlands is principally the same across Europe, the management of woodland resources is organized on different levels (national, regional, local) and within different legal frameworks. Besides the overall lack of awareness for the archaeological heritage in forests, this has led to a high degree of fragmentation within the professional heritage research community. Public ownership of forests is common in the Russian Federation, Central-East and South-East Europe. For Central-West Europe, North Europe and South-West Europe the average percentage of forests in public ownership is only around 30 percent. Particularly high proportions of privately owned forest are found in Austria, France, Denmark, Norway, Sweden, Slovenia, Portugal and Spain (fig. 5). In general, the proportion of private forests and numbers of private forest holdings has increased over the last 20 years. The main driving force behind the changes in ownership is the efforts towards privatization and restitution of forestland in countries formerly under centrally planned economies. Mixed ownerships or forests owned by institutions that are considered neither public nor private, like churches or charities, are extremely rare (Forest Europe *et al.,* 2011, 109 ff., fig. 67 and 68).

The responsibility and jurisdiction concerning forest management is as fragmented as the forested landscape itself or its ownership. Although most European countries have a national forest programme or a similar process in place, formally endorsed sub-national forest policy documents are still very important in several countries. However, many countries changed their legal/regulatory framework over the last few years, with most changes affecting silvicultural practice (often related to regeneration/tending, biodiversity provisions), enshrining institutional reorganization, and reorganizing financing arrangements. The most directly relevant international commitments are the European Union Regulations and Directives on forests, which have heavily influenced a range of national regulations in Member States (Forest Europe *et al.,* 2011, 145 ff.). The declining importance of local communities and regional authorities in favour of uniform EU rules should be used by archaeologists respectively heritage conservators as an opportunity to represent their interests with a common Pan-European voice as well.

4.3. Proposed European network

For this reason, the authors, in collaboration with interested partners across Europe, have endeavoured to establish a European network. Under the title ARCFOR (Recording and preserving archaeological heritage in forests), a network of archaeologists and heritage professionals engaged in woodland archaeology from several EU and associated countries is currently being established and is open to further interested colleagues. The main challenge will be to secure financial support. While various network members have been very successful in raising funds for individual research projects on a national, sometimes even on a European level, to our knowledge collaborative projects between more than two countries do not exist at present. We are currently exploring different funding opportunities on the European level to support transnational exchange and cooperation. Building on existing collaboration of researchers based at universities and heritage agencies who share a common interest in archaeological research in woodlands, the proposed European network will aim at the following goals:

To promote the use of modern technologies for archaeological prospection in woodlands.

- To foster communication, exchange and collaboration in order to streamline past and ongoing

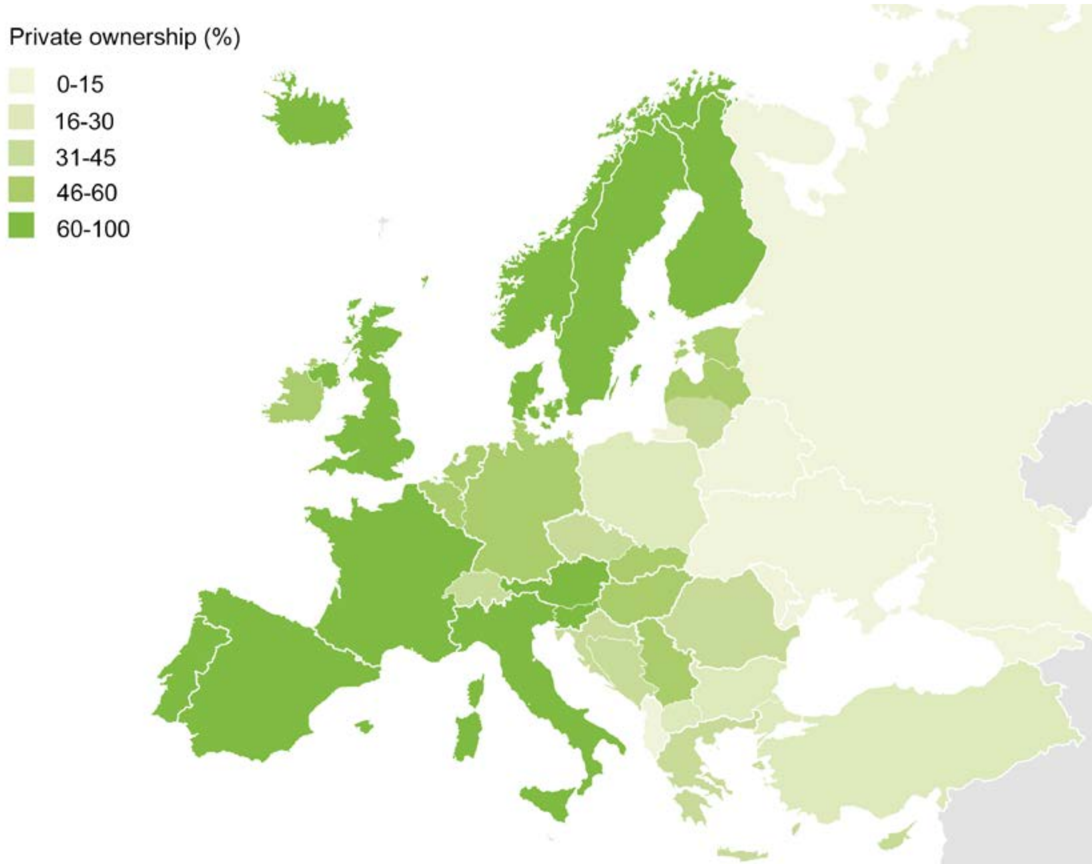


Figure 5 : Area of privately owned forest as percent of total forest. (after Forest Europe et al. 2011, 109)

research, and coordinate future activities, including training of students and young professionals, and the development of best practice.

- To involve other stakeholders engaged in the use and management of woodland resources, such as local communities, private and public land owners, forest industry and authorities, environmental agencies, as well as recreational and leisure activities and tourism; and
- To jointly negotiate, develop, test, implement, and disseminate best practice that enables the sustainable use of cultural, economic and ecological woodland resources by reconciling the interests of heritage conservation, woodland ecology, forestry, tourism and recreation.

The contacts and collaboration that the network seeks to establish between archaeologists and heritage professionals across Europe, and between them and other stakeholders engaged in forest management, are expected to provide a continuous platform from which an impact on public perception and policy making can be achieved long-term. It is supposed to give archaeology and cultural heritage management the much-needed additional weight when negotiating the sustainable use of forest resources with other stakeholders. The presentation of this initiative during a session on woodland archaeology at the CAA meeting in Paris in April 2014 is expected to further promote international cooperation to raise awareness of the important cultural legacy of woodlands.

Acknowledgements

The ideas expressed here are based on extensive discussions with European partners over the last two years. While we are grateful for their input, we assume full responsibility for all remaining errors. We thank the organisers of the session on woodland archaeology at the CAA conference in Paris in 2014 for inviting us to present this overview paper. We also thank Margaret Sloan and Anna Dowden-Williams for proofreading the manuscript.

References

ANGELSTAM, P., AXELSSON, R., ELBAKIDZE, M., LAESTADIUS, L., LAZDINIS, M., NORDBERG, M., PĂTRU-STUPARIU, I. and SMITH, M. (2011) Knowledge production and learning for sustainable forest management on the ground: pan-European landscapes as a time machine. *Forestry* 84 (5). pp. 581-596.

BECK, A. (2011) Archaeological applications of multi/hyper-spectral data – challenges and potential. In: COWLEY, D. C. (ed.) *Remote sensing for archaeological heritage management. EAC Occasional Paper 5,* Brussels: Europae Archaeologia Consilium, pp. 87-98.

COWLEY, D. C. (ed.) (2011). *Remote sensing for archaeological heritage management. EAC Occasional Paper 5.* Brussels: Europae Archaeologiae Consilium.

CROW, P. (2004). *Trees and Forestry on Archaeological sites in the UK. A review document.* Forest Research agency of the Forestry Commission. [Online] Available from: http://www.forestry.gov.uk/pdf/FR_archaelogical_review.pdf/$file/FR_archaelogical_review.pdf

CRUTCHLEY, S. and CROW, P. (2009). *The light fantastic: using airborne lidar in archaeological survey.* Swindon: English Heritage.

DONEUS, M. (2013). *Die hinterlassene Landschaft – Prospektion und Interpretation in der Landschaftsarchäologie. Mitteilungen der Prähistorischen Kommission 78.* Vienna: Austrian Academy of Sciences Press.

DONEUS, M. and BRIESE, C. (2011). Airborne Laser Scanning in forested areas – potential and limitations of an archaeological prospection technique. In: Cowley, D. C. (ed.). *Remote sensing for archaeological heritage management. EAC Occasional Paper 5,* Brussels: Europae Archaeologia Consilium, p. 59-76.

EUROPEAN COMISSION (ed.) (2010). Green paper on forest protection and information in the EU: preparing forests for climate change. Brussels: European Commission.

EUROPEAN ENVIRONMENT AGENCY (ed.) (2010). 10 messages for 2010: Forest ecosystems. Luxembourg: Office for Official Publications of the European Union.

FOREST EUROPE, UNECE and FAO (eds.) (2011) *State of Europe's forests 2011: status and trends in sustainable forest management in Europe.* Oslo: Forest Europe Liaison Unit.

FORESTRY COMMISSION (ed.) (2011). Forests and historic environment: UK forestry standard guidelines. Edinburgh: Forestry Commission.

HAMBERGER, J. IRLINGER W. and SUHR, G. (2012). *In Boden und Stein: Denkmäler im Wald.* Freising: Bayerische Landesanstalt für Wald und Forstwirtschaft.

HESSE, R. (2013). The changing picture of archaeological landscapes: lidar prospection over very large areas as part of a cultural heritage strategy. In: Opitz, R. S. and Cowley D. C. (eds.). *Interpreting archaeological topography. 3D data, visualisation and observations. Occasional Publication of the Aerial Archaeology Research Group 5.* Oxford: Oxbow Books, pp. 171-183.

MÖLLERS, S. (2004). *Zum Schutz archäologischer Denkmale in Waldgebieten: Kooperationsmöglichkeiten von Denkmalpflege und Forstwirtschaft.* Available from: http://www.schnippenburg.de/fileadmin/pdfs/Schnippenburg_Moellers_2004c.pdf. [Accessed: 9th April 2014].

NEUBAUER, W., TRINKS, I., SALISBURY, R. B. and EINWÖGERER, C. (eds.) (2013). *Archaeological prospection: proceedings of the 10th international conference, Vienna, May 29th – June 2nd 2013.* Vienna: Austrian Academy of Sciences Press.

OPITZ, R. S. and COWLEY, D. C. (eds.) (2013). *Interpreting archaeological topography. 3D data, visualisation and observation.* Occasional Publication of the Aerial Archaeology Research Group 5. Oxford: Oxbow Books.

RISBØL, O. (2013). Cultivating the 'Wilderness' – how Lidar can improve archaeological landscape understanding. In: OPITZ, R. S. and COWLEY, D. C. (eds.) *Interpreting archaeological topography. 3D data, visualisation and observations.* Occasional Publication of the Aerial Archaeology Research Group 5. Oxford: Oxbow Books, pp. 53-64.

RITCHIE, M. and WORDSWORTH, J. (2010). *Identifying the historic environment in Scotland's forests and woodlands. Forestry Commission Scotland Practice Guide.* Edinburgh: Forestry Commission.

ROTHERHAM, I. D. (2007). The implications of perceptions and cultural knowledge loss for the management of wooded landscapes: a UK case study. *Forest Ecology and Management* 249, p. 100-115.

ROTHERHAM, I. D. and ARDON, P. A. (2006). The archaeology of woodland landscapes: issues for managers based on the case-study of Sheffield, England and four thousand years of human impact. *Arboricultural Journal* 29, pp. 229-243.

SIPPEL, K. and STIEHL, U. (2005). *Archäologie im Wald. Erkennen und Schützen von Bodendenkmälern.* Kassel: Landesbetrieb Hessen-Forst.

STROMER, W. (1968). 600 Jahre Nadelwaldsaat, die Leistung des Peter Stromer von Nürnberg. In: SPERBER, G. (ed.) *Die Reichswälder bei Nürnberg – aus der Geschichte des ältesten Kunstforstes*. Mitteilungen aus der Staatsforstverwaltung Bayerns. 37. Nürnberg: Frankenverlag Lorenz Spindle, pp. 25-29.

VEČEŘA, J. (2004). Povrchové pozůstatky po těžbě rud a jejich vyhodnocení (Surface remains after ore mining and their evaluation). *Medievalia Archaeologica* 6, pp. 145-156.

WISEMAN, J. and EL-BAZ, F. (eds.) (2007). *Remote sensing in archaeology.* New York: Springer.

Archaeological Mapping of Large Forested Areas, Using Semi-Automatic Detection and Visual Interpretation of High-Resolution Lidar Data

Øivind Due TRIER,[1] Lars Holger PILØ[2]

[1]Norwegian Computing Center, Section for Earth Observation, Oslo, Norway
[2]Oppland County Administration, Section for Cultural Heritage Management, Lillehammer, Norway
trier@nr.no, lars.holger.Pilo@oppland.org

Abstract

This paper presents results from the on-going mapping of cultural heritage in Oppland County, Norway, based on airborne lidar scanning of forested land.

Automatic detection is used in combination with manual inspection of visualizations of the ground surface, derived from the lidar data, and targeted field survey. In this manner, large forested areas in Oppland County, Norway have been mapped for ancient monuments, especially moose hunting systems and iron extraction sites.

Airborne lidar data is now being used extensively in the mapping of cultural heritage in forested areas in Oppland County, Norway. Fieldwork may be done in a much more targeted and limited way than traditionally. If the lidar point density is sufficiently high, prognosis mapping of archaeological pits may be done without fieldwork. Much larger areas may be mapped than before the introduction of airborne lidar scanning, at the same time providing a much more accurate and complete mapping.

Keywords: Airborne Laser Scanning, Template Matching, Moose Hunting Systems, Iron Extraction Sites.

1. Introduction

Oppland County, Norway, has several on-going and planned road construction projects, and some areas are being zoned for building of mountain cottages. Also, many areas have commercial timber production. To reduce conflicts between cultural heritage protection and modern land use development, Oppland County Council is conducting a much more accurate and complete mapping of cultural heritage than has been done previously.

Starting in 2010, the Cultural Heritage Department of Oppland County Council is currently conducting a large project, using high-density airborne laser scanning (ALS) to map ancient monuments. Approximately 4100 km² were mapped in 2013, mainly forested areas, bringing the total up to about 5700 km². Automatic pit detection has been used to support visual inspection of a digital terrain model (DTM) generated from the ALS data.

For visual inspection of a DTM, a number of visualization methods exist, including the standard hill-shading and slope images, which are available in many software packages like ENVI and ArcGIS; sky-view factor (Kokalj, 2011) and local relief models (Hesse, 2010). Many of the main lidar visualisation techniques are available in the LiDAR Visualisation Toolbox (LiVT) at http://www.arcland.eu/outreach/software-tools. Several authors have mapped cultural heritage by visual inspection (e.g., see Bewley, 2005; Bollandsås, 2012)). However, for automatic detection of pit structures, no suitable method existed to our knowledge, so we developed our own method (Trier, 2012).

Oppland County Council undertook three separate ground-thruthings of objects in 2013, based on ALS data from 2012. The purpose was to investigate whether data collected during visual inspection of ALS data, supported by automatic detection, was of a sufficient quality to allow it to be entered into the national database of ancient monuments without a field control.

2. Data

Large forested areas in Oppland County, Norway, are known to contain ancient moose hunting systems and iron extraction sites. Today, these are manifested as pits in the ground. The iron production sites were used 1400–700 years ago, and consist of charcoal burning pits, often located in groups of three or more around a central oven. The hunting systems were used 2000–500 years ago, and consisted of concealed pitfall traps and wooden fences, located on moose trekking routes. The fences are gone, but the pits remain.

We have received ALS data in the form of LAS files, containing (x, y, z) points with the following information:

- x, y, z coordinates in UTM zone 32 and minimum 10 cm accuracy
- return number: 1-4 (each emitted lidar pulse may have up to four discrete returns)
- class label: ‘ground’ or ‘other’
- intensity

The three areas chosen for fieldwork are situated in Nord-Fron and Sør-Fron municipalities in the central part of Oppland County (Figure 1). The ALS data was acquired in 2012 and has at least five emitted pulses per square meter, i.e. the number of first returns per m² is 5 points or better. The individual points have a precision of better than 10 cm. The ALS data for all the three areas belong to the same scanning-project. The laser-scanning instrument was a TopEye System, with a frequency of 200 000 Hz, mounted on a helicopter.

The density map of ground returns shows that nearly all the areas have more than 2 ground returns pr. m² (colour coding in Figure 2-Figure 4). Some parts have more than 5 ground returns, due to overlapping fields of view from neighbouring flight lines.

The three areas are Stølssletta, Venlisætra and Fagerlisætra. The Stølssletta area in Nord-Fron municipality is situated in a forested area adjacent to modern farm settlement in the Skåbu Valley at an altitude of 760-910 m.a.s.l. The checked area covered 2.5 km² (Figure 2). The hilly terrain is sloping towards the Northeast. The area contains single charcoal pits of a medieval date, and occasional pit fall traps, dating to the Iron Age and Medieval Period.

The Venlisætra area in Sør-Fron municipality is a summer farm area at an altitude of 720-960 m.a.s.l. It covers 1.6 km² in a Northeast-facing slope (Figure 3). It contains Medieval iron extraction sites with clusters of charcoal pits, single charcoal pits from the Medieval Period, pitfall traps from the Iron Age and Medieval Period, and occasional above ground charcoal kilns mainly dating to between the 17th and 19th century.

The Fagerlisætra test area in Nord-Fron municipality (Figure 4) is situated in a lightly forested area with summer farms. It covers 1.1 km² at an altitude of 870-930 m.a.s.l. The area contains medieval iron extraction sites with clusters of charcoal pits, single charcoal pits from the Medieval Period, and occasional above ground charcoal kilns mainly dating to between the 17th and 19th century.

3. Methods

A computer-based image analysis system is used in combination with manual inspection of visualizations of the ground surface, derived from the lidar data, and targeted field survey. In this manner, large forested areas in Oppland County, Norway have been mapped for ancient moose hunting systems and iron extraction sites.

3.1 Automatic pit detection

The automatic method of (Trier, 2012) for the detection of pit structures in ALS data was used, and is summarised here. We received airborne laser scanning data in the form of LAS files (described above). The (x, y, z) points labelled as ‘ground’ were used to form a triangular irregular network, which was then used to form a digital terrain model with 20 cm grid spacing. Then, template matching by image convolution, also called image matching by correlation (GONZALEZ, 1992) was used to locate candidate pit locations. Each pit template is a hemisphere with a ring edge. A range of template sizes were used, covering the plausible pit sizes. Overlapping template matching detections were merged, keeping only the strongest one in each case.

The intermediate result after template matching is a large number of possible pit locations; however, most of these are false detections. The list of possible locations needed to be sorted, with the most likely pit detections at the top. For this purpose, a number of measurements were computed for each detection. The measurements were designed to quantify various deviations from the expected pit shape, and included: convolution value divided by radius, minimum and average depth, root-mean-square difference from pit template, eccentricity and elongation. Thresholds were used on the measurements to assign a confidence level (0-6) to each pit detection:

0 = not a pit, deleted
1 = pit with very low confidence
2 = pit with low confidence

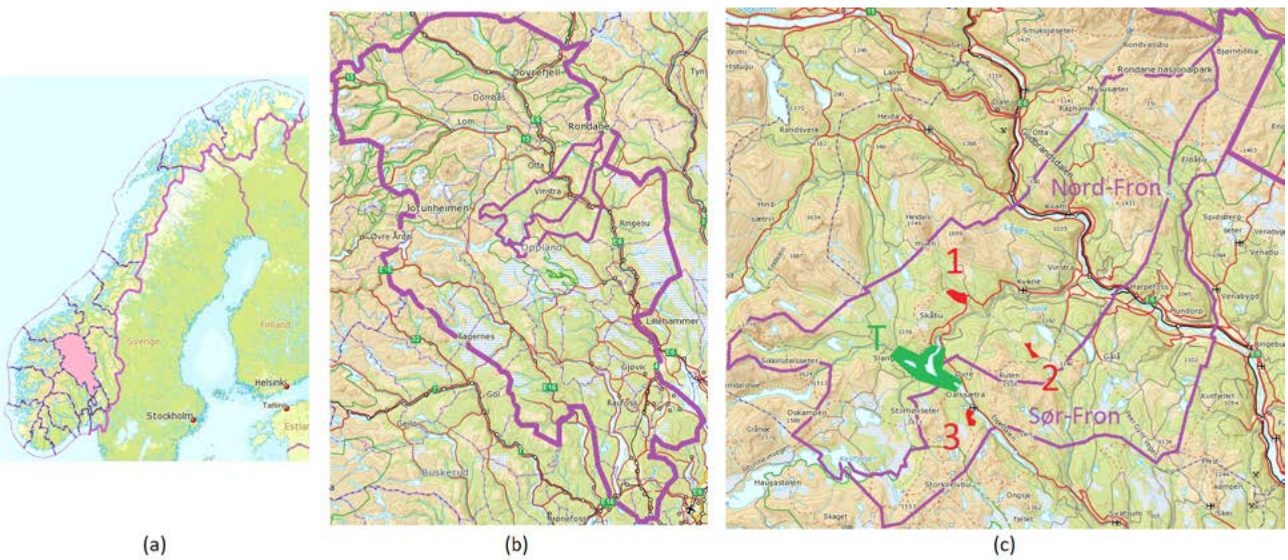


Figure 1: The location of the three fieldwork areas and the training data area. (a) Map of Norway, with county boundaries, and Oppland County in pink. (b) Oppland County, with the borders of Nord-Fron and Sør-Fron municipalities highlighted. (c) The location of the areas in Nord-Fron and Sør-Fron municipalities; 1 = Stølssletta, 2 = Fagerlisætra, 3 = Venlisætra, T = training data.

3 = pit with medium confidence
4 = pit with medium high confidence
5 = pit with high confidence
6 = pit with very high confidence.

The thresholds were set manually by ordering a training set on one measurement at a time, and considering how many true archaeological pits could be missed at each confidence level. The training set consisted of 129 true pitfall traps and 1000 false pit detections from a 29 km² area surrounding the lake Olstappen, in Nord-Fron and Sør-Fron municipalities (Figure 1).

The pit detections of each confidence level (1-6) were exported to a separate GIS layer. For a more detailed description of the method, please see (Trier, 2012).

3.2 Visual inspection

A highly detailed model was constructed from the ALS ground points, with a grid size of 0.25 m. The software used was Quick Terrain Modeller. An additional generalized model was constructed, with a 2 m grid size. The generalized model was then subtracted from the highly detailed model as a change analysis with an interval of 0.5 m. In this way a local contrast was produced, showing very local differences in height as differences in colour instead of as shadows, as in the case of hill-shade. In many ways this represents a local relief model (Hesse, 2010). However, the local relief model is more complicated, and time-consuming, to compute.

Visual inspection of the local contrast took place using one screen with an exported 2D-geotiff of the local contrast (azimuth 0, elevation 55) in ESRI ArcGIS and one screen with a 3D local contrast model in Quick Terrain Modeller. The visual inspection was supported by automatic pit detections marked in the ArcGIS-project. All automatic detections were visually checked on the screen. Some were interpreted to be ancient monuments or anomalies, while others were discarded during the inspection process. The automatic detection method produces a varying number of false detections depending on topography, data quality and modern activity. The main purpose of using automatic detections during the visual inspection is to achieve a consistent quality of archaeological data, cutting down on human error during inspection, i.e. missing objects.

Objects are marked in two different shape-files during visual inspection. Objects that are believed to be ancient monuments (based on experience) are geo-referenced with a point in the centre of the objects in an 'Ancient monument' shape-file. Other objects that could possibly be ancient monuments, but where interpretation of the DTM is more uncertain, are marked in an 'Anomaly' shape-file. The main reason for splitting the objects into two groups is that the 'Ancient monument' shape-file is made available to area-planners and other interested parties through a public website (http://open.innlandsgis.no/), while the 'Anomaly' shape-file contains too many false objects to be of use to planners.

3.3 Fieldwork

The purpose of the fieldwork is to assess to what extent data collected during visual inspection of ALS data, supported by automatic detection, is of a sufficient quality to allow it to be entered into the national database of ancient monuments without a field control.

In all three areas, a single archaeologist, using a handheld GPS with DPOS correction, undertook ground-truthing. The GPS contained a GIS-project, including the objects both from the 'Ancient monument' and the 'Anomaly' shape-file. To avoid visual clutter on the small GPS-screens the automatic detection data were not included.

Ground-truthing was undertaken by walking from 'Ancient monument' point to point and checking each object. Anomalies were also checked, but to a varying degree in the different areas (see below). No systematic surface survey was undertaken, but the terrain was surveyed when walking between checked objects, yielding some impression as to the presence of visible objects, not found during the visual inspection of the ALS data.

4. Results

Fieldwork at Venlisætra and Fagerlisætra was undertaken by Lars Pilø, while the ground truthing at Stølssletta was undertaken by Anna McLoughlin.

Anomalies were systematically checked during the ground-truthing at Stølssletta (Table 1, Figure 2). No additional objects were found during fieldwork.

Objects	True	False	Total	% True
Ancient monuments	100	5	105	95,2%
Anomalies	8	22	30	26,7%
Total	108	27	135	80,0%

TABLE 1: AUTOMATIC DETECTIONS VERSUS FIELD VERIFICATION AT STØLSSLETTA.

For Venlisætra, anomalies were checked in the first part of the survey, but as the first 11 yielded no true objects, the checking of these objects was discontinued, to allow for a larger number of checked objects in the 'Ancient monuments' group (Table 2, Figure 3). No additional objects were found during ground truthing.

Only two anomalies were checked in the Fagerlisætra area (Table 3, Figure 4). Both were small charcoal pits, adjoined to the same iron extraction site. Two additional objects were discovered during the ground-truthing. Both had been categorized as confidence level 2 (low confidence) by the automatic pit detection, but had been discarded during the visual inspection of the ALS data. As the automatic detections had not been included in the field GIS-project, it was only later discovered that they had in fact been targeted by the automatic detection.

To summarize, 197 of 205, or 96% of the objects that were interpreted as ancient monuments from visual inspection

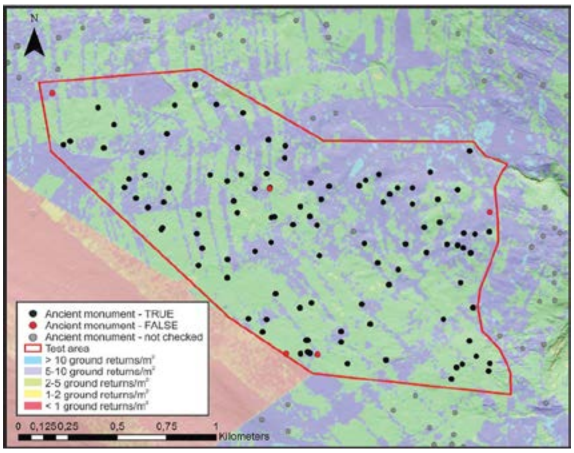


Figure 2: Result of field inspection at Stølssletta, Nord-Fron municipality.

turned out to be true ancient monuments according to the field inspection (Table 4). In addition, some anomalies detected by visual inspection turned out to be true ancient monuments by field inspection.

5 Discussion and conclusions

Objects	True	False	Total	% True
Ancient monuments	60	3	63	95,2%
Anomalies	0	11	11	0,0%
Total	60	14	74	81,1%

Table 2: Automatic detections versus field verification at Venlisætra.

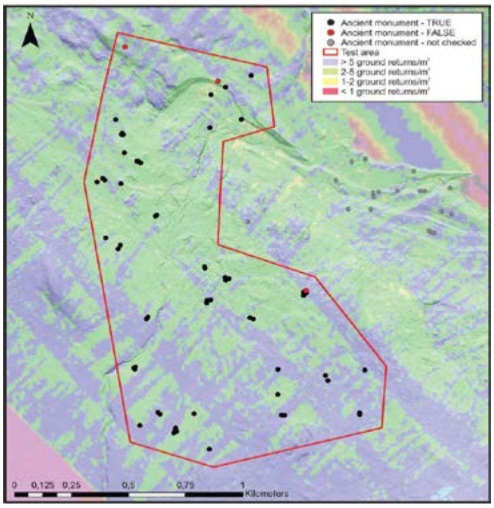


Figure 3: Result of field inspection at Venlisætra, Sør-Fron municipality.

The three checked areas are relatively similar. They are situated at an altitude of above 700 m.a.s.l. Two areas (Stølssletta and Venlisætra) are situated in a Northeast-facing slope while Fagerlisætra has a slightly undulating topography. All three areas showed a remarkable consistency in producing 95% or better of true objects in the 'Ancient monuments' category.

Objects	True	False	Total	% True
Ancient monuments	37	0	37	100,0%
Anomalies	2	0	2	100,0%
Total	39	0	39	100,0%

Table 3: Automatic detections versus field verification at Fagerlisætra.

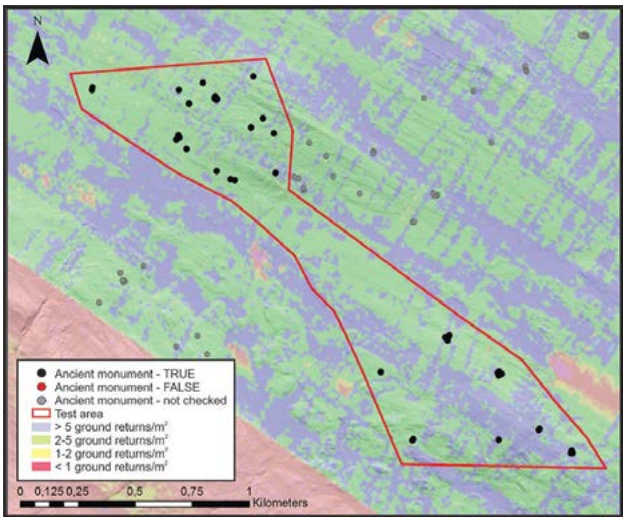


Figure 4: Result of field inspection at Fagerlisætra, Nord-Fron municipality.

On the basis of the evidence it seems reasonable to conclude that typical outfield monuments like charcoal pits and pitfall traps can be accurately mapped using visual inspection of detailed DTM, supported by automatic detection, provided that the number of ground returns is sufficiently high (> 2 ground returns pr. m² (TRIER, 2012)). It should thus be possible to produce large-scale maps of these monument-types, provided that ALS data is available. This work is already on going in Oppland County.

This study does not give much information on the presence and number of visible ancient monuments not found during the visual inspection of the DTM. Based on the limited surveys conducted by walking in the terrain between the checked objects the number of additional objects in the test areas is believed to be low. However, information from previous ground-truthing in Gravfjellet in Øystre Slidre municipality provides more systematic information on this question. This area also contained medieval iron extraction sites with charcoal pits in clusters, single charcoal pits and some pitfall traps. A 70 km² ground-truthing in 2011, based on visual inspection of a one-light-source hill-shade (and with less experience in interpretation of ALS data than now), led to the discovery of 1650 visible ancient monuments. In 2012 plans were initiated to develop about 10 km² of this area for cottages and infrastructure. This provided an opportunity to do systematic surface survey of this limited area (Tveiten, 2013). An additional ca. 10 % of single charcoal pits were found during the systematic survey, which is a remarkable low number, considering the visual inspection was undertaken using hill-shade, which is not really a suitable visualization technique for this

	Visual inspection: objects interpreted as ancient monuments	Field inspection		
Test area		True	False	% True
Stølssletta	105	100	5	95,2%
Venlisætra	63	60	3	95,2%
Fagerlisætra	37	37	0	100,0%
Total	205	197	8	96,1%

Table 4: Summary of results for the three areas.

kind of work. The impression of the distribution of single charcoal pits did not change after adding the extra objects. The additional pits were typically small, hidden in dense spruce forest or damaged.

As of 2013, the new ALS-based mapping has covered 5700 km² of forested land. At Gravfjellet in Øystre Slidre municipality, a 70 km² area has been zoned for mountain cottages. A total of 1650 archaeological features, mostly charcoal pits belonging to iron extraction sites, have been mapped and ground-proofed. This enables the municipality to plan the location of individual cottages, local roads, etc., in order to minimize the destruction of cultural heritage. In Gausdal municipality, prognosis mapping (as yet without ground-proofing) of 290 km² revealed about 1800 cultural heritage objects. These figures illustrate the density of cultural heritage in some areas in Oppland County, and the need for detailed archaeological mapping.

The purpose of using ALS in the heritage management in Oppland is mainly as a tool to map outfield monuments. In this it succeeds brilliantly, providing inexpensive, good quality and above all systematic data, which is of great value for cultural heritage management. It is, however, not a substitute for proper fieldwork on the ground, when an area is scheduled for development, as not all visible monuments are found in the ALS data, detail on some objects is lacking, and of course not all sites are visible above ground. In the forested areas of Oppland, however, most monuments are clearly visible above ground, making the interpretation of ALS data a very valuable tool for cultural heritage management. Under such conditions, ALS data provides excellent mapping opportunities for large areas (Figure 5-Figure 6).

In conclusion, airborne lidar data is now being used extensively in the mapping of cultural heritage in forested areas in Oppland County, Norway. The results of combining semi-automatic detection and visual inspection are good. Fieldwork may be done in a much more targeted and limited way than traditionally. Provided that the

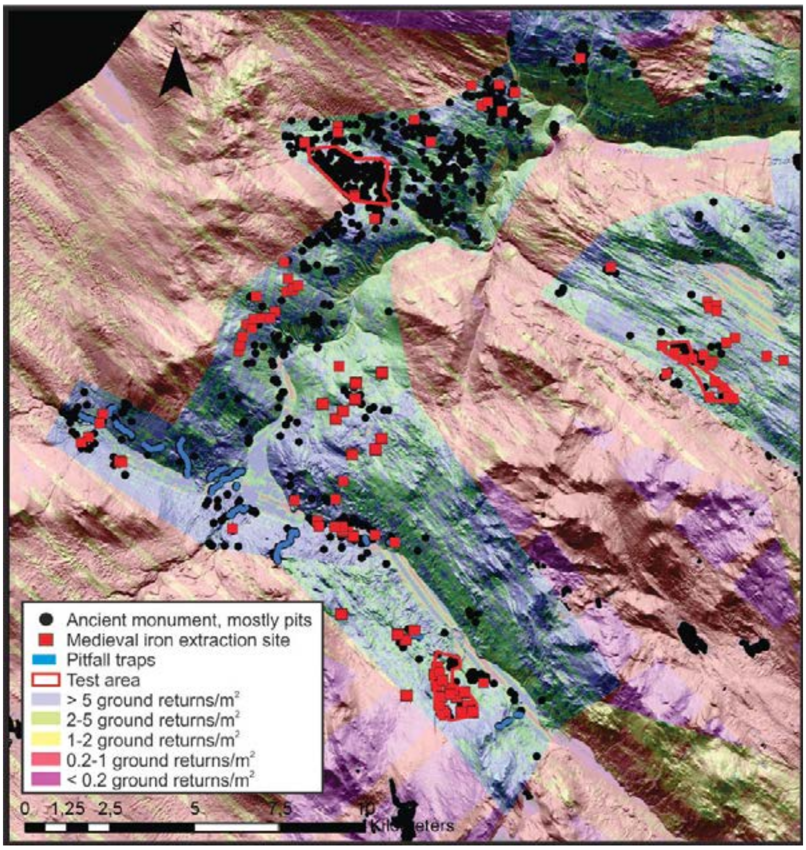


Figure 5: Overview of ancient monuments mapped from ALS data in Nord-Fron and Sør-Fron municipalities. The three field survey areas (test areas) are indicated with red polygons.

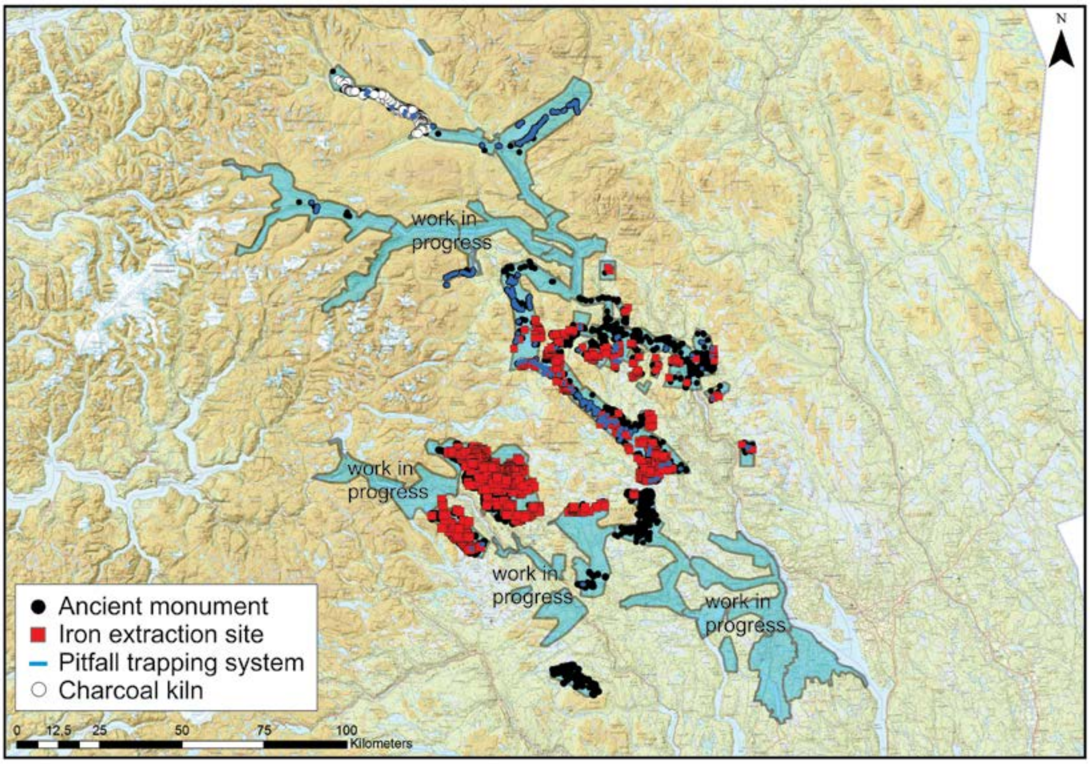


FIGURE 6. OVERVIEW OF ALL ANCIENT MONUMENTS MAPPED SINCE 2010 BASED ON ALS DATA.

quality of the lidar data is sufficiently high, that is, the digital elevation model derived from the lidar ground returns contains sufficient detail; prognosis mapping of archaeological pits may be done without fieldwork. Much larger areas may be mapped than before the introduction of airborne lidar scanning, at the same time providing a much more accurate and complete mapping of cultural heritage.

Acknowledgements

This research was funded by the Directorate for Cultural Heritage in Norway (Riksantikvaren).

References

BEWLEY, R. H., CRUTCHLEY, S.P. and SHELL, C.A., (2005). New light on an ancient landscape: lidar survey in the Stonehenge World Heritage Site. *Antiquity* 79, pp. 636-647.

BOLLANDSÅS, O. M., RISBØL, O., ENE, L. T., NESBAKKEN, A., GOBAKKEN, T. and NÆSSET, E., (2012). Using airborne small-footprint laser scanner data for detection of cultural remains in forests: an experimental study of the effects of pulse density and DTM smoothing. *Journal of Archaeological Science* 39, pp. 2733-2743.

GONZALEZ, R. C. and WOODS, R. E., (1992). *Digital Image Processing.* New York: Addison-Wesley, pp. 583-585.

HESSE, R., (2010). Lidar-derived local relief models – a new tool for archaeological prospection. *Archaeological Prospection* 17, pp. 67–72.

KOKALJ, Ž., ZAKŠEK, K. and OŠTIR, K., (2011). Application of sky-view factor for the visualization of historic landscape features in lidar-derived relief models. *Antiquity* 79, pp. 263-273.

TRIER, Ø. D. and PILØ, L. H., (2012). Automatic detection of pit structures in airborne laser scanning data. *Archaeological Prospection*, 19(2), pp. 103-121.

TVEITEN, O. and PETTERSSON, S., (2013). *Kulturminne i Gravfjellet. Kulturhistoriske registreringer i Gravfjellområdet, Øystre Slidre kommune i Oppland. (In Norwegian.), Lillehammer,* Norway: Oppland Fylkeskommune.

Laser scanning and Automated Photogrammetry for Knowledge and Representation of the Rupestrian Architecture in Cappadocia: Sahinefendi and the Open Air Museum of Goreme

Marco CARPICECI, Carlo INGLESE

Dipartimento Storia, Disegno e Restauro dell'Architettura - SAPIENZA University of Rome (DSRA), Italy
marco.carpiceci@uniroma1.it, carlo.inglese@uniroma1.it

Abstract

The survey of the '40 Martyrs Church' in Sahinefendi and churches located within the Goreme Open Air Museum is part of a research project oriented to test systems for surveying and representation applied to the architecture excavated rock. The data acquisition using old and new technologies and processing in post- production involves a necessary critical action that affects not only each specific monument, but the same methodology of acquisition and processing. In particular, we tackled the problem of systematization of procedures for the architectural survey in the field of rock art according to the 'emerging' technologies like Laser scanning, automated photogrammetry and immersive photography. The rock architecture, unlike the one built, has the need to represent its particular, irregular, morphological feature, so it was necessary to conduct a trial of representation techniques unusual for architecture, most commonly used in cartographic representation, as contour lines and planar development of complex surfaces.

Keywords: Cappadocia, Rupestrian Architecture, Laser Scanning, Automated Photogrammetry

Premise

We know that these are concepts taken for granted; but we must continuously remind ourselves that architectural survey is a 'critical form' of knowledge, a way in which we organize and put together a set of sizes, of analysis and of observations that make up a 'discovery'. We know because we meet something again; discovery is a form of iterative deepening; we observe, we reveal, what we know becomes clear, we recognize, we re-reveal what we recognize, and we observe again. There are no alternatives to the operating practices that bind man to historical-architectonic goods, and from this 'union' knowledge is generated. The human factor, therefore, is not a 'frill' to be avoided but is instead the means. Certainly methods of measurement tend to extract more and more reliable objective data, but every technology brings with it diverse characteristics from the others and at the same time its own peculiarities to take advantage of and to interact with so as to obtain a 'homologation' of a final product. LASER and digital technologies have reached a global level and dimension just in this decade that has made them easily usable, with previously unthinkable precision.

Revolutionized Technologies

Think about digital photography, which until recently had to share its laborious evolution with chemical photography, well established and stable. Today cameras and software can provide instruments of precision and processing of data for any application, for any 'idea'. With chemical photography we had the constant problem of exposure range. Today the chromatic range of an image can be increased enormously with HDR (High Dynamic Range). Radial distortion of the optics was one problematic characteristic for architecture and accurate reproduction, and only the photogrammetric optics had the 'costly' advantage of high correction. Nowadays every perspective, even with varying focus, can be potentially associated with a specific transformation that tends to cancel out any deformation in any shooting condition. With a quantity of photodiodes near to 40 megapixels on full-frame, digital sensors have come to a definition, reaching and accommodating optical capabilities that until recently remained overabundant and therefore not taken advantage of (Carpiceci, 2012a).

Digital photography and computer science then provided the possibility of evolution of the old stereo-photogrammetry into photomodellation. The refinement of the software of digital processing has permitted more and more of an increase in the accuracy of the algorithms based on epipolar geometry. The latest evolution of the architectonic survey, however, is certainly due to the laser-scanner, which only recently has gained the possibility of 'transportability'. With an instrument of 5 kg carried in a briefcase, resting on a carbon fibre photographic tripod, it's finally possible to carry out scans in (almost) any place on earth.

And now drones are ever more insistently and forcefully at the door of architectonic survey.

We are therefore witnessing a phase of 'epochal' transition. The instrumentation at our disposal is evolving exponentially. And just as quickly it's changing the horizons of possibility that such technological growth gives us.

But it's just that precision and speed of data acquisition that can be helpful in those fields of research that were difficult to explore up until now. Rocky habitats are one of those fields where it has been difficult to obtain a precise architectonic survey that is correctly 'representative' of the subject.

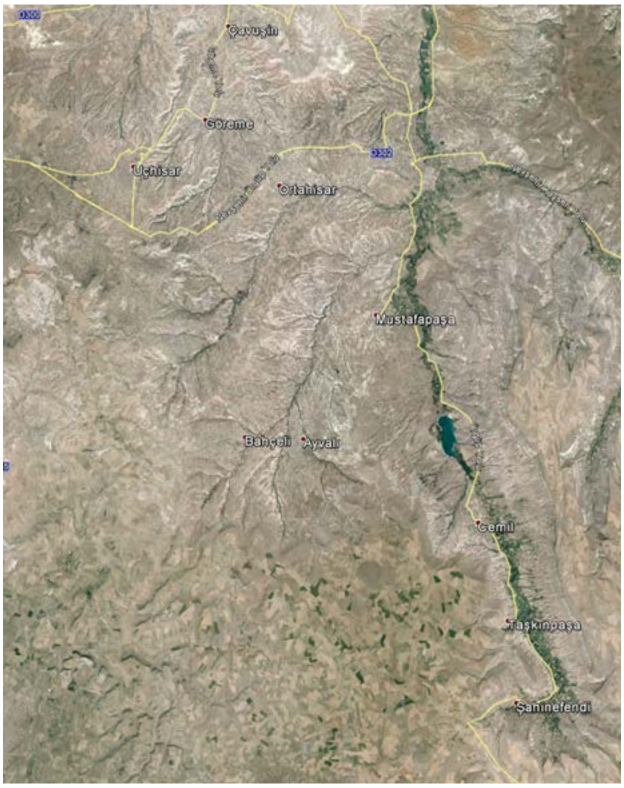


Figure 1: Turkey, Cappadocia, Goreme and the area between Sahinefendi.

You press the button, we do the rest

In 1889 Kodak used the motto 'You press the button, we do the rest' as an advertising phrase at the launch of automation in photography. Today the laser scanner is often considered in the same way as a Kodak camera, where it's enough to press the button 'scan' to obtain 'the rest'. And today it's often thought that panoramic photos are taken with a cellular telephone, assisted by software that recommends what movements we make. Today, it's often thought that the cloud of coloured points or the mapped mesh is architectonic survey but, unfortunately or fortunately, that's not the case.

The complexity of the technologies and their correct usage implicate a deep knowledge of their 'geometric' structure, in order to proceed with skill and a precise and correct realization of architectonic survey (Carpiceci 2012b, Carpiceci 2013a).

From the required concatenation of the scans in such a way to limit undetected zones as much as possible, to the realization of images for a correct 'chromatic mapping' of the mesh.

From the production of spherical photos through panoramic heads regulated on the front nodal point of the optic used, to the regulation of exposures as a function of the dynamic range of the subject (Carpiceci, 2011).

From the processing of metric data for the realization of a numerical uniform model, to the completion of elaborate

Figure 2: Turkey, Cappadocia, Sahinefendi, the rock village on the slopes Orta Tepe.

graphics capable of describing the chosen architectural structures completely.

A careful analysis of the processes of survey and representation is essential, in light of modern technologies. This 'revision' allows us to consider the various potentials, in order to then take advantage of them with awareness in all the phases of analysis, processing and of communication.

From Sahinefendi to Goreme

The survey campaign in Turkey is concentrated on two significant places: the rupestrian village of Sahinefendi and the Open Air Museum of Goreme (Figure 1). These are two different and characteristic places (Carpiceci, 2013b).

In the valley of Sahinefendi, as in other valleys, erosion of crumbling land and rocks not worn down has generated peculiar forms of cones topped by blocks of hard, dark rock (Figure 2). These pinnacles present a coniform aspect, with walls sloping about 67 degrees, sometimes topped with a chapel of more resistant tuffaceous material.

The ease of working this material has facilitated its excavation by humans, who over time have created a series of areas with still-recognizable functions. Residential areas, places of production and of commerce such as the 'pseudo wineries' which have clearly recognizable tubs for the deposit and compression of the fruit, places for settling and maturation, and places of storage, conservation and marketing. The places of worship are also very important and characteristic. The most important and complex is the Monastery that is found in the high zone in the North beneath the plateau (the Orta Tepe), with a central body made by the excavation of a cliff front and some conical conformations, where we can find convent areas and a chapel (Figure 3).

The church of the Forty Martyrs, towards the valley, is certainly the best-preserved place in Sahinefendi, given that a painting cycle is still present in its cone, datable to the

Figure 3: Turkey, Cappadocia, Sahinefendi, the Monastery.

eleventh century (Thierry, 1963; Thierry, 1975, Figure 4). Its restoration was concluded in 2013 by a large group of Italian and Turkish restorers, led by Professor Maria Andaloro from the University of Tuscia. The church is set in a cone of about 20 by 10 meters. A series of rooms are divided over three main levels. The lower level is for the most part underground and is composed of two rooms of approximately circular shape and a flat roof and a third, slightly higher and a bit smaller, having a parallelepiped form. The middle level is the main level, where the church was developed, with two naves, ending in apses and with barrel vaults. The third level is made up of two small circular rooms that face towards the outside and into the naves of the church.

In 2008, restoration was begun on the paintings. These restorations were completed with the church's transformation into a museum in 2013. During this time, Sapienza University of Rome with its Department of History, Drawing and Restoration of Architecture, specifically Professor Marco Carpiceci, carried out some survey projects aimed at the measurement and knowledge of the monument through laser scanning and immersive photography. In the latest survey campaign (2013), Professor Carlo Inglese and Architect Giovanna Cresciani were added to the collaboration. With the coverage of the most significant places through laser scanning, we wanted to embrace the area no longer simply on the singular episode but instead on a co-contextual framework; a sort of relation of the single 'cone' with those around it.

The Open Air Museum of Goreme represents a complex of unique anthropological value on top of its art historical value, declared a UNESCO World Heritage Site in 1985, for which there is not currently a workforce for architectural survey (Figure 5).

The area of Goreme is extensive and includes rocky conformations of various types. It's the biggest spread of cave habitats with a monastic function (Rodley, 1985).

The heart of the area is made up of a large natural rock hemicycle, with a radius of about 40 meters and a height of 15-20. The large exedra overlooks a series of conical shapes of various sizes (Figure 6).

Figure 4: , Cappadocia, Sahinefendi, Church of the Forty Martyrs, left aisle.

Figure 5: Turkey, Cappadocia, Goreme, the area of the Open Air Museum.

From an anthropological and functional point of view, the Museum makes up a complex aggregation of small groups of 20-40 people. Every group was independent from the others, was distinguishable by an autonomous series of rooms: a place of worship (church or chapel); a refectory; a kitchen; areas used for everyday activities and for rest.

Figure 6: Turkey, Cappadocia, Goreme Open Air Museum, the hemicycle natural rocky.

The historical hypotheses tend to date most of the wall paintings to between the tenth and eleventh centuries (Thierry, 1963; Thierry, 1975). This, therefore, must have been the moment of maximum splendor of this great monastic city. An orderly community of autonomous nuclei of various sizes, but all belonging, evidently, to one general organization.

In the exedra, the most numerous community was the Karanlik Kilise (dark church) that would go on to represent one of the principal objectives of the survey campaign of 2014. Slightly less extensive, but still important from an iconographical and architectural point of view, is the Carikli Kilise (church of the sandal) that was surveyed in 2013 together with the chapel of St. Catherine. In the cones in front of the hemicycle, there are a number of churches detected between 2011 and 2013: St. Barbara; Elmali Kilise (church of the apple) and St. Basilio. A bigger cone, finally, is thought to possibly have been used as a female community. Inside it, there is a church detected in 2013 called just that, Kizlar Kilisesi (church of the girls).

Outside the Open Air Museum there is Cappadocia's most significant church: Tokali Kilise (church of the buckle), in which there is a mosaic cycle on a lapis lazuli background from the eleventh century. Tokali underwent survey in 2013.

Acquisition Techniques

The morphological prerogative of excavated architecture is that this, apparently attributable to forms similar to 'built' forms, in reality is nearer to a sculpture, where we can make out features and similarities with another reality, but that instead is the result of the aggregation of complex shapes.

And so in the traditional measurement practices (for constructed architecture) we proceed to the scanning chain, numerical model, geometric model, and representation for orthogonal projections. In rocky habitats we face an almost total impossibility of determining the geometric model, and the only intervention on the numerical model consistent with a decimation of the point cloud where the formal variation allows for a descriptive simplification.

A determining factor for the execution of this type of survey is given by the transportability of the instrumentation, which in our case is 'unique' and depends on the size and the weight of 5 kg of the scanner FARO Focus3D. With this type of scanner we were able to easily 'climb' down the cliffs without needing to transport bulky crates of several tens of kilograms.

Another decision, in order to take advantage of the portability of the instruments, consisted in giving up the Total Station. We tried, as much as possible, to use targets or spheres scattered around the subject to make the stations interrelated without the necessity of closed (almost impossible) or open polygons for their topographical 'structuring'. The recognition of homologous points, sometimes unspecific but seen in multiple scans, was useful in the 'registration' of the clouds. Basically the laser scanner measures points according to a polar geometric concept (distance, horizontal angle, vertical angle), like the Total Station, and therefore by excluding the topographical instrument we just recognized the functions already included in scanning. It should be added that the easily possible 'compensations' only in closed polygons, in the registration of the clouds are carried out on a great number of recognized points, automatically by the software (Carpiceci, 2012b).

A fundamental and peculiar aspect that we had to confront was the survey of painted surfaces. The paintings should have their chromatic aspect registered as well, but this is wishful thinking. We already know that dimensional measurement registers one unique figure, more or less precise but constant in time with possible dimensional variations normally infinitesimal and mostly due to the metric tolerance of the scanner. For colour, we enter in a world made of variations, starting from the fact that we record reflected light and therefore the chromatic aspect of the subject is influenced by the lighting conditions. But the thing we'd like to record is not a particular lighting condition but exclusively the colour of the subject.

These considerations make us realize how complex a registration that 'excludes' or minimizes the factor determined by lighting could be. Try to think of an external corner of a cubical plastered building. The two contiguous facades will always appear to be of different colours at every moment of the day, even if we 'know' they are the same colour. And so even the internal painted surfaces will undergo different local alterations based on their inclination in relation to lighting sources. It is therefore necessary to separate the part of chromatic registration from the part of scanning, in such a way to carry out specific photographs for similar parts of surfaces and with the same specific lighting. These then are 'mapped' on the cloud (or on mesh) in such a way to obtain a chromatically correct model (Carpiceci, 2011).

The phase of scanning was therefore carried out with the exclusion of the chromatic recovery from the scanner, which also has a very limited photographic section, certainly not comparable to a high level external camera.

Figure 7: Turkey, Cappadocia, Goreme Open Air Museum, Church of the Tokali, photo lighting set up (left) and scan the dark with the reflectance (right).

Despite the exclusion of RGB (Red, Green, Blue), the laser scanner registers the light intensity (reflectance) of the surfaces (Figure 7). This intensity is influenced both by the objective light of the surface where the measured point is found, and by the light from the reflection of the laser ray from the scanner itself. For the interior, therefore, this problem of variability was solved simply by carrying out the scans in the dark. In this way the registration was done with the actual reflectance of surfaces lighted exclusively by the laser of the instrument.

Some differences can be found in the cases in which parts of the surface had been registered with a grazing laser and others with a frontal laser, but these are particular cases that have not affected the validity of the solution found.

Representation

Classic architectonic representation is characterized by lines that describe essentially apparent outlines, corners and surface discontinuity (difference of elements, colours, etc.). The latter can be successfully substituted by photoplans, which nowadays are finally possible even for subjects, or parts of subjects, not plans.

Excavated architecture, however, with its more 'sculpted' than architectonic aspect, presents notable difficulties. In summary, not being able to retrace real edges, if not random and certainly not descriptive of the geometry of the volumes, using a form of 'diverse' representation is necessary (Carpiceci, 2013b).

For this reason we decided to borrow the 'diverse' technique from another applied field of 'two-dimensional' representation: territorial. Cartography bases its capability of representing the morphology of the land thanks to the use of isoipse (contour lines). Contour lines have two fundamental characteristics that make them significant to 'form', the first (etymological) is that lines composed of coplanar points, at the same height. The second characteristic is equidistance, or the constant difference between the planes: this constant determines the only variability that is represented by the interval, or the different dimension of every line of maximum slope traceable between a contour line and the limit. In cartography, equidistance is canonically 1/1000 of the scale of representation expressed in meters, for example: for the scale 1:25,000 the equidistance is 25 meters with lines marked every 100 meters.

In 'rocky' architectonic representation it's been noted that this relationship would involve an excessive concentration of lines. A good compromise was found by bringing the relationship to 1/500 and therefore for a classic representation 1:50 we have contour lines every 10 centimetres of difference and lines marked every 50 centimetres (figure 8).

The vertical sections have the 'flat' lines with the same spacing.

A great work of post-processing is also represented by the cleaning of the undercuts. In fact, the operator must keep track of the visible and hidden zones and therefore must

Figure 8: Turkey, Cappadocia, Sahinefendi, Church of the Forty Martyrs, plant of the entrance level and the church, contour lines at 10cm equidistant.

emphasize the apparent outlines of the forms and delete what is found behind them.

This 'manual' job could be avoided if a function were implemented that was capable of deleting these zones. Basically the sections are usually carried out by the mesh. Therefore the reconstruction of the surface of the numerical model establishes a separation between the interior and exterior of the material. This allows a common visualization with exclusion of the 'back faces'. It would suffice to make it so the invisible parts were cancelled out and then the determination of the sections were applied, to get the automatic elimination of the undercuts.

Next phase: inaccessible places

Rocky habitats certainly represent a subject very far from the concept of traditional architectonic survey. Such diversity is emphasized by the imitative aspect of traditional (constructed) architecture. An aspect that to a careful analysis demonstrates itself far from easily geometricised forms. Excavated architecture assumes sculptural characteristics that make it a unique typological and formal example. Another characteristic generalized in Cappadocia is inaccessibility. Beyond the 'museum' areas, which are fitted with stairs and walkways that allow easy access and pathways, the rest has spaces only reachable with specific equipment, such as that of spelunkers and climbers. And above all of it, there are then the peaks of the cones and rocky formations that would require a flying device such as a helicopter.

Recently we have seen the use of UAV (Unmanned Aerial Vehicles), commonly called drones, more and more. The use of these 4-propellered helicopters gives the possibility of flying over restricted areas of territory and taking photographs or videos useful to the successive processing of multi stereo matching, of photomodellation.

There are three fields that converge together in SAPR (Pilot Remote Aircraft Systems): radio-controlled aircraft, digital photography, and photomodellation. More and more sophisticated algorithms, combined with the increasing ability of calculation and visualization have made the technique of photomodellation more precise and reliable. Together with this, digital photography has now reached results and processing potential that make it nothing like what it was just five years ago, bringing quality and operational dynamism even on smaller and lighter camera-optic systems. In the end the refinement of the technologies of radio and remote control even from tablets have made navigation with drones easier and easier. Starting from the enormous productive 'entertainment' development, the evolution of the drone has spread to a wide variety of activities wanting the affordable and manageable operational ability of remote flight.

The UAV (Unmanned Aerial Vehicles) can therefore resolve most problems faced in sites with difficult access. In other words we can finally proceed with survey not only for the subjects for which accessibility involves uncomfortable and risky pathways, but also for the parts of common subjects that are not reachable without special equipment or particular 'means'.

Laser scanning, digital photography, photomodellation and radio-controlled flight, four determining factors that right now (some slightly earlier and some more recently) are literally revolutionizing the way of understanding architectonic survey.

This revolution is even more deeply felt in those places (or areas of places) that until recently were unknown or were 'victims' of 'hypothetical' reconstructions or required large sums of money in order to acquire reliable data.

The application of new technologies to inaccessible places represents a fundamental test to control the relationship and coordination of various techniques and their rightful place in the more complex 'Survey System', which is nothing more than the ordered set of material regarding a determined architectonic subject.

Acknowledgements

This paper is the result of the synergistic collaboration between the authors. In particular, Marco Carpiceci for You press the button, From Sahinefendi to Goreme and Representation, Carlo Inglese for Revolutionized Technologies, Acquisition Techniques and Next phase.

References

BIANCHINI C. (2012). *La documentazione dei teatri antichi del Mediterraneo*. Roma: Gangemi editore.

CARPICECI M. (2011). Survey problems and representation of architectural painted surface. *The International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences* XXXVIII-5/W16, pp. 523-528.

CARPICECI M. (2012a). Fotografia digitale e architettura. Storia, strumenti ed elaborazioni con le odierne attrezzature fotografiche e informatiche. Roma.

CARPICECI M. (2012b). Modelli geometrici e costruzioni grafiche per il rilevamento architettonico. Roma.

CARPICECI M. (2013a). Siamo solo agli inizi del rilevamento digitale: alcune considerazioni sullo sviluppo delle attuali tecnologie. *Quaestio* XV, 27, pp. 53-64.

CARPICECI M. (2013b). Cappadocia Laboratorio-Rilievo (2007-2015). In *XXXV Convegno Internazionale dei Docenti della Rappresentazione: Patrimoni e Siti Unesco, memoria, misura e armonia, 24-26 ottobre 2013 Matera*, Gangemi Editore, pp.221-229.

CARPICECI M., INGLESE C. (2014). Laser Scanning and Automated Photogrammetry for the knowledge and the representation of the architecture cave in Cappadocia: Sahinefendi and the Open Air Museum in Goreme. Communication to the CAA2014, Paris, Panthéon-Sorbonne University, 22-25 April 2014.

CARPICECI M., COLONNESE F., INGLESE C., CRESCIANI G., (2013). Angelini Andrea. Dalla roccia alla città. Il rilievo del villaggio rupestre di Sahinefendi, In *XXXVI Convegno Internazionale dei Docenti della Rappresentazione: 'Italian Survey & International experience', 18-20 settembre 2014 Parma*, Roma: Gangemi editore, pp.603-610.

CORMACK R. (1967). Byzantine Cappadocia: the archaic group of wall-paintings. *Journal of the British Archaeological Association*, 1967, Ser.3, 30, pp. 19-36.

DAVIS-WEYER, C. (1998). Pittura. Secoli 6°-10°. In *Enciclopedia dell'arte medievale*. Roma: Istituto della enciclopedia italiana, pp. 447-466.

GIOVANNINI L. (Ed.), (1971). *Arts of Cappadocia*. London: Barrie and Jenkins.

JERPHANION G. de (1925). Une nouvelle province de l'art byzantin. Les églises rupestres de Cappadoce. Paris: Paul Geuthner.

JOLIVET-LEVY C. (2001). L'arte della Cappadocia. Milano, pp. 41-47.

KALAS V. (2006). The 2004 Survey of the Byzantine Settlement at Selime-Yaprakhisar in the Peristrema Valley. *Dumbarton Oaks Papers* LX, pp. 271-293.

KOSTOF S. (1972). Caves of God: *The Monastic Environment of Byzantine Cappadocia*. Cambridge, MIT Press, pp. 51-64.

MATHEWS T. F., DASKALAKIS-MATHEWS, A.-C. (1997). Islamic-Style Mansions in Byzantine Cappadocia and the Development of the Inverted T-Plan. *Journal of the Society of Architectural Historians* , pp. 294-315.

MIGLIARI R. (2003). *Geometria dei modelli: rappresentazione grafica e informatica per l'architettura e per il design*. Roma: Kappa.

NICOLETTI M. (1980). *L'architettura delle caverne.* Roma: Laterza.

OUSTERHOUT R. 1995. Historical Design in the Environment: An Examination of a Byzantine Settlement in Cappadocia. In MACKENZIE Dorothy, Design for the Environment: The Interdisciplinary Challenge. Urbana-Champaign, pp. 13-19.

OUSTERHOUT R., G. (2005). A Byzantine settlement in Cappadocia. Washington, D.C.

RESTLE M. (1968). *Byzantine Wall Painting in Asia Minor.* trans. I. R. Gibbons, 3 vols. Greenwich, Conn: New York Graphic Society.

RODLEY L. (1985). *Cave monasteries of Byzantine Cappadocia.* Cambridge: Cambridge university Press.

THIERRY N. (1963). *Nouvelles églises rupestres de Cappadoce. Region du Hassan Dagi*. Paris: Klincksieck.

THIERRY N. (1975). L'art monumental byzantin en Asie Mineure du XIe siècle au XIVe. *Dumbarton Oaks Papers* 29, pp. 75-111.

THIERRY N. (1983). *Haut moyen-âge en Cappadoce: les églises de la région de Cavusin*. Paris: Paul Geuthner.

VERZONE P. (1962). Gli [sic] monasteri de Acik Serai in Cappadocia. *Cahiers Archéologiques* 13, pp. 119-136.

WHARTON A. J., (1988). *Art of Empire: Painting and Architecture of the Byzantine Periphery.* University Park Pennsylvania State Univ. Press, pp. 13-18.

Advantages and Disadvantages of Digital Approach in Archaeological Fieldwork

Carlo BIANCHINI, Francesco BORGOGNI, Alfonso IPPOLITO

Dipartimento Storia, Disegno e Restauro dell'Architettura (Sapienza - Università di Roma), Italy
carlo.bianchini@uniroma1.it, francesco_borgogni@yahoo.it, alfonso.ippolito@uniroma1.it

Abstract

Graphics and photographs have always played a fundamental role in archaeological documentation and data interpretation. Both may be affected by subjective interpretation. In the last decades the digital approach has helped to speed up fieldwork and provide enough information for GIS platforms.

Since the 1990ies researches have applied digital numeric models to the representation of phenomena also in disciplines like archaeology. Nowadays this approach has become common practice. Nevertheless exploiting increasing amounts of data often implies more complex procedures for investigation and categorization.

The whole process depends on the quality of models adopted to gather, order out and process this information. It is aided by 3D numeric models that are becoming widely available for automatic 3D acquisition, digitalization and real time browsing. Digital gathering of fieldwork should consider the need of collecting formation in the shortest time and budgetary constraint shortest time and budgetary constraints.

budgetary con-straints. Digital gathering of fieldwork should consider the need of collecting maximum information in the shortest time and budgetary constraints.

Keywords: Survey, Archaeological Fieldwork, Documentation, Low Cost Technologies.

1. Introduction

The cognitive process in archaeology aims at reaching the most precise and complete understanding of the object of research. The methodologies that exploit the potentialities of digital instruments for non-contact surveying have lately increased their relevance actually becoming part and parcel of an integrated workflow somehow combining tradition and high-tech. The study of an archaeological site or element often engages also different expertise and professionals (archaeologists, architects, historians, engineers, ICT experts, etc.): this fact entails the setup of a common working protocol already at the excavation stage for both data capturing and modeling. This goal can be achieved through a double effort: on the one hand, experts in digital technologies and techniques will have to better respond to the needs connected with the archaeological research; on the other, instead, scholars will have to become better acquainted with capturing/modeling technology in order to master its potentials. In this paper we shall describe the procedure worked out by our research group taking into account both the main techniques used for 3D capturing and the methods for their optimization when applied to archaeology. In particular we shall describe and compare models obtained with active (laser scanner), passive (photomodeling) 3D no contact capturing systems and traditional direct surveying methods.

The survey and the methodological comparison presented in the following paragraphs refer to an excavation campaign developed in 2011 in the necropolis of Crustumerium, an ancient city situated a few kilometers north of Rome. The work has entailed the documentation of the various excavation stages of one of the tombs[1] as well as of some findings.

2. The necropolis at crustumerium

The majority of ancient Roman authors describe Crustumerium as a Latin city. It occupied an area of approximately sixty hectares along the present via Marcigliana, north of Rome. The site was strategically very important being on the main commercial road which connected Etruria with Campania: this particular location made the city flourish in archaic times and, most probably, it also represented one of the reason why the neighboring cities of Veio and Rome were so much interested in maintaining relations with Crustumerium. Rome expanded and in 500-499 B.C. finally conquered Crustumerium which kept its strategic its strategic relevance until 369 B.C. when, after the destruction of Veio, the Romans were no longer interested in controlling roads in the area. Thus the ancient city gradually fell into the state of semi-abandonment and disappeared completely during the V

[1] Experiments with large objects have been purposefully excluded because of the sheer bulk of information would make it difficult to carry out elaborations.

Figure 1: Archaeological site at Crustumerium. General view and photograph of the entrance to the reserved area of the site at Crustumerium

Figure 2: Typologies of Etruscan tombs. Tomb 'a calatoia': Photo of a tomb with single lateral recesses at the site of Crustumerium.

century A.D. The ancient site of Crustumerium was then discovered in 1976 thanks to the research of the CBR conducted by the archaeologists Lorenzo and Stefania Quilici who produced the first survey documentation concerning this ancient settlement (Fig. 1).

The site is so vast that, at least so far, an exhaustive study of the internal structure of the city is still lacking, being its necropolis the only well analyzed part. The finds collected to date have allowed the researchers to differentiate and define various funeral typologies related to different historical periods.

Our research was focused on one of the tombs excavated in July 2011 (n. 310[2]), a rectangular well of 106x231 cm. approximately 84 cm. deep, with a lateral loculus (80x260 cm., 63 cm. high) closed with tufo tiles placed vertically in which the archaeologists discovered the skeleton of a woman with a complete funeral dowry.

3. Problems inherent in data acquisition

Excavations inside a necropolis imply the re-opening of structures which had been conceived and built to remain closed forever and that, in addition, are presently located under layers of materials deposited through the centuries. It is paramount then that any archaeological investigation must be regarded as a highly invasive, destructive and above all irreversible process and that documentation is the only means to keep track of the original status of the site. Furthermore, the necropolis of Crustumerium contains various types of funeral structures differing in shape and size. The common features are big stone slabs – horizontal and vertical – designed with the function of sealing the corpse inside his or her own burial recess (fig. 2). The documentation of these sealing slabs (which inevitably need to be removed in order to bring to light the dowry and the skeleton) represents one of the main problems to be addressed. Too often, in fact, they have been too hastily documented or simply destroyed without considering the relevance of these burial structures. Nowadays the stratigraphic approach allows the accurate documentation of each excavation phase and of all structures belonging to the different layers encountered by the researchers. In this framework, the burial site has been specifically studied to test out different fieldwork documentation methodologies (both traditional and high-tech) in order to assess advantages and disadvantages in terms of accuracy, cost, time, etc. Even if the traditional approach (direct survey, hand sketching, etc.) still appears somehow irreplaceable, nevertheless we have concentrated especially on 3D capturing techniques comparing procedures and results coming from 3D scanning and photomodeling.

One particular parameter has been closely monitored during our work developed together with the team of archaeologists: the user-friendliness of the system. Any technical development, in fact, will turn to be effective only if the main user group is likely to adopt it: in our case it means that 3D capturing/modeling will deploy its real potential in the archaeological field only if archaeologists will feel confident enough of using it widely. Not only the cost of the equipment can be seen as a great obstacle but also a general suspiciousness towards technology and a lack of confidence in handling digital instruments both

[2] The number refers to the cataloguing useful for identifying the tomb within the archaeological site.

Figure 3: Chronology of excavation.
From arrangement of the tomb area to the discovery of archaeological finds.

prevent the 3d capturing /modelling to become widely a accepted method.

Still too often, in fact, researchers regard the traditional direct surveying as the best way to obtain controlled and reliable results in a reasonable time at a sensible cost. Quite apart from the technological issues related to the comparison of 3D capturing techniques, our work aimed also at demonstrating how in a real context (the excavation of tomb 310 in Crustumerium) the application of semi-automatic 3D capturing procedures could better respond to the needs of the archaeological team.

In this framework, the most significant stages of the excavation were documented using at the same time a TOF laser scanner (Leica ScanStation2),[3] a 12 Megapixels compact camera (Nikon s3000) and the typical methodologies of graphic documentation used by archaeologists. The excavation work entailed in total ten different stages corresponding to an equal number of layers: we have documented each of them and finally reconstructed a stratigraphic model of the digging sequence from the ground to the tomb level. All structures and findings have thus found their correct position in a homogeneous 3D space minimizing the loss of information due to their physical removal made by archaeologists. The survey was not only focused on the 'geometry' of the structures and findings: but it has also been able to document the superficial state of the different surfaces by texturizing the 3D model with the images acquired by the scanner or captured by the camera (fig. 3).

4. Data acquisition

4.1 Direct surveying

As already mentioned in the previous paragraphs, the integrated approach used in Crustumerium relies on a close cooperation between archaeologists and experts in digital capturing/modeling technologies. Besides, we all intended to compare data and procedures commonly used by different teams in order to assess benefits and drawbacks. For this reason, different teams worked on the same site each acquiring information according to their 'standard' protocol. In this framework, archaeologists have collected data using their traditional procedures: observation of the object, understanding of the context in which it is immersed, direct measure of a limited number of significant points. The information has then been stored in the shape of two dimensional drawings (plans and sections) and photographs. Both these, together with other studies and analyses (historical, sociological, etc.), would lead to hypothesis about the traditions of the ancient civilization of the inhabitants of Crustumerium as well as to place the archaeological finds in their proper historical and geographical context. It's paramount that in this case there is a direct relationship between the level of acquaintance of the surveyor towards the site and the quality of the survey itself as any bit of information reported on a piece of paper depends inevitably and directly on the skill, experience and will to communicate specific aspects of the object. In other words the information collected are the result of a direct choice of the surveyor. Besides, two more crucial aspects must be considered about traditional surveying methodologies: firstly, it is not always possible to use direct surveying procedures without altering or damaging the find or the context; secondly, as we shall demonstrate in the following lines, the direct approach not always represents the best combination of time, cost and information for documenting the various steps of an excavation campaign (fig. 4).

4.2 Non contact surveying

Long Range Laser Scanner

Long Range Laser Scanner In surveying tombs buried underground one runs the risk of encountering problems while trying to apply the procedures – like those of the ground undercut or the positioning of the target which lies at a certain depth and hasrestricted dimensions. These have to be solved in order to proceed with the step of registering between separate scans. As far as the undercuts are concerned, if proper considerations are not taken into account before initiating the survey, one runs the risk of not obtaining complete data or that of getting a plethora of information. As to the problem of targets, the choice was made to avoid applying them to directly on surveyed surfaces for two fundamental reasons: one, in order not to alter the surveying data and simulate as closely as possible

[3] Instruments which can acquire data at a minimum sample spacing of 2x2 mm, capable of surveying objects at the distance od 300 m. Sample spacing is the distance between two consecutive measurements in relation to two principal axes of reference, x and y.

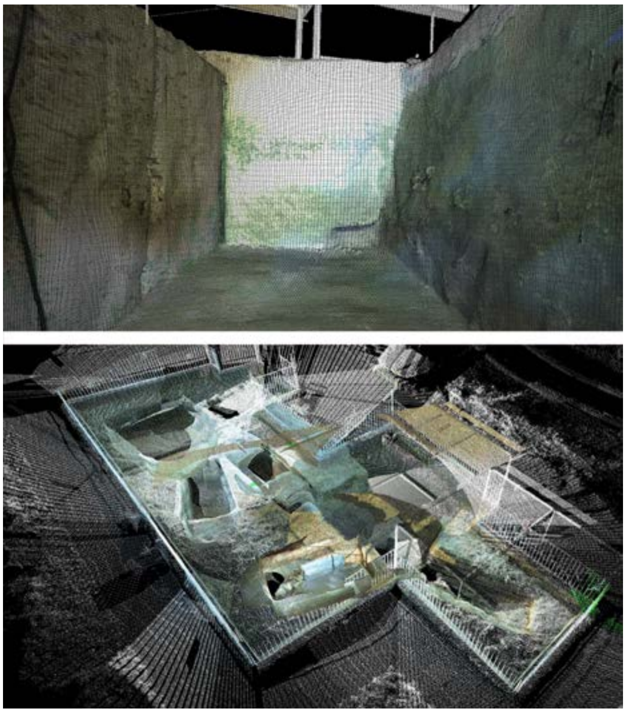


Figure 4: Surveying with 3D laser scanner Scanstation2. Surveying data acquisition, point cloud. To eliminate a series of undercuts in was necessary to immerse the scanner into the sepulchers. This position impeded the aiming of the targets necessary for the successive stage of data registration.
Montage of scans, point cloud: total point cloud with combined single scans.

Figure 5: Example of a photo set. The surveying project envisaged taking 12 photo sets for each excavation stage. The necessity of overlapping two or three photograms was taken into account.

the conditions of the fieldwork; two, the necessity to excavate would not allow to keep the objects in their fixed places inside the tomb for the whole time of carrying out the excavations. These preliminary considerations guided an optimized positioning of targets during the surveying of all excavation stages of tomb n. 310. In order to proceed as expeditiously as possible with the digging, four targets were fixed on four wooden poles, driven into the ground around the area to be studied and positioned in such a way that they were visible from all places without disturbing the archaeologists' work. What is more, in order to scan the inside of the tomb, independently from its depth, a tubular scaffolding – easy to dismantle - was erected and used to fix four more targets above the excavation site.

Ten different excavation stages have been documented. For each of them two separate scans were performed in such a way as to fill out various undercuts with a 2x2 mm. spacing. As long as the excavations was not too deep, the scanner was placed on its tripod. When the excavation got deeper, it was placed directly on the ground around the edges of the hole. In this way the lower shadow zone was limited[4] and the scanning provided more data. The Cartesian coordinates and respective RGB and reflectivity data characterize each point of every acquired cloud.[5] Thanks to the overlapping of clouds it was possible to unequivocally represent the real development and succession of the ten excavation layers adding somehow a fourth-dimensional character (time) to the conventional two and three-dimensional representations. The stage of registering point clouds follows the acquisition phase: it aligns the local coordinate systems of the different scans to an uni-vocal and homogeneous Cartesian system of reference generally using tie points. This operation creates the numeric model, that is to say the general point cloud composed by all the aligned 'sub-clouds'. Leica Cyclone 6.0, the software used at this stage, ensured also a real-time control of the level of uncertainty of the procedure thanks to semi-automatic tools for registering clouds. In this framework, we chose to consider the average error of collimation for targets (about 2 mm.) as a reference for the general uncertainty value of the data acquired with the laser scanner and, more generally, the benchmark value also for comparing different methodologies. All the analyzed stages of the excavation have been thus captured as independent scans leading to independent numeric models. For this reason, after optimizing the clouds with Meshlab 1.3.0 software (i.e. eliminating redundancies), they have been put together in a single 3D space actually reconstructing the exact timeline that had generated the capturing. 4Things that are no visible from a definite positioning of the instrument cannot be surveyed. In order to have complete data, it is necessary to carry our a further series of scans that will

[4] Things that are no visible from a definite positioning of the instrument cannot be surveyed. In order to have complete data, it is necessary to carry our a further series of scans that will include places hidden from sight and will simultaneously guarantee the overlapping satisfactory for assembling the parts. The lower shadow zone is the area 'unsurveyable' for technical mechanical reasons and extends at the angle of 90 degrees.

[5] The value of reflectivity indicates the amount of light that a given surface can reflect.

include places hidden from sight and will simultaneously guarantee the overlapping satisfactory for assembling the parts. The lower shadow zone is the area 'unsurveyable' for technical mechanical reasons and extends at the angle of 90 degrees.

Exploiting Poisson's interpolation algorithm[6] for extracting mesh surfaces from a point cloud, it was possible to build polygonal three dimensional models made up of triangular surfaces which had to be partially corrected manually to eliminate again redundancies and to keep only the portion of the model significant for our purposes. Chromatic surface characterization is yet another datum necessary to compare the 'scanned' 3D model with the image-based one. For this specific reason we used a special function developed within MeshLab that transfers the RGB value from the point cloud directly to the vertexes of the mesh surface and then to the texturized model.

The application of this particular technology to the surveying of tomb 310 pointed out some specific issues:

– natural illumination of the site with shadows changing through the different phases of the day; the context which, though cooperative towards our experimentation, could not be too modified to accomplish our needs;
– the surveyed object, which could not be moved or turned because of its particular structure. Finally, the use of low cost instruments for non-contact surveying (compact commercial camera with 12 Megapixels) called for an optimization of the survey project in terms of number and framing of photographs to be taken. On the other hand, as our activity aimed at comparing various different surveying methodologies, it has been necessary to provide a photographic documentation of the same excavation stages already documented through laser scanning.

A preliminary recognition of the actual equipment potentials of our teamin hardware/software terms, we decided to take a set of pictures that would not exceed the total of 30 units, a number that would anyway guarantee a satisfactory overlapping of images without overloading the calculation system.

Twelve photographs taken around the excavation site together with 18 close up shots formed the whole dataset yielding to a perfect balance between the general pictures and the detailed ones. Unfortunately, because of the uneven geometry of the structure and some operational limits of the laser scanner, certain well-hidden parts of the object could not be surveyed at all. In this case the point cloud has been obtained only by correlating homologous points on two photograms.[7] After cleaning the datum with Poisson's interpolation algorithm to get rid of redundant information, it was possible to build a three dimensional polygonal model. The result of this process underwent an optimization leading to the transfer of chromatic values of the initial point cloud to the vertexes of the mesh surfaces and finally to the 3D texturized model.

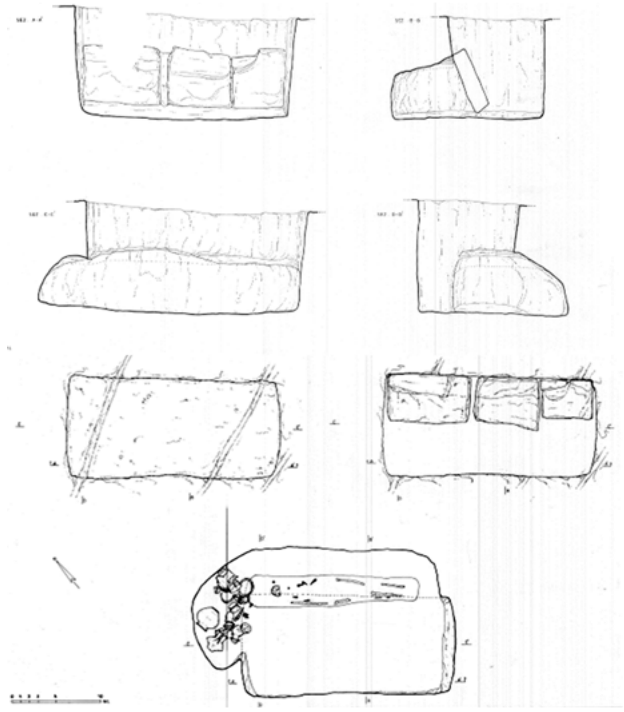


Figure 6: Direct surveying.
Two dimensional elaborations: plan and sections of various excavation stages

Comparison

The first evaluation concerned a methodological comparison of non-contact surveying methodologies.

When working on any excavation site, especially such as that of Crustumerium gone through a series of clandestine excavations, it is very important to extract the various finds as fast as possible. Keeping an excavation site open too long also means risking all kinds of problems connected with weather conditions. While conducting our survey, in fact, an unexpected night rain filled all the tombs with water seriously compromising the fieldwork. Not only did the unfortunate accident affect negatively the rhythm and the pace of our work (the whole area had to be dried) but it also threatened the state of conservation of the finds, which had already been partly uncovered. Such events, rare as they are, bring to light the need of procedures that guarantee the fastest collection and documentation of the finds without reducing excavation standards. In our experience, time proves to be one of the main parameters that distinguishes the two approaches analyzed: data

[6] For further reading on this particular algorhythm, see texts in the bibliography.

[7] The automatic correlation of homologous points is the stage in which the photomodelling process is articulated. It follows the acquisition of photo images precedes the construction of the numeric model. At this stage homologous points in the photograms are transformed automatically into points in space characterized by Cartesian coordinates XYZ and the chromatic ones – RGB. Such an operation is necessary in order to find the position in the very centre of the projection or of the camera in relation to the scene at the moment of acquiring the image. It determines the precision and number of points gathers later.

acquisition with the laser scanner, though extremely fast,[8] imposes nevertheless some operational constraints related with the positioning of the instrument that can be quite time consuming. Moving the scanner around the tomb has been in fact not so easy: the scanner itself weighs approximately 18,5 kg and any shift generally involves also its tripod, the batteries and the laptop that controls the instrument.[9] On the contrary, the procedure based on the photographic data acquisition is much easier: the only instrument necessary is in fact a camera (much more convenient to carry around), no particular positioning is required before each picture, the overall time necessary to complete the work is that needed to target and shot each photograph. Here is a comparison of the two methodologies in terms of the time used for the acquisition phase: TOF laser scanner Number of scans: 2 Average time of a single scan: approx. 45 mins. (taking into account the necessity to move the instrument around as well as targets' acquisition) Average total time necessary for the acquisition of a single excavation stage: 90 mins. Photomodeling Number of photograms: 30 Average time per single photogram: app. 20 sec. Average total time for a photographic set: 10 mins.

The differences between the two approaches are quite evident: by saving up about 80 minutes for each stage of the excavation means that for the 10 stages the reduction would be of about 13 hours.[10] It must be stressed, however, that the point cloud acquired by the 3D laser scanner can be navigated and controlled directly on-the-go, while this important feature is precluded when gathering data from photographs. In this second case, though, it is possible to take in only 10 minutes the 30 photograms necessary to guarantee a complete modeling and a point cloud comparable to that produced with the laser scanner. Another parameter of the comparison of the two non-contact surveying methodologies is the cost of the surveying campaign. From this standpoint, photomodelling proves to be much cheaper than 3D scanning especially for the cost of the equipment itself and its deploying on site. The comparison of the two methodologies has unequivocally demonstrated that photomodelling, at least in the capturing phase, is much more advantageous than the TOF laser scanner.

5. Data elaboration

The assessment of the level of uncertainty of the captured data represents one of the most significant parameters related with survey. When using the technique of photomodelling, though, it is impossible to assess this datum during the capturing phase. Any comparison between the data obtained with this methodology and those acquired by laser scanning requiressomepreliminary operations so to make the two clouds compatible: in particular, the image-based cloud needs to be conveniently scaled as the raw numeric model is dimensionless. Following this crucial step, we started to comparesystematically the 3D models of each excavation stage (laser scanning and image-based) analyzing homologous elements. This operation turned to be particularly tricky: the acknowledgement of homologous elements in both models (i.e. a line or an edge) implies the tracing of a number of corresponding points in the twoclouds, often very different in shape and color.

Once the image-based models were properly scaled, there followed an alignment phase performed by means of rigid roto-translations which helped to find the best fitting positionbetween the scanner and the image-based models. The procedure was carried out with the Geomagic Studio 10 software: after a manual pre-alignment, we used an automatic procedure to reach the best positioning always controlling the average error.[11] Justbefore performing the models comparison, theyhave been optimized by cleaning out edges, that is to sayregularizing the contours of homologous surfaces. Finally, we went through the assessment of deviations: point by point we pointed out their distancein the two different models (the minimum distancesegment) assuming the 3D scanner model as a reference. The following tables summarize the models comparison referring to the initial, intermediate and the final excavation stages of tomb n. 310.

6. Conclusions

Analysis of the values obtained by comparing 3D models built with the data from the 3D TOF laser scanner and from the photomodelling brings forward a number of particularly interesting issues. Some characteristics intrinsic to each technology fundamentally influence the final result in terms of chromatic output and uncertainty level of the model obtained. Texture mapping for both models – The model obtained from the scanner and the image-based one, are both carried out by transferring RGB values of the initial cloud directly onto the vertexes of the final mesh surface. In principle, a very dense point cloud yields a mesh surface rich in geometry(high number of vertexes and edges) and thus an excellent chromatic result and ahigh level of detail in the shift from the cloud to the surface. The test carried out on real conditions (survey of tomb n.310) has revealed quite a different situation. The scanner used for documenting the excavation did collect a high number of points which could be used to construct a particularly detailed 3D model, but the acquired color datum was almost useless forthe following main reasons:

- the resolution of the 3D scanner camera (1024x1024 pixel) did not guarantee a convenient level of detail;
- the way the scanner acquires images: the internal mirror drives in fact light onto the CCD sensor often generating many undesired reflection effects;

[8] ScanStation2 has the maximum acquisition velocity of 50 000 points per second.

[9] In the new models the problem of the weight of the instrument has been at least partially solved and the user can control the scanner directly on the instrument thanks to an integrated display.

[10] Just as reference, in the time the laser scanner worked on the tomb n. 310, 4 other tombs on the excavation site and a funeral memorial stone discovered in the area were successfully surveyed.

[11] The value of the average error is in inverse proportion to the precision level of the alignment operation

	Scanner laser 3D data	Data from photographs
Number of scans	1	-
Number of photograms	-	10
Exported dimension file point cloud	55.222 Kb	25.226 Kb
Initial numer of points	1.570.721	371.444
Number of points after cleaning excessive data (.ply)	947.928	208.874
Mesh model obtained through Poisson's interpolation	659.374 faces	310.741 faces
Final dimension model (.ply)	16.488 Kb	7.962 Kb

Excavation stage 1

Scale factor calculation = distance of model from scanner (m) / distance from model da photographs (m)	Average value in the data obtained: **1,57**
Data registration	Manual pre-alignment by signalling of four homologous points Automatic alignment with error calculation Average error = **0,004 m**
Deviation calculation	Maximum Distance positive: +0,021 m Maximum Distance negative: - 0,017 m Average Distance: 0,000 m Average Distance positive: +0,004 m Average Distance negative: - 0,004 m Standard Deviation: **0,005 m**

Table 1: Excavation stages 1. Comparison between scanner laser 3D and surveying from photographs, data acquisition and data elaboration.

- the time needed for scanning produces a set of images with different exposition and shadows and finally a chromatically non-homogeneous point cloud.

3D models obtained from photomodelling, instead,although less dense in terms ofvertexes and edges, are chromatically uniform enough to enrich with information a geometry which is less detailed.[12]

Another issue refers to the geometric detail level of the 3D model. While the scanner guarantees a high geometric accuracyforacquired data being not influenced by the brightness or glare of the object, photomodelling is instead deeply affected by the 'visible' aspect of the object that needs to be sufficiently well lighted to be adequately photographed. From this specific standpointthe laser scanner yields more homogeneous data thanphotomodelling where, on the contrary, there is a strong lack of homogeneity caused by differences in lighting.[13]

Other aspects connected with any surveying campaign are cost and time. There is no doubt that the photomodelling technology is zero cost in comparison with the price of a 3D laser scanner but also in comparison with the instruments used in photometry and topography. The budget would most likely be comparable instead to that of a direct surveying campaign, practically affordable by anyone. Moreover, these image-based technologyallow for surveys 10-20 times faster than those carried out with a 3D laser scanner. Worth considering is also the transport and positioning of the instruments where cameras operativityshow no constraints while 3D laser scanners might seriously be limited.[14] We cannot affirm thatphotomodelling can always substitute the use of 3D scanners: in our specific case, though, where archaeologist had toget their documentation as soon as possible in order to describe exhaustively the state of the excavation, it has proved very productive and accurate. Even if an uncertainty value of ±3-4 cm. was considered acceptable by archaeologists, the image-based technique yielded models with uncertainty level of ±0,5-0,4 cm. and with a great save of time. In our opinion, finally, we think thatthis 'zero-cost capturing technology'will deeply revolutionize the archaeological documentation in the very close future. It proved to be a valid alternative to the graphic documentation on the spot, a powerful instrument for 3D capturing as well as scanners, an indispensable equipment of all future researchers.

[12] In some 3D models based on photographs chromatic values were changed because of different exposition to light of the object surveyed. This is the case of a part of the loculus and the parts that were in shadow.

[13] Some of the zones particularly dark like e.g. the bottom of the recess and the so-called 'sky' have a very low detail level in comparison to the rest of the tomb.

[14] These considerations are connected to the typologies of the 3D laser scanner used in the experiment described. They cannot be, therefore, regarded as valid for all laser scanners available on the market.

	Scanner laser 3D data	**Data from photographs**
Number of scans	1	-
Number of photograms	-	10
Exported dimension file point cloud	55.222 Kb	25.226 Kb
Initial numer of points	1.570.721	371.444
Number of points after cleaning excessive data (.ply)	947.928	208.874
Mesh model obtained through Poisson's interpolation	659.374 faces	310.741 faces
Final dimension model (.ply)	16.488 Kb	7.962 Kb

Excavation stage 5

Scale factor calculation = distance of model from scanner (m) / distance of model from photographs (m)	Average value in the data obtained : **1,42**
Data registration	Manual pre-alignment by signalling four homolgous points Automatic alignment with error calculation Average error = **0,008 m**
Deviation calculation	Maximum Distance positive: +0,029 m Maximum Distance negative: - 0,052 m Average Distance: 0,003 m Average Distance positive: +0,006 m Average Distance negative: - 0,006 m Standard Deviation: **0,006 m**

Table 2: Excavation stages 5. Comparison between scanner laser 3D and surveying from photographs, data acquisition and data elaboration.

Number of scans	1	-
Number of photograms	-	10
Exported dimension file point cloud	55.222 Kb	25.226 Kb
Initial numer of points	1.570.721	371.444
Number of points after cleaning excessive data (.ply)	947.928	208.874
Mesh model obtained through Poisson's interpolation	659.374 faces	310.741 faces
Final dimension model (.ply)	16.488 Kb	7.962 Kb

Excavation stage 10

Scale factor calculation = distance of model from scanner (m) / distance of model from photographs (m)	Average value obtained: **0,50**
Data registration	Manual pre-alignment by signalling four homologous points Automatic alignment with error calculation Average error = **0,002 m**
Calcolo delle deviazioni Deviation calculation	Maximum Distance positive: +0,015 m Maximum Distance negative: - 0,018 m Average Distance: - 0,000 m Average Distance positive: +0,002 m Average Distance negative: - 0,002 m Standard Deviation: **0,003**

Table 3: Excavation stages 10. Comparison between scanner laser 3D and surveying from photographs, data acquisition and data elaboration.

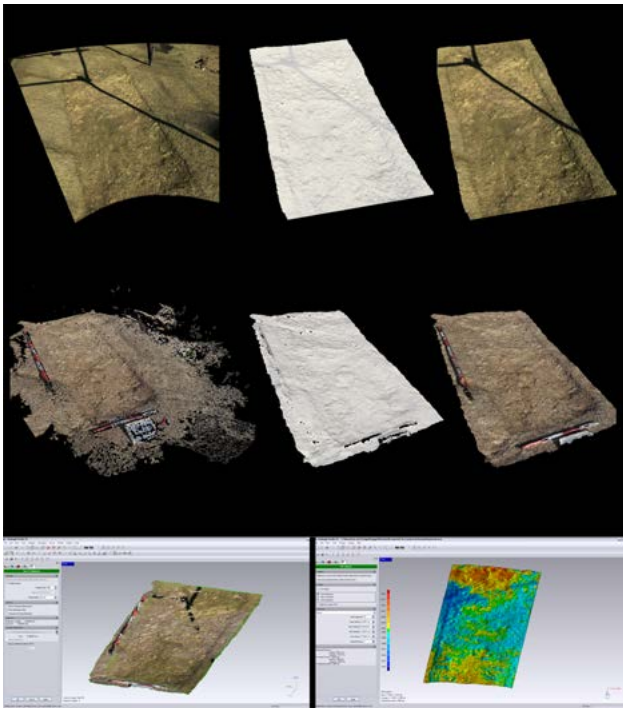


Figure 7: Excavation stages 1.
Point cloud from 3D laser scanner: original point cloud.
Mesh model from 3D laser scanner: mesh model geometry and mesh model with photographic characterization. Point cloud from photomodelling: initial point cloud.
Mesh model from photomodelling: mesh model geometry and mesh model with photographic characterization.
Model alignment: Automatic operation of alignment and average error calculation.
Deviation calculation: by calculating deviations it is possible to quantify the gap between two models. The image presents chromatic values referring to the so called deviation scale. The maximum values are in red and intensive blue.

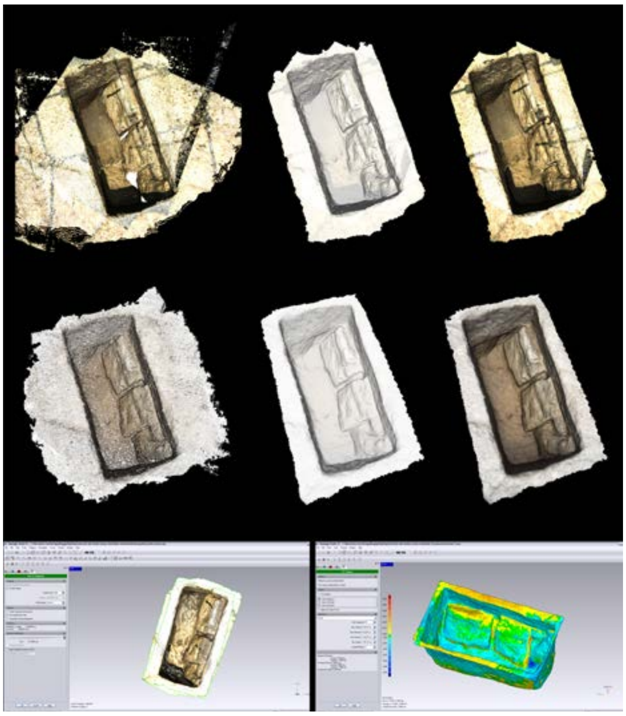


Figure 8: Excavation stage 5.
Point cloud from 3D laser scanner: initial point cloud.
Mesh model from 3D laser scanner: mesh model geometry and mesh model with photographic characterization. Point cloud from photomodelling: initial point cloud.
Mesh model from 3D laser scanner: mesh model geometry and mesh model with photographic characterization.
Point cloud from photomodelling: initial point cloud.
Mesh model from photomodelling: mesh model geometry and mesh model with photographic characterization,
Model alignment: Automatic alignment operation and average error calculation.
Deviation calculation: by calculating deviations it is possible to quantify the gap between two models. The image presents chromatic values referring to the so called deviation scale. The maximum values are in red and intensive blue.

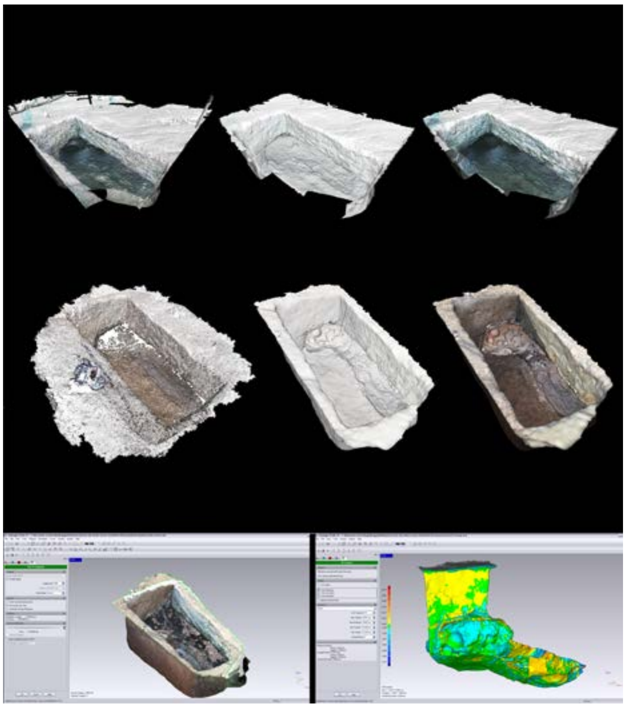


Figure 9: Excavation stage 10.
Point cloud from 3D laser scanner: initial point cloud.
Mesh model from 3D laser scanner: mesh model geometry and mesh model with photographic characterization. Point cloud from photomodelling: initial point cloud.
Mesh model from 3D laser scanner: mesh model geometry and mesh model with photographic characterization.
Point cloud from photomodelling: initial point cloud.
Mesh model from photomodelling: mesh model geometry and mesh model with photographic characterization,
Model alignment: Automatic alignment operation and average error calculation.
Deviation calculation: by calculating deviations it is possible to quantify the gap between two models. The image presents chromatic values referring to the so called deviation scale. The maximum values are in red and intensive blue.

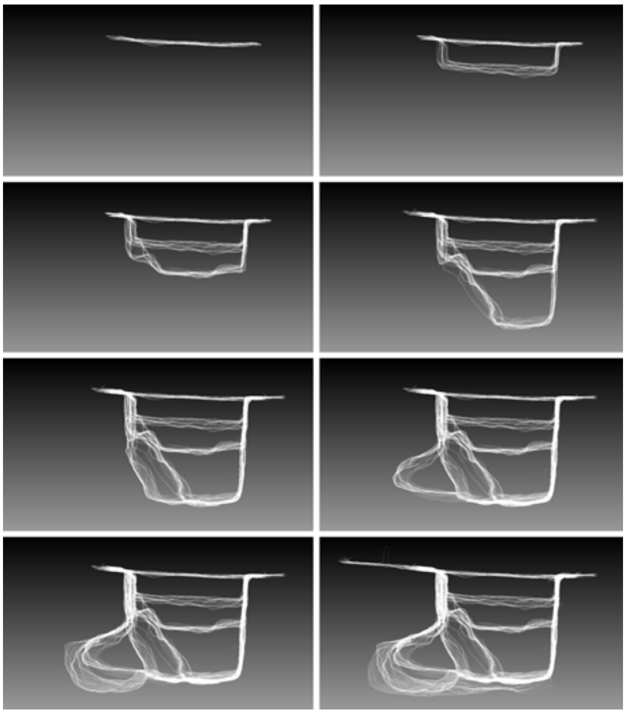


Figure 10: Excavation stratigraphy reconstruction. Placing the various stages of the excavation into a unified Cartesian system of reference makes it possible to reconstruct in 3D the stratigraphy of the documented excavation.

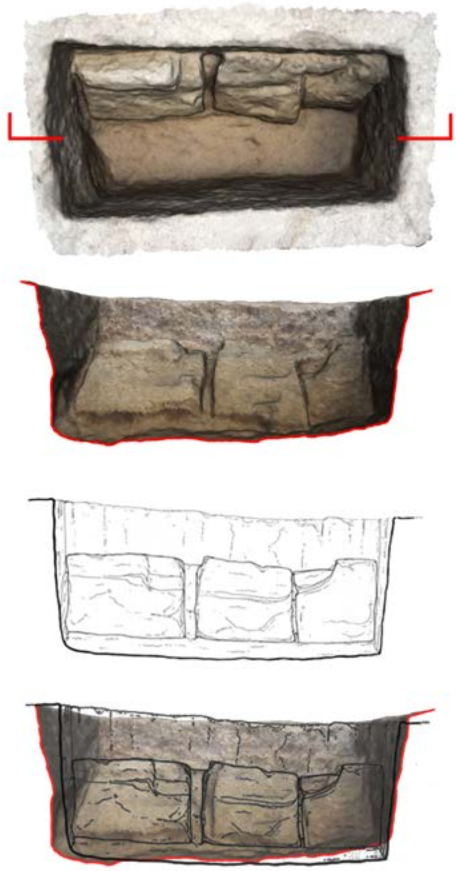


Figure 11: Base elaboration mining. By intersecting models with reference plans it is possible to obtain also two dimensional representations and compare them with elaboration worked out by direct surveying.

References

BARKER, P. (2005). *Techniques of archaeological excavation.* London: B. T. Batsford.

BIANCHINI, C. (2001). Modelli discreti e modelli continui nel rilievo e rappresentazione informatizzata dell'architettura. In: *Disegnare. Idee, immagini,* 22, pp.51-60.

BIANCHINI, C. (2005). Laser Scanning X. In: *Metodologie innovative integrate per il rilevamento dell'architettura e dell'ambiente, ricerca Cofin 2004. A cura del Dip. di Rilievo, Analisi e Disegno* dell'Ambiente e dell'Architettura - Università di Roma La Sapienza. Roma: Gangemi Editore, pp. 24-31.

BIANCHINI C. (2012). *La Documentazione dei Teatri Antichi del Mediterraneo sul Attività del Progetto Athena Mérida / Documentazione di Antichi Teatri del Mediterraneo: le attività di Athena a Mérida.* Gangemi Editore, Roma.

BIANCHINI C., BORGOGNI F., IPPOLITO A., SENATORE L.J. (2013). The surveying and representation process applied to archaeology: a quest for invariants in a highly variable context. In Computational Modeling of Objects Presented in Images, *Lecture Notes in Computational Vision and Biomechanics* 15, pp. 1-29.

BORGOGNI, F. (2011). Rilievo integrato alla grande scala: il caso di Mérida. In Chiavoni E., Filippa, M., Docci M. (a cura di). *Metodologie integrate per il rilievo, il disegno, la modellazione dell'architettura e della città, Prin 2007. Roma: Gangemi Editore, pp. 41-46.*

CALLIERI, M., DELL'UNTO, N., DELLEPIANE, M., SCOPIGNO, R., SODERBERG, B., AND LARSSON, L. (2011). Documentation and Interpretation of an Archeological Excavation: an experience with Dense Stereo Reconstruction tools. The 12th International Symposium on Virtual Reality, Archaeology and Cultural Heritage VAST, 2011.

CORSINI, M., DELLEPIANE, M., PONCHIO, F., SCOPIGNO, R. (2009). Image-to-Geometry Registration: a Mutual Information Method exploiting Illumination-related Geometric Properties. *Computer Graphics Forum* 28(7), pp.1755-1764. Available from: http://vcg.isti.cnr.it/Publications/2009/CDPS09'

D'AMELIO, S., LO BRUTTO, M. (2009.) Analisi e comparazioni di modelli di superficie di reperti archeologici realizzati tramite laser scanner e fotogrammetria. In: *Atti 13° Conferenza Nazionale ASITA, Bari 1-4 Dicembre 2009,* pp. 841-846.

DELL'UNTO, N., WALLERGÅRD, M., DELLEPIANE, M., LINDGREN, S., ERIKSSON, J., PETERSSON, B., AND PAARDEKOOPER R. (2010). An Experiment of Integrated Technologies in Digital Archaeology: Creation of New Pipelines to Increase the Perception of Archaeological Data. In: Contreras, G., Farjas M., Melero, F.J. (a cura di) *CAA2010: Fusion of Cultures. Proceedings of the 38th Annual Conference on Computer Applications and Quantitative Methods in Archaeology, Granada, Spain, 6-9 April 2010.* Oxford: Archaeopress, pp. 463-468.

DE LUCA, L. (2011). *La Fotomodellazione architettonica.* Palermo: Dario Flaccovio Editore.

DI GENNARO, F. (2006). Crustumerium e la sua necropoli, Roma. In: M. A. Tomei (a cura di). *Memorie dal sottosuolo - Ritrovamenti archeologici 1980/2006.* Verona: DemetraEdizioni, pp. 222-223.

DI GENNARO, F., VERGANTINI, L. (2000). *Via della Marcigliana. Crustumerium. Realizzazione di un primo itinerario di visita della necropoli e dell'area urbana. Archeologia e Giubileo. Gli interventi a Roma e nel Lazio nel Piano per il Grande Giubileo del 2000.* Napoli: Electa, pp. 459-465.

DOCCI, M., GAIANI, M., and MIGLIARI, R. 2001. Una nuova cultura per il rilevamento. In: *Disegnare, Idee immagini* 23. Roma: Gangemi Editore.

DOCCI M. (2011). Tradizione e tecniche innovative nel rilevamento e l'analisi dell'architettura, della città e dell'archeologia. In: Chiavoni E., Filippa, M., Docci M. (a cura di). *Metodologie integrate per il rilievo, il disegno, la modellazione dell'architettura e della città, Prin 2007.* Roma: Gangemi Editore, pp. 21-26.

DOCCI, M., MAESTRI, D. (2009). *Manuale di rilevamento architettonico e urbano*. Roma: EditoriLaterza, pp. 70-73;113- 127.

EL-HAKIM, S. (2006). A sequential approach to capture fine geometric details from images. In: *ISPRS Symposium: Image Engineering and Vision Metrology, Dresden, Germany September 2006,* pp. 97-102.

GRUEN, A., HANUSCH, T. (2004). Image- And Model-Based Reconstruction Of The Small Buddha Of Bamiyan, Afghanistan. In: *The International Archives of the Photogrammetry, Remote Sensing and Spatial Information Science*s, Vol. XXXIV, Part 5/ W12, 2004.

GRÜN, A., BEYER, H. 2001. System calibration through self-calibration. In: Grün A., Huang, T.S. *Calibration and orientation of cameras in Computer Vision.* Springer, pp. 163-193.

IPPOLITO, A. (2005). Dalla nuvola di punti alla superficie. Analisi e problematiche. In: *Metodologie innovative integrate per il rilevamento dell'architettura e dell'ambiente, ricerca Cofin 2004. A cura del Dip. di Rilievo, Analisi e Disegno dell'Ambiente e dell'Architettura - Università di Roma La Sapienza.* Roma: Gangemi Editore, Roma 2005. pp. 32-43.

IPPOLITO, A., BORGOGNI, F. (2011). I modelli 3D nei rilievi di architettura. In: Chiavoni E., Filippa, M., Docci M. (a cura di). *Metodologie integrate per il rilievo, il disegno, la modellazione dell'architettura e della città, Prin 2007.* Roma: Gangemi Editore, pp. 71-77.

IPPOLITO A. (2011). La costruzione dei modelli per il rilievo archeologico. In: Chiavoni E., Filippa, M., Docci M. (a cura di) *Metodologie integrate per il rilievo, il disegno, la modellazione dell'architettura e della città, Prin 2007. Roma: Gangemi Editore, pp. 27-34.*

SENATORE, L.J. (2011). *La scala del modello digitale. In: Chiavoni E., Filippa, M., Docci M. (a cura di) Metodologie integrate per il rilievo, il disegno, la modellazione dell'architettura e della città, Prin 2007.* Roma: Gangemi Editore, pp. 47-52.

WOJTAS, A.M. (2010). *Off-the-shelf close-range photogrammetric software for cultural heritage documentation at Stonehenge. In: International Archives of Photogrammetry, Remote Sensing and Spatial Information Sciences, Vol. XXXVIII, Part 5 Commission V Symposium, Newcastle upon Tyne, UK. 2010.*

From Survey to Representation of the Model A Documentation of Typological and Chronological Sequences of Archaeological Artefacts: Traditional and Innovative Approach

Alfonso IPPOLITO,[1] Luca J. SENATORE[1]
Barbara BELELLI MARCHESINI,[2] Gabriella CEROLI[3]

[1]Department of History, Representation and Restoration of Architecture
Sapienza University of Roma, Italy
[2]Post-Doc Researcher, Groningen Institute of Archaeology, The Netherlands
[3]Freelance archaelogycal restorer graduate from the
Central Institute of Restoration in Rome ICR, now IsCR, Italy
alfonso.ippolito@uniroma1.it, luca.senatore@uniroma1.it,
b.belellimarchesini@virgilio.it, gabriella.ceroli@gmail.com

Abstract

The integration of the traditional approach with the continuous technological advancement and in particular the great digital revolution offers new research tools that can be easily applied to science. Even though the fundamentals of the survey disciplines haven't changed, they have recently been interacting with new tools and applications for a long time. Not only in the discipline of architectural survey but also in the archaeological survey, technological advancements lead to the development of the integration between traditional methods and new low cost solutions of non contact survey for data collection.

This paper discusses different experiences in the digital-catching of information and aims at discussing advantaged and disadvantages of non-contact documentation in respect of the 'traditional' approach.

Keywords: Non contact survey, 3D models, Virtual restoration, Pyrgi, Dolium

Introduction

In the field of survey and representation, the introduction of new information technologies has quickly revolutionized the way we understand and address the matter. The technological advancements lead to the development of three different fields of research: the first deals with data collection; the second restricted one aims at the development of new low cost solutions; the third one concerns the traditional methods and their integration with the former two. We believe that it's not appropriate to focus our attention on a specific technique but to integrate all different tools and methodologies for the elaboration of 2d and 3d models, in order to achieve the best interpretation of the analyzed subject.

This premise makes it clear that the problem must be addressed as a whole, through the design, testing and concrete application of innovative management tools, suitable to deal case by case with the complex issues we have outlined. Such tools, however, cannot yield their full potential without a substantial change in perspective: archaeology cannot and must not be regarded as an isolated approach, but as part of a larger 'cultural cluster'. Obviously, this implies that a broader management and the implication of different skills is necessary: archaeologists, architects, restoration specialists, administrators.

This paper discusses different experiences in the digital-catching of information and aims at discussing advantaged and disadvantages of non-contact documentation in respect of the 'traditional' approach.

The aim is to discuss the advantages of digital catching during the phase of documentation and restoration of the artifacts. For this case-study we experimented the geometrical virtual reconstruction of hand-made irregular artifacts.

The aim of research is to evaluate if it's possible to virtualize their shape and profile and whether the digital approach is too extreme/stiff for such an application.

The selected case-study helps to highlight the benefits of (preliminary) virtual reconstructions in respect of expensive restoration of huge objects.

As regards digital-catching of information, the paper focuses the use of image-matching and compares the results which have been obtained, both in terms of precision and in terms of economic benefits.

This study aims to suggest a modus operandi for the standardization and regulation of data collection, processing and recovery procedures applied in order to make the final scientific results more objective and correct.

1. Case study: Dolium from Pyrgi

The case-study which is discussed in this paper is a dolium from the site of Pyrgi, which is a relevant maritime archaeological complex on the Mediterranean sea, about 60 km north of Rome (Colonna, 1996); (Figure 1).

In the Etruscan period (VIth-IIIrd century BC) Pyrgi was the main harbour of the town of Caere and held the function of international port of call. It was the seat of a Monumental Sanctuary, celebrated by ancient authors for its extraordinary wealth and attended by both Greeks and Phoenicians, and of a minor sanctuary which probably housed rites of passage (Colonna, 2000; Baglione, Gentili, 2013; Baglione, 2014). The sacred complex was erected next to the arrival of a huge extraurban road which connected Caere and Pyrgi, comparable with the one which linked Athens to the Pyreus. North of the sacred area, the Etruscan settlement extended over 10 hectares along the bay up to a rocky promontory, which was occupied in the IIIrd century by a roman maritime colony (Belelli Marchesini, 2014, with references).

Pyrgi is being explored since 1957 by Sapienza University of Rome, mainly focusing the sacred area. From 2009, a new line of research was addressed to the intersciplinary reconstruction of the topographical frame and the original landscape, which have been substantially modified by the sea ingression, the rise of the sea level and the disappearance of coastal lagoons. Fieldword is now focussing the intermediate area between the Monumental Sanctuary and the settlement, where a main pebbled road departing from the extraurban track and leading to the harbor has been brought to light (Baglione *et al.,* 2011).

Such road is flanked by different buildings endowed with both a residential and a public function. On the side of the sacred area, the building erected next to the connection of the two roads was provided with a portico and with a late archaic painted roof decorative system; its main function as a storeroom, possibly related with the commercial trades, is highlighted by the presence of several dolia.

Unfortunately the building and its furniture have been badly damaged by ploughing activities. Only one of the dolia was almost entirely preserved, since it was found in 2010 lying horizontally and squashed on the floor level. The dolium was broken into several pieces, most of them still in connection; nevertheless, the bottom, the mouth and part of body were partially missing, due to the presence of a plough furrow. Because of the dynamic of its collapse and compression, the dolium was mainly broken into two halves but some substantial portions were shifted and superimposed (Figure 2).

In the differents stages of the excavation, the dolium has been documented with 2D photographs and drawings. Each fragment was given a progressive number which mirrors the sequence of their removal, as to support further activity of conservation.

Figure 1a & 1b: Archaeological site of Pyrgi
Aerial view

Figure 2: Dolium's fragment
Images of excavation

Taking into account the field documentation, the fragments were immediately selected and assembled by the restorers, in order to reconstruct meaningful portions of the dolium, trace its profile and evaluate its volume. Because of its partial preservation and huge bulk, a complete reconstruction of the dolium itself seemed not to be worthwhile.

In the meantime, the dolium was selected as a peculiar case-study for virtual reconstructions.

The dolium was made by hand, shaping the bottom and progressively adding lumps or strips of clay in order to build up the walls (Caruso, 2003, pp. 60-85; Cuomo Di Capriio, 2007: pp.165-170; Emiliani, Corbara, 2002: pp.262-265); the strips were joined with liquid clay and pressed with the hands and the outer surface was progressively smoothed with proper tools. The craftsmen had to be skilled enough to prevent deformations of such huge artifacts, due to the dessication of the clay and the long-lasting process of fabrication (Geoponika, 6.3.4).

The irregular shape and the fragmentary preservation, as well as the lack of wheeling marks suggesting the orientation of the reconstructed portions of the dolium, were the main features to take into account in the preliminary approach to the case-study.

2. Virtual reconstruction

The first aim of our research was to find out how much the features of the dolium could affect the possibility of a 3D reconstruction. The reconstruction was mainly based on four assembled portions of the dolium: two of them including the rim and the shoulder (A and B); the others including the body (C) and the only preseved bit of the flat bottom (D) (Figure 3).

2.1 Approach 1 – Non contact survey

The first step of this study was aimed to understand if a digital approach exclusively based on new methods of low cost non-contact survey, as the technique of photomodelling, could allow the virtual reconstruction of the fragmentary dolium from Pyrgi

As regards methodolology, the photomodelling allows to obtain realistic 3d model from photos. The object is reconstructed by projecting in three-dimensional space of points and lines generated by image-matching algorithms, which allow the correlation of homologous points in the different frames in a totally automated. The connection between the pairs of frames is accomplished by dividing the images in a grid of small squares. For each frame of one of the two images, the algorithm allows to search automatically the one that has the most similarities in the other image color. The size of the grid squares, such as the minimum number of correlations between the frames, is represented by numerical values that must be set for a correct and conscious data extraction.

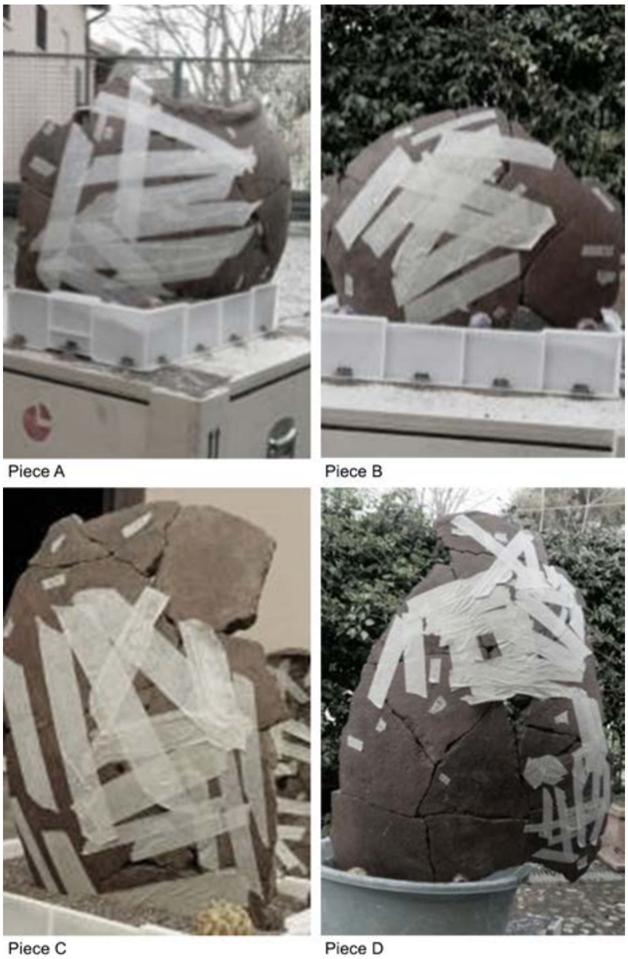


Figure 3: Dolium's fragment
Images of recomposition made by restorers

Nowadays there are many low cost and open source softwares for photomodelling. Each of these use always the same methodology that is divided into the following stages:

- Acquisition of photo images of the surveyed object, including the possibility to do shots from various angles, ensuring that part of them overlap. Thanks to this, the algorhythm can recognize homologous points in various photograms.
- Photography orientation: the homologous points in the photos are automatically transformed into points in 3d space characterized by cartesian XYZ coordinates and the chromatic RGB ones. This operation is indispensable to establish the exact position in the center of the projection, i.e. that of the photo camera in relation to the scene at the moment of acquiring the image; it determines precision as well as the number of points taken successively.
- Data elaboration and realization of 3D model throught specific software for management and control of dense point cloud. The point cloud is transformed into a 3D polygonal mesh that constitutes the surface of the object and makes it possible to vary the degree of detail. The mesh can be transferred into other 3D editor software to make

corrections like removing standing out or redundant points or closing holes inside the polygonal mesh.

- Restoring the outward look: here texture is added to the 3D model. This is done automatically and guarantees the perfect correspondence between geometric and chromatic data avoiding manual texture mapping operations which might result in an elevated level of uncertainty of the 3D model.

The photographs were done with a Canon 400d Reflex camera with 10.1 Megapixels and followed a scheme that guaranteed the whole object to be covered and each photograph overlapped with the successive one in at least 50-60 %; the phases of orientation frames and construction of the three-dimensional model were made through the software Agisoft PhotoScan. The survey data were elaborated through the viewer MeshLab, whereas the operations of scale and orientation of the 3D model were made with CAD software (Figures 4, 5).

For this case-study, scale operations were carried out taking into account real measures but also spheres of known size included in the photographs.

Due to the fragmentary conservation of the dolium, it was not possible to establish the correct orientation of its fragments in three dimensional space.

In this first step, the non-contact survey and digital approach did not therefore provide enough information to reconstruct the geometry and the shape of the case-study object.

2.2 Approach 2 – Non contact survey + Direct survey

Step 1-piece A + piece D

After the failing of the above procedure, we decided to integrate the data obtained from photomodelling with some information from direct survey, that is to say matching the digital and the traditional approach in order to better evaluate the 3D extension of the artifact, the procedure involved a new data acquisition: the direct survey allowed to obtain one vertical and the diameter of both the rim and the flat base as to support the 3D reconstruction of the artifact.

In spite of the irregular shape of the artifact, it was assumed its real shape could be reconstructed through the rotation of its hand-drawn vertical section around a vertical axis. In this way, the solid has been reconstructed following an ideal process that led to its virtual reconstruction; on the surface of the ideal solid we attempted to position the 3D model of piece A and piece D, previously acquired and processed with photomodelling tecniques.

The result was not fully satisfying because our starting point was just an ideal model obtained through a 2D approach.

Figure 4: Data acquisition Piece D. Example of dataset.

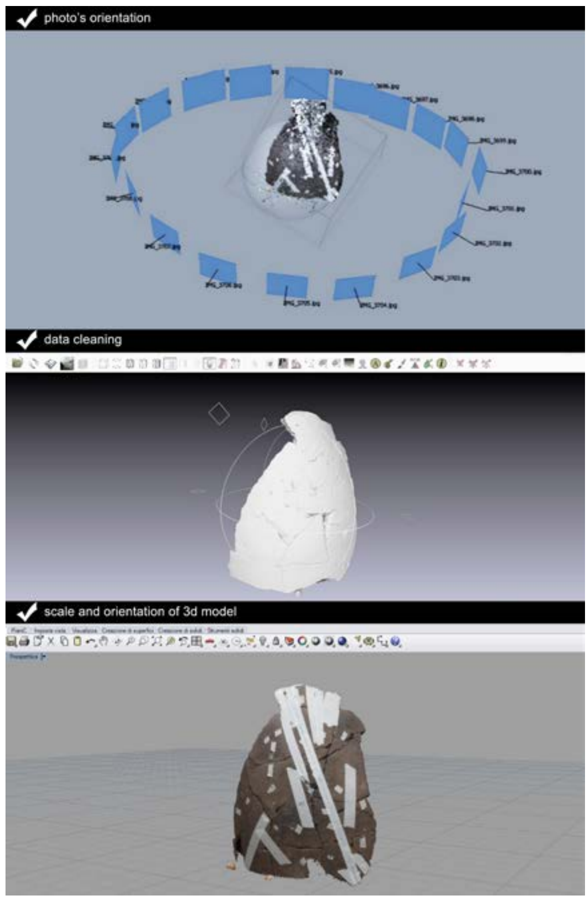


Figure 5: Data elaboration Piece D
1: photo's orientation: software Agisoft Photoscan 0.9.1
2: data cleaning: software MeshLab 1.3.3
3: operation on 3D model: software Rhinoceros 5.0

Nevertheless, this ideal model was the starting point for further analysis of the 3D model, such as the extraction of several horizontal sections along its vertical profile, proved that the obtained diametes were not aligned with the same vertical axis, but falling in an average range of about 8 cm off the ideal radius (Figures 6 & 7).

Virtualization throught data from direct survey of horizontal sections (edge and bottom) and vertical profile and position of piece A and piece D reconstructed throught photomodelling.

The 3D ideal model appears to be archaeologically unconvincing because it does not conform to the known types of dolia.

Step 2-piece A + piece D + piece B

In the following step we decided to increase the information from direct survey, adding more 2D horizontal sections of the dolium, in order to better identify the rotation's axis of the artifact and understand the geometry of the object: the new horizontal sections were traced on pieces A ans D including the rim the bottom, the shoulder and the body of the artifact and allowed to obtain diameters of the dolium at different heights.

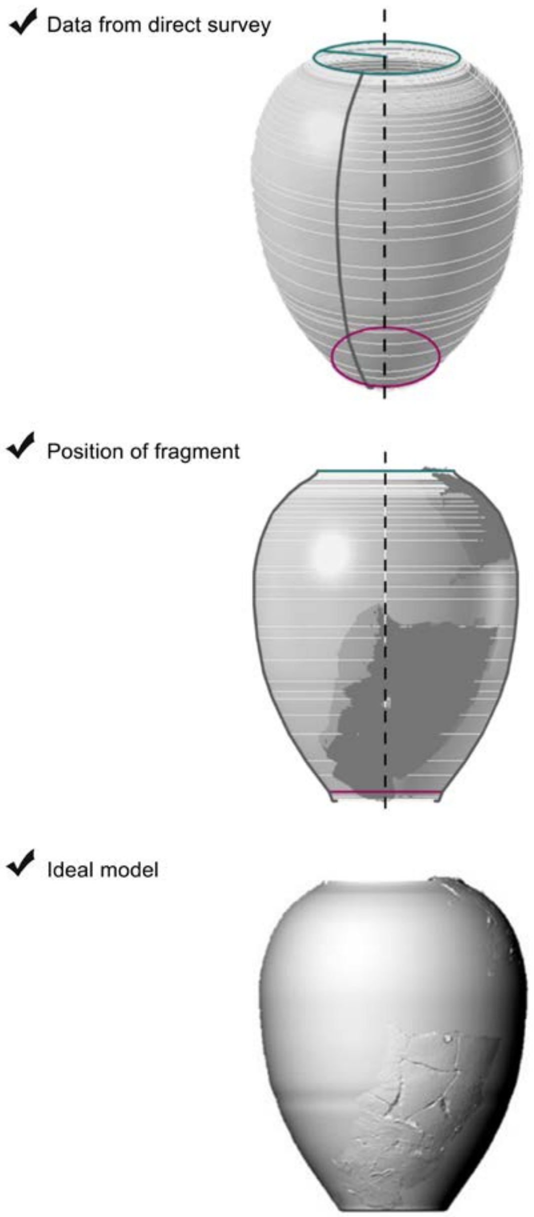


Figure 6: Ideal model
Virtualization throught data from direct survey of horizontale section (edge and bottom) and vertical profile and position of piece A and piece D reconstructed throught photomodelling.
The 3D ideal model appears to be archaeologically unconvincing because it does not conform to the known types of dolia.

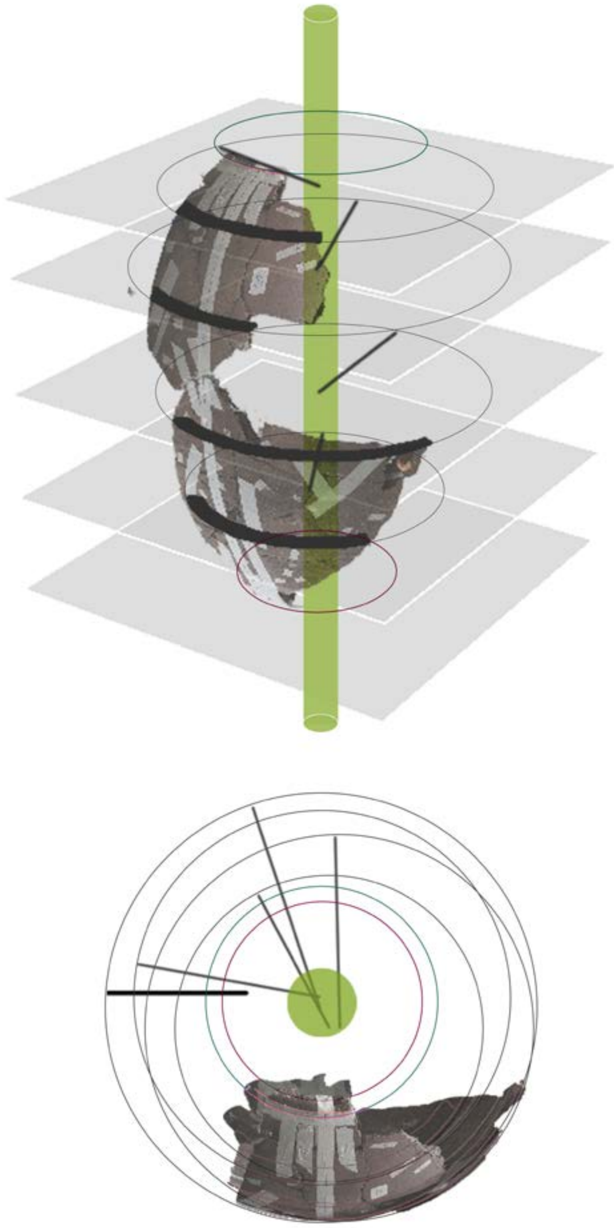


Figure 7: Real model
Centers are not aligned, they fall in a range with maximum radius 8 cm.

This phase of our procedure has led to unexpected results. The vertical profile of pieces A and D were in fact not compatible to the same vertical axis and were not matching as to reconstruct the same ideal geometric profile of the dolium. It was not possible to relate the different horizontal sections, obtained from the direct survey, to the same vertical profile and the axis of rotation therefore appeared not to be a vertical but a broken line.

The previous step proved that both the 3D and 2D reconstruction of the dolium from Pyrgi highlight its extremely irregular profiles and geometric shape, due to its technique of fabrication.

In the final step, we decided to understand what could be the most reliable position of another fragment of the upper part in relation with the lower one (piece B).

This second approach, in which data acquisition from direct survey has been essential, has allowed to elaborate two types of models: the ideal one obtained through the rotation of the vertical section along a vertical ideal axis (step 1), and the real one obtained through the horizontal and vertical profiles measured by the direct method (step 2) (Figure 8).

Approaching to a virtual reconstruction of irregular artifact through the integration of traditional approach and techniques of non contact survey has allowed us to have a wider view of the problem and to exploit the potential of both. Non-contact survey and technologies for digital reconstruction make possible to reconstruct the solid in 3d space; the traditional approach has allowed us to derive the horizontal and vertical profiles without which the virtual reconstruction would have not been possible.

Some theoretical and practical problems were encountered during the acquisition and processing of ta: the excessive fragmentation of the dolium and the lack of wheeling marks suggesting their exact orientation of the fragments themselves in the in the 3d space; the difficulty to obtain a continuous profile while matching the different fragments and their disalignment from an ideal central axis.

In conclusion, what we could reconstruct is not the real solid, but the most probable one.

This solid is the result of the the revolution along a vertical axis of a profile, which stems from the interpolation of the vertical profile of the ideal solid (step 1) and the vertical profile of the real solid (step 2). This profile is the most

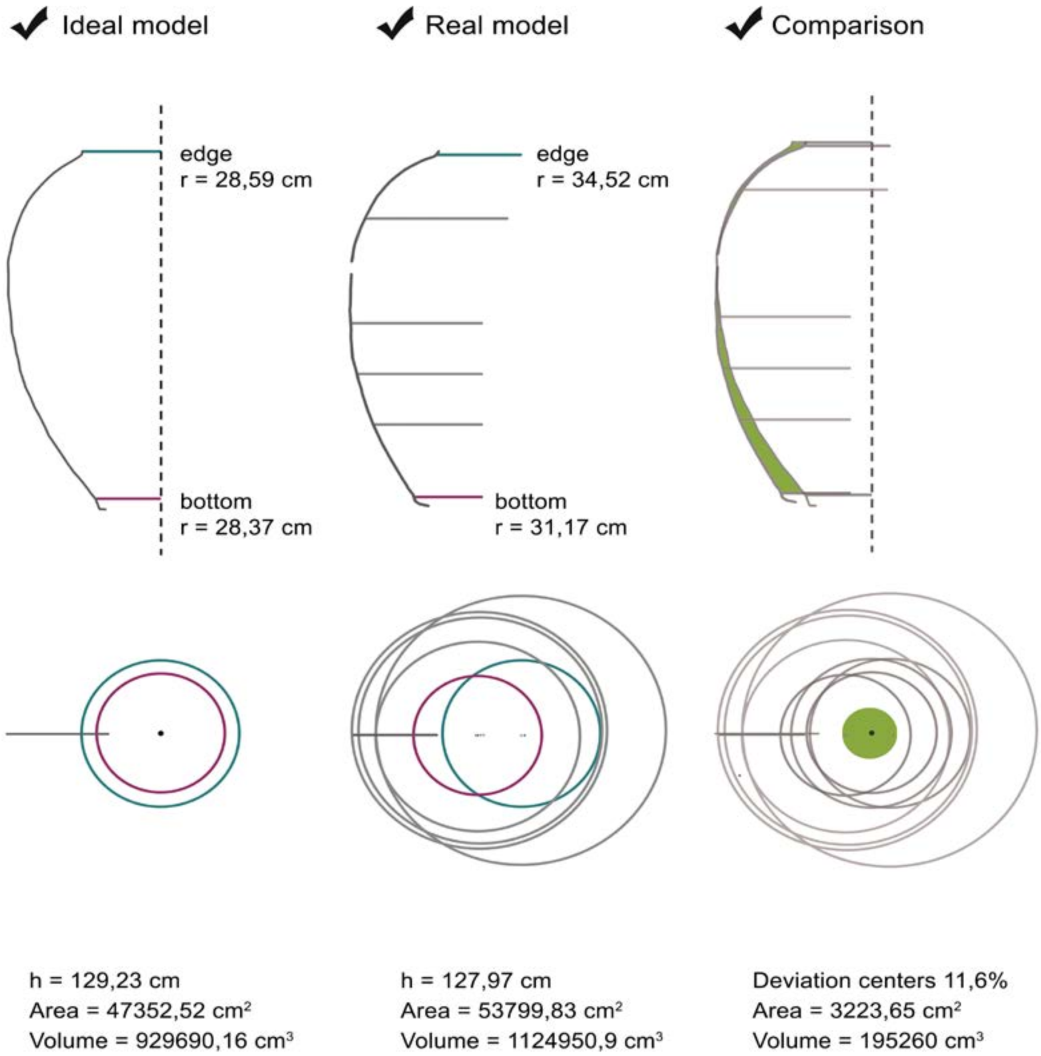


FIGURE 8: COMPARISON BETWEEN IDEAL MODEL AND REAL MODEL.

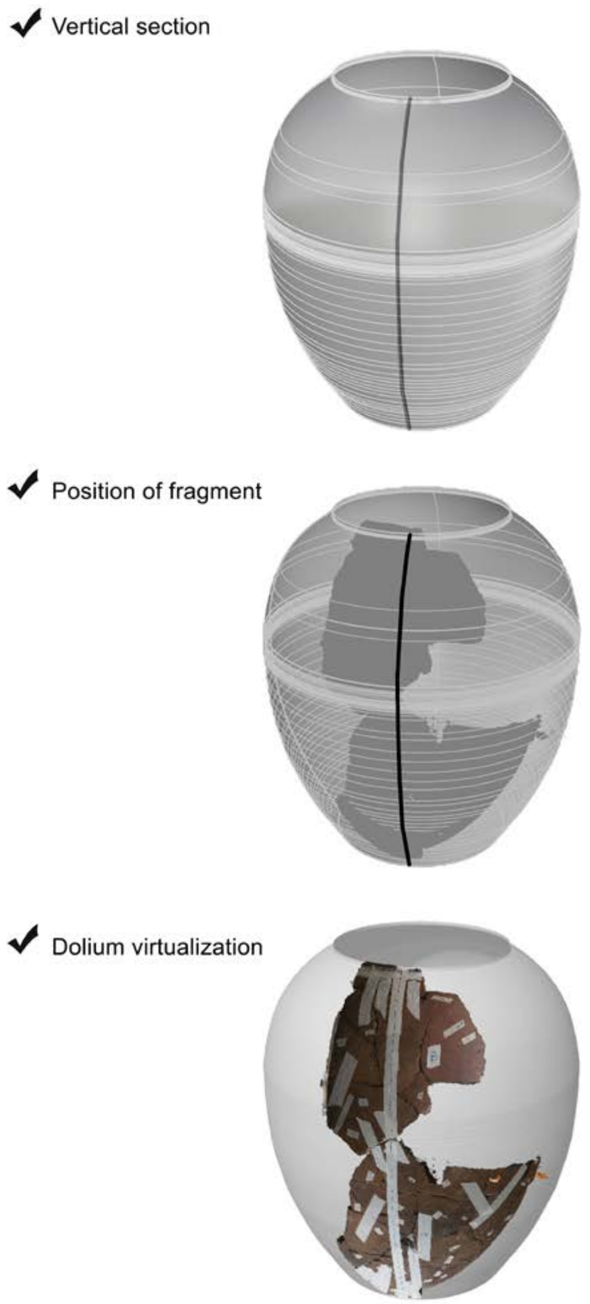


Figure 9: Virtual reconstruction of the most probable solid. The mean curve represents the most probable vertical profile of dolium.

probable geometric shape of the dolium from Pyrgi in 3d space, since it can be turned into a 3D solid where pieces A, D and B can easily fit (Figure 9).

The mean curve represents the most probable vertical profile of dolium.

Conclusions

The traditional approach aims at a selection of meaningful profiles and is affected by the preservation of the artifacts. The section drawn by an archaeologist turns into a shape which is meaningful from a typological and chronological point of view. From a typological point of view, the dolium of Pyrgi fits into the evolutive line of a group of artifacts featured by an elongated ovoid body and a flat base, dating from the VIIth century up the Hellenistic period.

When dealing with irregular artifacts, the interaction between traditional and 3D approach provides a precious opportunity to fully understand their peculiar features and modelling techniques. In fact, the obtained sections demonstrate that, in spite of the craftsman's skill, the dolium was certainly mishaped or not perfectly vertical.

Even though the 3D approach seems not to be suitable in such cases, our experience demonstrates that 3D model may help to measure the deviation of the many profiles from the 'ideal' model the Etruscan artisan had in mind.

References

ARBACE, L., ELISABETTA, S., CALLIERI, M., DELLEPIANE, M., FABBRI, M., ANTONIO, I. AND SCOPIGNO, R. (2013). Innovative uses of 3d digital technologies to assist the restoration of a fragmented terracotta statue. *Journal of Cultural Heritage* 14(4), pp. 332–345, July-Aug. [Online]. Available from: http://vcg.isti.cnr.it/Publications/ 2013/AECDFAS13

BAGLIONE, M.P. (2014). Pyrgi, un santuario nel cuore del Mediterraneo. In F. Gaultier, L. Haumesser; P. Santoro (eds.). *Gli Etruschi e il Mediterraneo. La città di Cerveteri (catalogo della mostra)*. Paris:Somogy, pp. 204-219.

BAGLIONE, M.P., GENTILI, M.D. (eds) (2013). *Riflessioni su Pyrgi. Scavi e ricerche nelle aree del santuario.* Roma: Erma di Bretschneider.

BELELLI MARCHESINI, B. (2014). Analisi delle fortificazioni della colonia romana e rapporti con l'abitato etrusco. In Bellelli V. (ed.). *Caere e Pyrgi: il territorio, la viabilità e le fortificazioni. Atti della Giornata di Studio*, Roma-Pisa: Fabrizio Serra Editore, pp. 199-263.

BAGLIONE, M.P., BELELLI MARCHESINI, B., CARLUCCI, C., MICHETTI LAURA M. (2010). Recenti indagini nel comprensorio archeologico di Pyrgi 2009-2010. *Scienze dell'Antichità* 16, pp. 541-560.

BIANCHINI, C., BORGOGNI, F., IPPOLITO, A., SENATORE, L. J. (2014). The Surveying and Representation Process Applied to Archaeology: A Quest for Invariants in a Highly Variable Context. In: Di Giamberardino, P. , Iacoviello, D., Natal Jorge, R., Tavare J. M. R. S. (eds.). *Computational Modeling of Objects Presented in Images*. Berlino: Springer, pp.1-29.

BIANCHINI, C., BORGOGNI, F., IPPOLITO, A., SENATORE, L., CAPIATO, E., CAPOCEFALO, C., and COSENTINO, F. (2012). Dal Rilievo di Rappresentanza. Background teorico, questioni pratiche, possibili linee guida. In: *VSMM 2012, 18 ° Conferenza internazionale sui sistemi virtuali e multimediali. Atti del Convegno Internazionale Milano*, pp. 507-513.

CARUSO N. (2003). *Ceramica viva. Manuale pratico delle tecniche di lavorazione antiche e moderne, dell'occidente e dell'oriente*. Milano: Hoepli, pp.60-85.

CHIABRANDO, F., NEX, F., PIATTI, D., and RINAUDO, F. (2010). Integrazione della fotocamera TOF e multi-picture approccio di corrispondenza per il rilievo dei beni culturali. In: *Archivio Internazionale di Fotogrammetria, Telerilevamento e Scienze informazioni spaziali*, XXXVIII, Parte 5, V Simposio Commissione. Newcastle upon Tyne, pp 143 148.

CIGNONI,P.CALLIERI,M.,CORSINI,M.,DELLEPIANE, M., GANOVELLI, F. and RANZUGLIA, G. (2008). MeshLab: an Open-Source Mesh Processing Tool. In *Sixth Eurographics Italian Chapter Conference*, Eurographics Association, Germania, pp. 129-136.

COLONNA G. (1996). *s.v. 'Pyrgi'. EAA, II suppl.* pp. 678-684.

COLONNA G. (2000). Il santuario di Pyrgi dalle origini mitistoriche agli altorilievi frontonali dei Sette e di Leucothea. *Scienze dell'Antichità* 10, pp. 251-336.

CORSINI, M., DELLEPIANE, M., PONCHIO, F., SCOPIGNO, R. (2009). Image-to-Geometry Registration: a Mutual Information Method exploiting Illumination-related Geometric Properties. In *Computer Graphics Forum* 28(7), pp.1755-1764. Available from: http://vcg.isti.cnr.it/Publications/2009/CDPS09'

CUOMO DI CAPRIO N. (2007). *Ceramica in archeologia 2, antiche tecniche di lavorazione e moderni metodi di indagine*. Roma: L'Erma di Bretschneider, pp.165-170.

DELL'UNTO, N., WALLERGÅRD, M., DELLEPIANE, M., LINDGREN, S., ERIKSSON, J., PETERSSON, B., AND PAARDEKOOPER R. (2010). An Experiment of Integrated Technologies in Digital Archaeology: Creation of New Pipelines to Increase the Perception of Archaeological Data. In: Contreras, G., Farjas M., Melero, F.J. (eds.). *CAA2010: Fusion of Cultures. Proceedings of the 38th Annual Conference on Computer Applications and Quantitative Methods in Archaeology*, Granada, Spain, 6-9 April 2010. Oxford: Archaeopress, pp. 463-468.

DE LUCA, L. (2011) *La Fotomodellazione architettonica.* Palermo: Dario Flaccovio Editore.

DE LUCA, L., BUSARAYAT, C., STEFANI, C., RENAUDIN, N., FLORENZANO, M., VÉRON, P. (2010) An Iconography-Based Modeling Approach for the Spatio-Temporal Analysis of Architectural Heritage. In: *Shape Modeling International Conference (SMI '10) Washington, USA*, pp. 78-89.

EL-HAKIM, S., BIZZARRI, L., VOLTOLINI, F., GIRARDI, S., RIZZI, A., REMONDINO, F., and WHITING, E. (2007) Dettagliata modellazione 3D di castelli. *International Journal of Architectural Computing* 5(2), pp. 199-220. Available from: http :/ / www.inf.ethz.ch/personal/whitinge/resources/pubs/IJAC-5-2-El-Hakim.pdf

EL-HAKIM, S. (2006). A sequential approach to capture fine geometric details from images. In: *ISPRS Symposium: Image Engineering and Vision Metrology,* Commission V Symposium Germany September 2006, IAPRS&SIS 36(5), Dresden: ISPRS, pp. 97-102.

EMILIANI G.P., CORBARA F. (2002). *Tecnologia Ceramica. La lavorazione. Vol.II*. Faenza: Faenza editrice, pp. 262-265.

FRATUS DE BALESTRINI, E.,GUERRA, F. (2011). New instruments for survey: on line softwares for 3d reconstruction from images. In: *The International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences* XXXVIII-5/W16, pp. 545- 552.

GAIANI, M., BENEDETTI, B., and APOLLONIO, F. (2011) Teorie rappresentare e per i siti archeologici di Comunicare ATTRAVERSO Modelli Critici. In: *Scires-IT Ricerca Scientifica e Tecnologia Ricerca Scientifica e Tecnologie dell'Informazione*. CASPUR - CIBER Publishing, 1 (2), pp. 37-70.

GRUEN, A., HANUSCH, T. (2004) Image- And Model-Based Reconstruction Of The Small Buddha Of Bamiyan, Afghanistan. In: *The International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences*, XXXIV, Part 5/ W12, 2004.

IPPOLITO A. (2011) La costruzione dei modelli per il rilievo archeologico. In: Chiavoni E., Filippa, M., Docci M. (eds.) *Metodologie integrate per il rilievo, il disegno, la modellazione dell'architettura e della città*, Prinn 2007. Roma: Gangemi Editore, pp. 27-34.

IPPOLITO, A. (2007) Dalla nuvola di punti alla superficie. Analisi e problematiche. In: *Metodi e TECNICHE integrate di rilevamento per la Costruzione e Modelli di fruizione virtualizzazione 3D dell'Architettura e della Città. Ricerca COFIN 2004* Roma: Gangemi Editore, pp 32 43.

KERSTEN, THOMAS P., LINDSTAEDT,MAREN (2012). Image-Based Low-Cost Systems for automatic 3D Recording and Modelling of Archaeological Finds and Objects. In: Ioannides M., Fritsch D. *et al.* (eds.), *Progress in Cultural Heritage Preservation Euro-Med 2012*, LNCS7616. Berlino: Springer-Verlag, pp. 1-10.

KLEBER, F. AND SABLATNIG, R. (2009). A survey of techniques for document and archaeology artefact reconstruction. In: *Proceedings of the 2009 10th International Conference on Document Analysis and Recognition*, ser. ICDAR '09. Washington, DC, USA: IEEE Computer Society, 2009, pp. 1061–1065. [Online]. Available from: http: //dx.doi.org/10.1109/ICDAR.2009.154.

MOLYNEAUX, B. (2011) *La vita culturale di immagini rappresentazione visiva in Archeologia.* London: Routledge, p. 296.

WOJTAS, A.M. (2010) Off-the-shelf close-range photogrammetric software for cultural heritage documentation at Stonehenge. In: *International Archives of Photogrammetry, Remote Sensing and Spatial Information Sciences*, Vol. XXXVIII, Part 5 Commission V Symposium, Newcastle upon Tyne, UK. 2010.

Archaeology in the Third and Fourth Dimensions: A Case Study of 3D Data Collection and Analysis From Prince Rupert, BC, Canada

Alyssa PARKER, Morley ELDRIDGE

Millennia Research Limited, Victoria, BC, Canada
alyssa@millennia-research.com, morley@millennia-research.com

Abstract

During a large excavation project in Prince Rupert, BC, Canada, a very large data set of 3-dimensional data points was collected. A Total Station was used for 3D provenience of all materials; supplemented with tablets for attribute data collection, and linking the attribute and spatial data with barcodes. This 3D dataset was augmented by photography of features for conversion to 3D models using SFM software. These models have been georeferenced and visualized with the surrounding point cloud data, which is symbolized by a variety of different attributes. This visualization has allowed detection of patterns that were not apparent in the field. The point cloud and models allow an unprecedented view of the archaeology of this region. This paper discusses the methods used to collect, store, and analyse this data, including different methods of 3D visualization and data sharing using web-based applications.

Keywords: 3D Archaeology, Digital Recording, Total Station, Barcode, Tablet

Introduction

The Kaien Siding Archaeology Project, Prince Rupert, BC, Canada (Figure 1) was one of the largest ever to be conducted in the province in terms of the scale of the excavation, the number of artifacts and faunal remains found, and the in-depth research undertaken afterwards. The Kaien Siding project was also a world-leader in creating an archaeological excavation and data management system built around concepts of using 3D space as the basic organizational principal, and making obsolete many of the traditional pillars of archaeological method, such as excavating rectangular units.

Figure 1: Project location, Prince Rupert, British Columbia, Canada.

The objective of the project was for the totality of cultural materials to be excavated prior to construction commencing. Undisturbed cultural materials were to be excavated using hand techniques while previously disturbed materials could be excavated with machine assistance. This approach, though fairly common in Europe and Asia, had not been taken previously with Northwest Coast shell middens.

The project formed a new methodological and technological paradigm for archaeological excavation. The main underlying concept beneath this new paradigm was that of 'unit-less' infinitely scalable archaeology. To this end, a system was designed to be able to quickly and efficiently perform a 100% excavation of two large archaeological sites while retaining or improving on the quality and amount of data collected as compared to traditional excavation techniques.

The systems created to complete this very large archaeological project were both novel and involved considerable risk. The risk stemmed from several sources, principally the effective volume rate of systematic excavation which we could expect crews to average; whether full-scale excavation requiring boat access could safely continue through the winter; whether the novel, unproven approach taken to both excavation and data recording would work as planned; and whether the volumes of archaeological deposit predicted from AIA testing were accurate. As it turned out, one site had less volume than expected, but the other contained a great deal more: in total, twice as much volume had to be removed compared to the original combined estimate. Thankfully the combination of technology and approach resulted in much faster than expected excavation rates compared to conventional methods, yet with very good data quality. This allowed completion of the project within the allocated budget and just slightly later than initially anticipated. The approach and methods are described in much more detail in following sections of this paper, but in brief, it combined an abandonment of traditional archaeological grid and unit-based excavation in favour of 'unitless' block excavation, the use of barcodes to replace hand labelling of notes and samples, and total station survey instruments combined with tablet computers to replace conventional measuring and detail recording.

Digital recording techniques in various forms have been used in archaeology world-wide for many years. Dibble *et al.* (2007) described a system which was the inspiration for ours. Dibble's system consisted of a total station connected to a barcode scanner and laptop computer for increased efficiency and accuracy with recording archaeological features. Other researchers have employed tablets in field excavation for paperless recording (Caraher, 2013; E'se'get Archaeology Project, 2012). However, to our knowledge, a system such as the one we used has not previously been used for a large-scale excavation project.

In particular, our system and resulting data analysis are unique in consulting archaeology. This paper will describe the technology used, our data management procedures, some of the analysis techniques applied to the resulting data, and a discussion of the benefits and drawbacks of this excavation technique.

1. Methods

1.1 Field techniques and technology

Robotic total stations were used to record spatial data in three dimensions and in real-world geographic coordinates; tablet computers with forms for collection of attribute data about the excavated materials and features (Figure 2); barcode scanners; and a suite of software products for collecting, compiling, analysing, and storing the data gathered.

The robotic total stations, Trimble S3 with TSC3 handheld data collectors/controllers, were vital to the project. Surveyed control points were initially placed by a professional surveyor, and thereafter all data was recorded in absolute geographic coordinates in three dimensions. This meant that the strict control using rectangular units in a local Cartesian grid with datums and levels was not required to provide spatial context. Thus, the excavation could proceed in order to maximize efficiency and two widely separated areas ('sites') could be combined into a single recording system.

Tablets proved to be a particularly useful tool in the field data collection process. Their main purpose, of course, was to reduce the need for note taking on paper, and thus removing the need for tedious and time-consuming transcription while eliminating the errors inherent in such transcription. Also, through auto filled date, time, and personnel stamps, dropdown menus, and validation controls on the forms, data collection was standardized.

FIGURE 2: A CREW MEMBER RECORDS DATA USING A TABLET COMPUTER AND A ROBOTIC TOTAL STATION.

Photographs could be taken with the tablets as well, and attached to the form entry, providing a link between photograph and textual data. Though the forms were the primary data collection format on the tablets, there was also a 'sketchpad' application that allowed the user to sketch diagrams, or to annotate photographs; these files could also be attached to a form entry. The forms were generated through an online forms service, which provides several benefits over static forms (such as PDF forms) saved on the tablets. These benefits include being able to modify form design in the office and deploy the modified forms to all users without requiring the user to do anything: dropdown lists of values could be easily modified by authorized personnel and instantly distributed, without ending up with multiple non-synchronized versions of the forms.

The online forms service we used was ProntoForms. While there are several other applications similar to this one, ProntoForms was selected because it worked on a variety of mobile device operating systems, it had most of the form features that we wanted, and it was simple to use and administer the forms. Of the applications initially tested, ProntoForms combined the majority of the most critical design and administration features. The number of forms we could create was unlimited. We were able to create forms for specific types of data – such as Artifacts, Features, Bucket Groups, Stratigraphic Breaks, etc. The forms were not limited to archaeological data. A smartphone form was used to record the beginning, end, timing, and destination of boat trips and passengers, which was used for safety, security, and payroll. Each form was customized with drop-down menus and other data fields that were relevant to the particular type of data being recorded (Figure 3). Date and time fields were auto-populated when a new data entry form was started. Fields could be limited to certain types of data (e.g. numeric only fields) to ensure conformity of the data; and fields could be made 'required fields', which ensured that critical data was always recorded and not left blank. The forms were not static once they had been created: forms were modified throughout the project as additional needs were identified; for example, dropdown menu items could be changed to reflect changing conditions at the site. In fact, dropdown menu changes that needed to occur more frequently than the office staff could accommodate were set up so that the field coordinator could simply edit a Google Docs (now Google Drive) spreadsheet, and the forms server pulled the information from this spreadsheet.

One minor drawback to using the online forms service was that internet access was required in order to download updates to the forms, as well as to submit completed forms. However, completed form data records were saved on the tablet when internet access was not available, and the forms, once downloaded to the tablet, could be used without requiring internet access. When the tablet was next connected to the internet, uploads of completed form data records and downloads of any form updates or data source updates were processed automatically. Therefore, there was no delay in recording due to the availability of internet. We did have internet available on site through the use of a mobile wireless internet network hub which connects to the cellular phone network. Those working out of range of this wireless network could bring their tablets within range at the lunch break or end of the day so that forms were updated and data records uploaded.

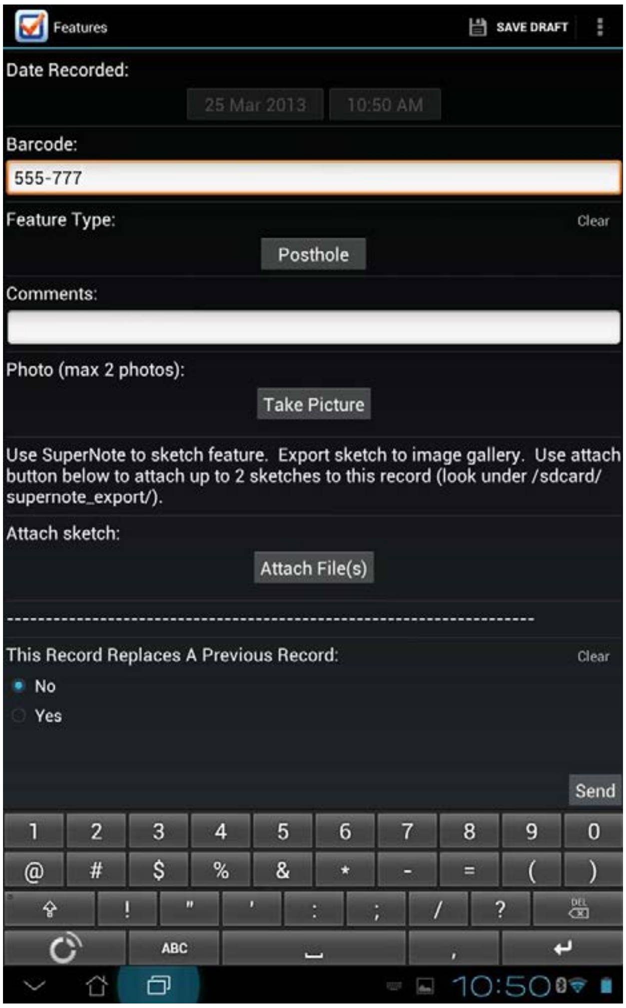


FIGURE 3: SAMPLE VIEW OF A FORM ON THE TABLET.

While we have described above the individual components of our system, the key of course was to link the spatial data, collected with the total stations, to the attribute data, collected with the tablets and create a system that was useful for analyses by a wide range of specialists. Theoretically, the total station data collectors could have been used to collect the attribute data and the two data types then would be directly linked; however, this would have greatly reduced efficiency, as one person would be responsible for recording both spatial and attribute data, which would have slowed down the work considerably. This would have required far more total station units to be available, and tablets are a fraction of the cost of a total station or RTK instrument, the other choice for maintaining spatial control. Another issue is the limited data collection fields available in the built-in total station software, which would have required customized software. Thus, rather than directly connecting the total station equipment to the attribute data collection equipment, we simply used barcodes to uniquely identify each and every bit of data collected. Every spatial data point was associated with a barcode, and every attribute data record was associated with a barcode (Figure 4). Barcodes also provided a vital physical tag for labelling collection bags at the time of collection. The barcode became the key component of this data collection system.

Barcodes were chosen (over other identification systems) for several reasons. One is the ability to use barcode scanners to scan the barcodes and thus avoid data entry errors. The barcodes also comprised a pre-defined unique ID system, which limits errors of using the same ID for two different objects. Moreover, use of barcode scanners is not simply for data entry, but also for use in the lab or office to quickly scan and find all relevant data related to a particular barcode.

FIGURE 4: A BARCODE SERVES AS THE UNIQUE IDENTIFIER FOR EACH RECORDED OBJECT, SUCH AS THIS BRACELET FRAGMENT.

For field purposes, we used handheld wireless laser barcode scanners. These could be connected by Bluetooth to the field tablets or the total station data collectors. In the lab, wired barcode scanners were connected to the lab computers for easy scanning of barcodes to both inventory items coming into the lab, as well as to easily search for the preliminary field data collected about a specific barcode.

A brief description of how this system worked together in the field follows. Artifacts found *in situ* were given barcodes and recorded spatially with the total station before they were removed; attributes about the artifact were recorded using the form on the tablet – photos were taken with the tablet, and, in some cases, also with a digital camera. Where possible, the barcode was included in the photograph – however, as the tablet photo filenames included the barcode they could be used to cross-reference the point-and-shoot higher resolution photographs of the artifacts where the barcode might not have been visible or legible in the photograph. Artifacts were bagged with their barcode, and transported to the lab. Features were also recorded *in situ*. Initially, features were recorded with a dense number of total station points taken around and within it, in order to create a 3D representation of the feature. However, later in the project, the methodology changed to using structure from motion (SFM) photogrammetry techniques to generate models, as this method was far superior in the quality and efficiency of creating the model. For this, at least 4 barcodes were placed around the feature to act as control points, and total station shots taken for these. Then a series of photographs were taken around and above the feature (ensuring that the barcodes were visible and legible in the photographs). These were processed in the office as described below.

Stratigraphy was recorded using several methods. One was a variant of traditional stratigraphic profiling. Once a trench or block wall profile was exposed the stratigraphic breaks visible were 'traced' using a prism. The robotic total station would automatically follow the prism, and would take data points at specific time or distance intervals, allowing the recorder to simply move the prism following the stratum and a series of points for that break were recorded. In addition, four control point barcodes were placed at the top and bottom, left and right extents of the profile, and total station points taken for these barcodes. Photographs were then taken, ensuring the control point barcodes were visible and legible. Some of these were done with the intent of using SFM to create a 3D model of the profile wall. Another method of recording stratigraphic data was through recording the surface of a clear stratigraphic change as it was exposed during excavation. An irregular judgemental array of total station shots was taken over this surface, and attributes about the newly visible stratigraphy were recorded using the tablets. Each stratigraphic break was given a unique barcode as an identifier.

Between these stratigraphic breaks, the soil composition was recorded through the use of 'bucket groups', which were the functional equivalent of a unit level or lot. As areas were excavated, the excavated materials were put into buckets for transfer to the screening station. Generally, several buckets worth of sediment would be excavated at a time from a relatively constrained area – these were considered a 'bucket group'. Each bucket group was given a unique barcode, and the number of buckets (for volume) and a description of the sediment were recorded in the tablets. The location this material was excavated from was recorded spatially with around five total station points (one in the center of the area, and four more around the edges to roughly delineate the area that the material came from; this allowed the option of calculating bucket group volume using geometry, as well as providing points for mapping). On average there were six buckets per bucket group, though this number was subject to change dependent on site conditions – for example, if a change in soil matrix became apparent before six buckets were filled and there were constraints on expansion sideways, excavation would pause and a new 'bucket group' would be started for the new stratum.

The screening shed crews were also equipped with tablets. All bucket groups were transferred to the screening shed to be screened for faunal material, artifacts, (for transport to the lab) and for FCR and other clasts, which were weighed then discarded. Bucket groups therefore received two entries in the tablets – the first entry was the excavator's description of the bucket group; the second was the screening crew's notations about the bucket group, including checkboxes for whether or not artifacts, faunal, or other material of interest was found. In addition, both the excavator and the screening crew recorded the number of buckets in the bucket group; this data was important for volumetric estimates, and recording it twice provided a check and redundancy in case of mis-entry of data. The screening crew also recorded information about artifacts and faunal material, and placed these in separate bags, with a copy of the barcode for the bucket group (multiple barcodes were pre-printed so that this could occur). Artifacts found in screening were given the barcode of the bucket group from which they came, but were otherwise processed the same as those found *in situ*.

1.2 Data management

Field data was uploaded immediately to the ProntoForms servers via the internet connection; from there it could be accessed in a number of ways. A web browser interface allowed viewing of submitted data records by both the field user as well as the office and lab personnel. In addition, each submitted data record was sent in a PDF (portable document file) format to our office server via FTP. Attachments could be embedded in the PDF or sent as separate files. The latter was the option we used for most of the data records, so that photos and sketches would be available as JPG images. The automated file naming of both the PDF and JPG included the barcode and the date and time of data entry; the JPGs also included a suffix of 'Photo' and the photo number (in the case of multiple photos attached to one data record) or 'Sketch' and the sketch number.

While the PDF files were useful for viewing individual data records, they are not suitable for managing large quantities of data. The data in CSV (comma separated values) format was also downloaded. In fact, ProntoForms provided the ability to set up automated export routines, including specification of filters including which forms' data records, specific date ranges, etc. were included in the export, as well as the formatting of the fields of the CSV. Using this functionality we configured an automatic daily export of all field-collected data in CSV format, which created a flat table of all the data from selected forms and e-mailed the file to the lab personnel and the GIS personnel (both at the project and at the main company office). CSV format is ideal for import into any number of different software; for our project, we used Microsoft Access as database software to manage the attribute that was submitted.

This field data was imported (semi-automatically) on a daily basis. A relational database was created in Microsoft Access to store and manage the data. Import procedures were programmed into the Access database, so that the lab personnel had only to manually import the CSV file into the database; all other import procedures were automated through the use of queries, macros, and code. These import procedures included data checking for issues such as incorrectly formatted barcodes (a result of manual data entry of barcodes when barcode scanners were not working). Also, cross-checks between excavator data entry and screening station data entry were performed, to check for barcodes that were included in one data set but not the other. These issues could therefore be discovered early and relayed to the field crews, and, where possible, the missing data provided. Other checks included comparison of field-submitted data records for artifacts against items actually received in the lab. The database was then used to generate reports that could be printed for use in lab processing procedures (checklists etc.), and data entry by lab personnel to record processing procedures and additional information.

The total station data was emailed on a daily basis to both the lab personnel and the office GIS personnel. This was again made possible through the on-site wireless internet network. This task had to be done manually by the field crews, but the files could be e-mailed directly from the data controller without requiring connection to a computer. Once received in the office, these files were imported both into ArcGIS, as well as into the Microsoft Access database, where they could be checked for errors. Queries were set up to identify attribute data records without matching spatial data, and vice versa, which allowed for quick follow-up corrections in the field where data was missed.

ArcGIS 10 with 3D Analyst was used as the GIS for this project. The ArcScene data viewer was invaluable for viewing the 3D point data. The spatial data was stored in a File Geodatabase, and each day's data appended to the file, so that all data was stored in a single file, making it simple to manage the data. Once imported, the spatial data was checked for errors such as incorrect elevation (often apparent when one or more points were floating either well above or below the rest of the cloud), which was often a case of an incorrect Height of Target being entered in the total station (and therefore simple to correct). Through an OLE database connection to Microsoft Access, the attribute data could be joined and preliminary results viewed, to inform the continuing excavation. While we did not have the capacity to have a dedicated GIS person on-site, that is certainly possible using our system, which would allow near-instantaneous viewing of collected data in 3D.

1.3 3D GIS Analysis

The excavation and data collection resulted in over 140,000 spatial data points being recorded. This data, initially stored in a single File Geodatabase feature class, was too large for efficient processing and analysis, and so was broken down into several smaller datasets once the final data collection was complete. Feature classes were created for major data types, such as artifacts, stratigraphic breaks, bucket groups, etc. The Microsoft Access database tables and queries were then linked to the total station data to allow spatial analysis and visualization of the attribute data.

Visualization was the primary analysis technique used. By symbolizing the data points to represent some attribute or combination of attributes, patterns could be detected in the data. For example, bucket group data was symbolized using graded colour by the percentage of shell (Figure 5). This visualization allowed detection of patterns including the location of a possible house. These visualizations of the point cloud were enhanced by the addition of other 3D object types. These include multipatch objects created from 3D models (generated from SFM), or from convex hulls of sets of data points (Figures 6, 7). TINs (Triangulated Irregular Networks) of both the surface of the site prior to excavation, and of stratigraphic breaks or other surfaces within the dataset, were also generated and used to inform analysis. 3D line features were created to represent landmarks or features such as canoe runs on the beach in front of the site, and extruded 3D polygons to simulate house posts where posthole features had been recorded. The combination of these various 3D objects and the symbolized point cloud within the 3D ArcScene viewer created a powerful visual re-creation of the sites and permitted identification of features and patterns which had not been readily apparent during the field excavation.

In addition, 3D models of some artifacts were generated through use of a laser scanner, and these models were placed into the GIS at the 'in-situ' coordinates where the artifact was found (or as close to as possible). Where the data had been recorded in sufficient detail, the artifacts were also oriented in the manner in which they were found. This provides an enhanced visual experience when assessing patterns in the location of artifacts.

Further analysis relied on dividing the site into smaller units to allow detection of grouped data trends not easily apparent in the classified point cloud. There were several discrete areas for each site, and these were then subdivided vertically into temporally-based 'analysis units'. Although the stratigraphy was extremely complex, as normal for shell middens (Stein, 1992) a basic breakdown was accomplished using the location and age of radiocarbon dates combined with symbolizing the point cloud according to a number of different attributes, (such as the percentage of shell and the dominant shell species). Relatively clear stratigraphic differences became apparent which appeared to correspond with the ages of the deposits. TINs were created of these stratigraphic breaks, and these were then transformed into volume objects that could be used to subdivide the point cloud attribute data. The two main datasets that were analysed this way include artifacts and faunal data. All of the items from these datasets were classified by Analysis Unit and Excavation Area.

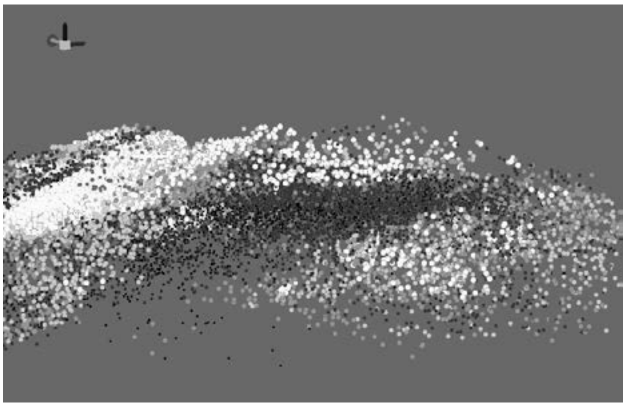


FIGURE 5: 3D BUCKET GROUP DATA POINT CLOUD SYMBOLIZED BY PERCENTAGE OF SHELL. SMALLER, DARKER DOTS INDICATE LOW SHELL PERCENTAGES, WHILE LARGER, LIGHTER-COLOURED DOTS REPRESENT HIGH SHELL PERCENTAGES.

FIGURE 6: EXAMPLE OF 3D MODELS OF FEATURES GENERATED USING SFM TECHNIQUES AND CONVEX HULLS.

FIGURE 7: DETAIL OF A HEARTH FEATURE MODELLED USING SFM PHOTOGRAMMETRY TECHNIQUES.

Comparisons of counts of different artifact types or faunal elements were performed using this breakdown, to attempt to discern patterns and trends. Figure 8 is an example, showing the dominant bird species that were identified in the faunal materials, broken down by excavation area and analysis unit. There appears to be a significant increase through time in the dominance of a small seabird, rhinoceros auklet, not locally available. The figure also shows a near total absence of birds in Area 5, AU3 both 'within house' and 'south of house' (although this pattern was not repeated for mammals or fish, and the samples are also relatively small). While the explanation of the patterning is beyond the scope of this paper, the capabilities of the system are evident.

1.4 Data sharing

A key question with any 3D data stores is how to share and present the data, when most media are inherently

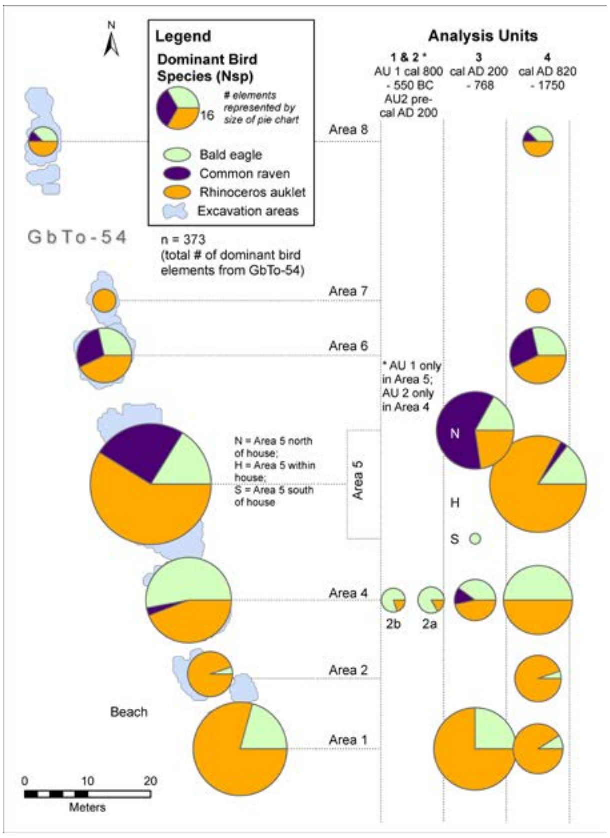


FIGURE 8 : DOMINANT BIRD SPECIES BY EXCAVATION UNIT OVERALL, AS WELL AS BY TEMPORAL ANALYSIS UNIT.

2-dimensional. However, the internet is rapidly expanding the options for sharing 3-dimensional data. While it is out of the scope of this paper to discuss this issue in detail, and there are others involved in such a detailed discussion (Gabii Goes Digital, 2013), we will discuss briefly the methods we have used to make our data available to the public and other researchers. We have used social media as a platform for sharing our data, including our Blog, Facebook page, and YouTube. Videos of animations of the data in ArcScene are the primary method – the benefit of video is being able to control exactly what is shown and how it is shown, like a 2-dimensional image would, but with the additional visual of being able to see the data in motion, which is arguably essential for comprehending 3D data via a 2D medium (the computer screen). However, the drawback of a video is that it does not allow user interactivity with the data, and it is essentially the 3D version of a static image.

To address this issue, we explored the use of other, interactive 3D web viewers. The one that proved the most useful for our purposes was ArcGIS's CityEngine Web Viewer. With settings set appropriately, an ArcScene document can be exported to the CityEngine format, which can then be viewed and explored in the Web Viewer. While an ArcGIS license is required to be able to create, and to privately share, the scenes, any publicly-shared scenes are accessible to anyone with the requisite computing requirements (WebGL-capable video card and web browser being the primary requirements). CityEngine allows a user to turn layers on and off, perform (limited) queries and searches, and to rotate, pan, and fly through the scene. It also allows side-by-side comparison of data sets (Figure 9). However, a scene is static once exported from ArcGIS; if the data is updated in the ArcGIS, the scene needs to be re-exported and re-uploaded. Another drawback of the CityEngine is that attributes such as the colour, size, and shape of features is set according to how they are displayed in ArcScene when the scene is exported – the user cannot change these. Also, the amount of data that can be included in a scene is somewhat limited (more by the computing capacity of the user's hardware than by limitations in CityEngine itself; overly large scenes displayed without issue on a higher-powered computer, but caused a lower-powered computer's browser to crash when loading). These caveats aside, it is a good tool for sharing subsets of the data, allowing the user to move about the scene at will and to query individual data points or objects. In particular, we have used CityEngine to share data with other researchers working on specific analyses (such as faunal and human remains analyses), who do not have access to nor the expertise to use ArcGIS software; yet for whom the spatial patterning is important to inform their analyses.

2. Discussion

Overall the system performed very well. Despite a number of issues that came up during the project, the data collected and the analysis we were able to perform is unprecedented in BC consulting archaeology. Errors were primarily those that would affect any large excavation project. One of the principal sources of error was that of incorrectly entered barcodes; this is fairly critical as the barcodes were the basis of linking the spatial and attribute data. This was largely due to technological issues with connecting the barcode

FIGURE 9: SAMPLE VIEW OF CITYENGINE WEB VIEWER SCENE, COMPARING THE DISTRIBUTIONS OF TWO DIFFERENT FISH SPECIES (GRADUATED SPHERE SYMBOLS) THROUGH A TOP/BOTTOM SPLIT SCREEN. OTHER 3D FEATURES SUCH AS SFM MODELS ARE INCLUDED FOR CONTEXT.

scanners to the tablets and the total station data collectors, which meant that barcodes were frequently entered by hand, leading to data entry error. The other issue was with missing data – data that simply did not get recorded. Both of these errors would be possible (and likely) to occur in a traditional excavation scenario – increased time available for excavation would likely alleviate some of these issues, but at great additional cost. Given the amount of data collected overall, the errors were negligible, and did not significantly impact any of the analyses.

The quantity of data collected means that there are almost unlimited number of different analyses that can be performed; while this project only permits exploration of a few of these avenues, the data collected is in a relatively simple and sharable format that can be used by researchers in the future, particularly as software for analysis of 3-dimensional data continues to improve and expand.

Acknowledgments

The authors would like to acknowledge the following groups and individuals:

The Metlakatla and Lax Kw'alaams (Coast Tsimshian) First Nations, the aboriginal communities with whom we worked; Canadian National Railway Company, the project sponsor; and Rachel Opitz (University of Arkansas), for information and advice on the use of SFM photogrammetry techniques.

References

CARAHER, W. (2013). iPads in the Field and Reflections on Archaeology's Digital Future. *The ASOR Blog* [Online]. Available from: http://asorblog.org/?p=4614 [Accessed July 22, 2014].

DIBBLE, H. L., MAREAN, C. W. & MCPHERRON, S. P. (2007). On the Use of Barcodes in Excavation Projects with Examples from Mossel Bay (South Africa) and Roc de Marsal (France). *The SAA Archaeological Record* 7**,** pp. 33-38.

E'SE'GET ARCHAEOLOGY PROJECT. (2012). Going Paperless. *E'se'get Archaeology Project* [Online]. Available from: http://coastalarchaeology.wordpress.com/2012/07/07/going-paperless/ [Accessed July 22, 2014].

GABII GOES DIGITAL. (2013). *21st c. Data, 21st c. Publications. Rich Digital Data and 3D Model Publication.* [Online]. Available: http://gabiiserver.adsroot.itcs.umich.edu/gabiigoesdigital/index.html [Accessed July 30 2014].

STEIN, J. K. (1992). *Deciphering a Shell Midden,* Toronto, Ont.: Academic Press Inc.

Integrated RTI Approaches for the Study of Painted Surfaces

Eleni KOTOULA, Graeme EARL

University of Southampton, Archaeological Computing Research Group (ACRG),
Southampton, UK
ek12v07@soton.ac.uk, graeme.earl@soton.ac.uk

Abstract

The present paper serves to show the potential of integrated Reflectance Transformation Imaging (RTI) approaches, namely microscopic, multispectral, trans illumination and trans irradiation RTI, with particular emphasis on painted surfaces. Classical-Hellenistic painted ceramic vases, egg-tempera paintings on wood and oil paintings on canvas were visualised using the proposed integrated RTI methodologies. Results indicate that the advanced visualizations obtained broaden our understanding about artefacts, assisting condition reporting, investigation, documentation and artefacts studies. In particular, microscopic RTI catalogues the surface topography at a microscopic scale and shed light on minor details of West Slope ceramic fragments, infrared RTI gave an insight into the red-figure vase painting technique and the hidden features of painted panel paintings and transmitted RTI enabled an advanced visualization of paint application and assisted condition reporting.

Keywords*:* R.T.I., Microscopic-R.T.I., I.R.-R.T.I., Trans-R.T.I., Painted surfaces

Introduction

RTI is a useful, non-destructive, affordable and easy imaging technique (Mudge *et al.,* 2005). Previous work has already proved that RTI and the most common fitting algorithm, Polynomial Texture Mapping (PTM), developed in 2001 at Hewlett Packard Laboratories (Malzbender *et al.,* 2001), significantly contribute to analysis, conservation and representation (Earl *et al.,* 2010). More details about the RTI workflow, including open source software and guides for data capture (Cultural Heritage Imaging, 2013a) processing (Cultural Heritage Imaging, 2011) and viewing (Cultural Heritage Imaging, 2013b) are available from Cultural Heritage Imaging website. Other useful resources can be found on the Hewlett-Packard website (Hewlett Packard, 2009) and the recent RTI guide published by English Heritage (Duffy, 2013).

Although RTI has found applications in a broad range of materials and artefacts types, for the purposes of the present paper studies focused on painted artefacts, including works of art and antiquities will be discussed. One of the earlier applications of PTM was that of the National Gallery. PTM provided detailed visual representations of the paintings and mock ups used as case studies and enabled researchers to detect and display features related to materials and techniques employed by the artists and damage (Padfield *et al.,* 2005). Researchers at the UniversitatPolitécnica de Valencia, Spain, revisited recently the same research questions with emphasis on the identification of manufacture techniques, condition assessment, conservation monitoring and evaluation of conservation materials, emphasising the potential of RTI in the field of paintings conservation (Manrique Tamayo *et al.,* 2013). Numerous paintings have been visualised in RTI form, including Henry O. Tanner paintings (Baade *et al.,* 2012), Lovis Corinth (Klausmeyer 2005), DomenicoBeccafumi (Castriota & Serotta 2011) and El Greco (Bakirtzis& Georgiou, 2014).

Moreover, RTI sheds light on the relief lines and dots, characteristic feature of the red figure technique, and enabled researchers of the Worcester Art Museum to reach conclusions about the tools used during the decorative process (Artal-Isbrand *et al.,* 2011). Notable is also the visualization of the head of a Roman statue of a young woman using a dome-based RTI system developed at the University of Southampton. Researchers reported the excellent surface detail and the information on the application and layering of paint obtained (Beale *et al.,* 2013). Moreover, RTIs of painted ceramics at the Fitzwilliam museum, demonstrated that capture of colour and surface properties even in cases of minor surface relief is possible (Bridgman & Earl, 2012).

1. Integrated RTI methodologies

One of the most significant components encouraging the development and adoption of RTI has been the flexibility of its methodology. Different imaging set-ups can be implemented presenting new possibilities for research, which are also able to benefit from a broad related literature on image processing, computer vision and graphics. The integrated RTI methodologies proposed will be presented in the following sections.

1.1. Microscopic RTI

Microscopic RTI combines the advantages of close up photography and photomicrography with those of RTI. In that way, it catalogues the shape and topography of the various components of artefacts at a microscopic scale. Microscopic RTI can be captured either with the

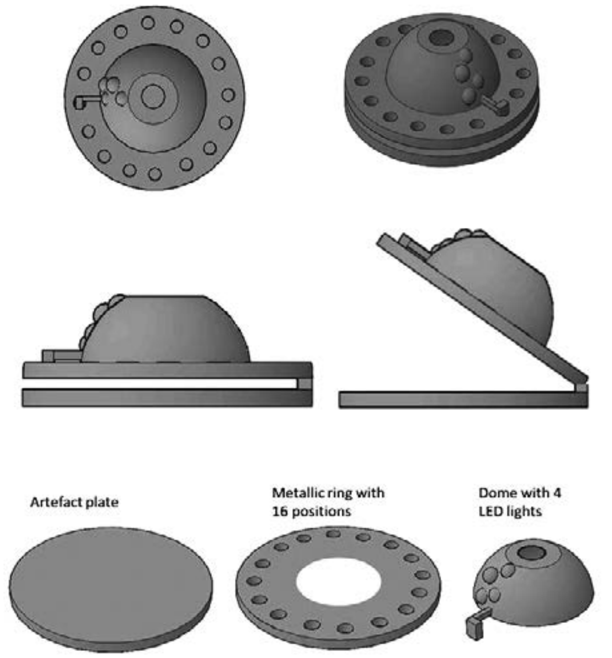


FIGURE 1: GRAPHICAL REPRESENTATION OF THE MINI-DOME AND ITS ASSEMBLY. THE DOME, EQUIPPED WITH FOUR LED LIGHTS IS JOINED WITH THE METALLIC RING AND MOVING CLOCKWISE AROUND ITS 16 POSITIONS IN ORDER TO CAPTURE 64 MICROSCOPIC IMAGES. THE DOME CAN BE TURNED 90 DEGREES, IN ORDER TO PLACE THE OBJECT SECURELY ON THE ARTEFACT PLATE. DIMENSIONS: PLATE'S DIAMETER 19CM, DOME'S DIAMETER 10CM, MAX HEIGHT 8,5CM.

Highlight RTI method (Kotoula, 2013) or with a mini-dome (Kotoula, 2012). For this paper microscopic RTI was applied to fragments of west slope ceramic vessels using a dome (Figure 1).

1.2. Infrared RTI

The experimentation for the simultaneous capture of spectral and textural data by Hewlett-Packard and Cultural Heritage Imaging (CHI) researchers reached the conclusion that the technique is feasible (Redman, 2007; Redman & Mudge, 2007). CHI has been experimenting with multispectral RTI referring to 'the hidden topological landscapes disclosing under-painting and drawings in the infra-red and the fine surface information disclosed in ultra-violet wavelengths' (Castriota & Serotta 2011; Schroer & Mudge, 2012). IR and UV RTI have been applied to an oil painting, as appears in the RTIiCAN blog, with the Osiris IRR camera/scanner (Gabov, 2010).

A combined RTI and infrared methodology was used for the study of ancient Dead Sea Scrolls, in order to reveal the text and examine the condition of the parchment (Caine &Magen, 2011).This project demonstrated a promising methodological approach for the study of documentary artefacts but also the necessity for further experimentation. The motivation for testing the IR-RTI on painted vases was the successful application of static infrared imaging undertaken previously. For example in the case of the pseudo-cypriote Amphora from Derveni the application of IR imaging enabled better understanding of the painted pattern (Figure 2).

FIGURE 2: PSEUDO-CYPRIOTE AMPHORA FROM DERVENI TOMBS (HEIGHT 0.281 M). DIGITAL IMAGE (LEFT), DESATURATED IMAGE (MIDDLE) AND IR IMAGE (RIGHT). IN THE IR IMAGE THE PAINTED PATTERN OF THE AMPHORA IS EMPHASIZED AND THE LEAVES ARE DIFFERENTIATED FROM THE BACKGROUND. THE BLACK AND BROWN PARALLEL LINES APPEAR MORE VIVID, WHILE THE REDDISH AND THE WHITISH ONES FADE OUT.

In the present study the IR-RTI data set is comprised of a series of infrared images, captured with a UV-VIS-IR modified DSLR camera and a range of filters (720nm, 760nm, 850nm, 950nm). The subjects were irradiated by IR light emitting diodes (LED) from different angles. This methodology was applied to red figure pelikes (storage jars) from the Derveni Tombs (Themelis & Touratsoglou 1997) as well as to icons (painted wood panels). The IR-RTI datasets were captured with an RTI arm inspired by the Quadri Flash Mantis system (Webb &Wachowiak, 2011), but presenting more possibilities for advanced data capture of medium size objects, proved to be helpful. A cylindrical base provides a secure foundation for a vertical axis, where the arm relies on. The arm can follow a vertical and clockwise movement, assisting data capture in that way. The infrared light can be securely placed on the arm's additional extension so as to be positioned in different points of the arm having various orientations (Figure 3). Conventional RTI provides an enhanced view of the artefacts under examination. The enhancement was significantly improved by the subsequent implementation of IR-RTI, where the three dimensionality features of the inner layer are highlighted. While these characteristics appear partly hidden and obscured under visible light, they are emphasised in IR-RTI. This provides the opportunity not only for enhanced condition reporting and documentation, but also examination of painting techniques and stylistic comparison of artefacts.

1.3. Transmitted RTI

Transmitted imaging (TI) is an imaging technique which has received only limited attention in the area of cultural heritage imaging. TI, proposed for the first time in the 1970s and further developed during the 1980s is considered as standard photographic techniques for the examination of artworks (Cornelius, 1977; Kushel, 1985). In TI the camera records the radiation transmitted through the investigated surface. Trans illumination and trans irradiation are useful for the examination of the painting history of such works of art, and also for their condition reporting. A notable value of the TI approach is its efficacy

Figure 3: The RTI arm in use.

in providing evidence for craqueleure. Moreover, TI differentiates successfully the painting techniques applied as well as the variable thickness of the painted layers in different areas of paintings (Cucci *et al.*,, 2012).

In our study imaging of a series of canvas painting testers revealed the potential of each technique (Figure 4). In turn, transmitted RTI was tested in an attempt to provide an enhanced RTI visualization, complementary to reflected visible and IR-RTI. Considering the efficiency of transmitted imaging in the examination of canvas paintings, two such objects were used as a case study for RTI experimentation. The set-ups tested were reflected and transmitted RTI in the visible (trans illumination RTI) and IR spectral region (trans irradiation RTI).

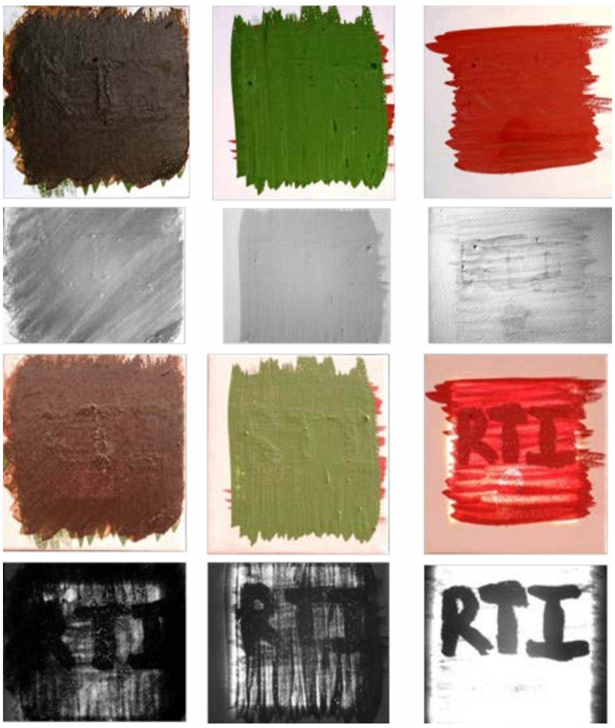


Figure 4: Canvas testers 1 (left), 2 (middle), 3 (right). Digital images (top row), reflected IR images (second row), trans illumination images (third row) and transmitted IR images (forth row). Stratigraphy of testers: Tester 3, first layer writings with carbon black, second layer cadmium red pigment, tester 2, the same as tester 3 plus a cement green pigment layer, tester 1, the same as tester 2 plus a raw ombre pigment layer.

Three different data capture strategies were tested: transmitted-dome-RTI, highlight transmitted RTI, and a two-camera methodology. The transmitted-dome-RTI methodology requires the use of a modified dome. A normal dome positioned upside down was used for testing. The only modification was the addition of a stand so as to support the dome in an upside down position.

The successful capture of highlight transmitted RTI datasets requires pre-capturing of the lighting positions as well as the accurate marking of these positions, in order to be used during data capture and processing of the transmitted RTI data. This process is very time-consuming and mismarking the lighting positions is very easy to occur. One of the techniques tested for transmitted-H-RTI employs two cameras. One camera is focused on the one side of the object. The second one is placed on the opposite side and is focused on the shiny sphere beside the object. The latter acquires the necessary data for the construction of the light position file, while the former captures the object under examination in transmitted illumination. An alternative option is the use of a light arm in predefined lighting positions. The advantage of the two cameras method is that the lighting positions and the transmitted RTI dataset are captured simultaneously, minimising the possibility for mismarking the lighting positions. Nevertheless, the technique is impractical in particular for trans irradiation applications, where not only two cameras but also two different light sources are necessary. At least two people (if not three) need to be employed.

The dome method is recommended for transmitted data capture due to the ease of use, the considerable less time needed for data capture and processing and moreover the quality of data captured. In cases where the use of a dome for data capture is impossible, a light arm can be of great assistance and proved to be a useful additional tool. There is no considerable delay in data capture and processing. On the contrary the two cameras methodology for transmitted H-RTI data capture is the one less preferred.

2. Results and discussion

2.1. Microscopic RTI case study: West Slope wares

West Slope wares, named after the large number of over-painted pottery found on the West Slope of the Acropolis, is one of the most important sequences of vases in the Hellenistic period, produced in many places around the Mediterranean including Athens, Corinth, Macedonia, Crete, Egypt and Pergamon. West slope wares originate from the gilded black glaze vases of the later fourth century, but instead of gilding the decoration consists of added yellow and white colour and incisions on a black glazed background (Williams 1999). The colour, provided

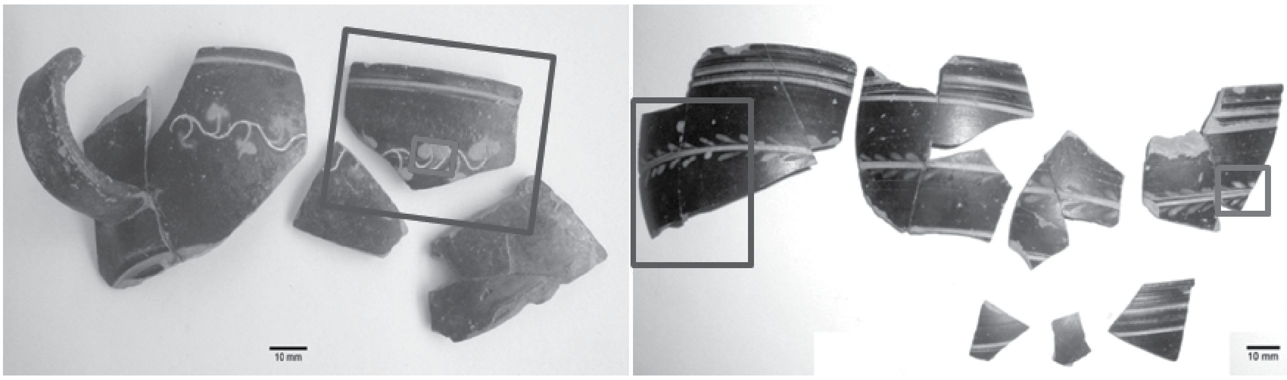


Figure 5: Fragments of the skyphos C3e and cylix C3th II. The rectangles demonstrate the areas for macro and microscopic RTI captures.

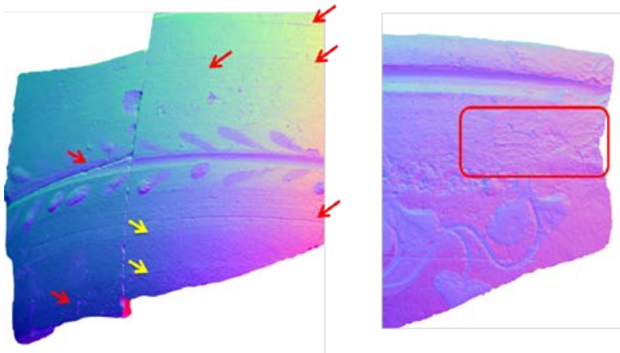


Figure 6: Macroscopic H-RTI visualizations of West Slope ceramic fragments from Derveni, fragment from skyphos (3cm across) with emphasised incisions (red arrows) and adjacent concavities (yellow arrows) (left) and fragment from kylix (2cm across) with emphasized a surface anomaly under the rim, possibly created by the fingers of the potter or by a sponge (right).

by dilute solutions of two different kinds of clay, was added before firing.

Rotroff (1991) discusses the problems in dating West Slope wares: shape and context do not provide enough data, while the study of the decoration and its development assists in refining chronology. In such an approach the poor state of preservation, not only due to common degradation but also because of the nature of West Slope painting, is problematic. The flaking of the painted designs and sometimes of the entire decoration is a usual common phenomenon. Usually the only evidence found of the decoration are remains which appear as a ghost on the surface of the glaze.

For this study two West Slope vessels in fragmentary state, a skyphos (C3e) and a kylix (C3th II) from Derveni Tombs (Themelis & Touratsoglou 1997) currently located at the Archaeological Museum of Thessaloniki in Greece, were visualised in RTI form. The H-RTI method was employed at macroscopic scale, and the dome method used at a microscopic level (Figure 5). The state of preservation of the paint decoration was bad, although the pattern was still recognizable. Even in cases of total loss, colour remained as a ghost on the surface of the glaze.

RTI offered an insight into the West Slope decoration technique and enables the advanced study of relief detail by emphasizing even on minor surface variations. Hardly recognizable remains of colour, areas where colour had been either totally or partially lost were highlighted. Moreover, in the case of yellow-tan details the relatively

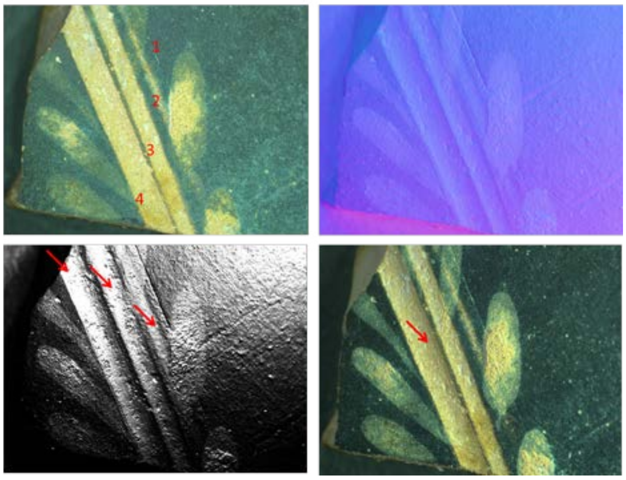


Figure 7: Fragment of the skyphos C3e (0.8cm across), clockwise from top left, microphotograph, normal map, and RTI visualizations.

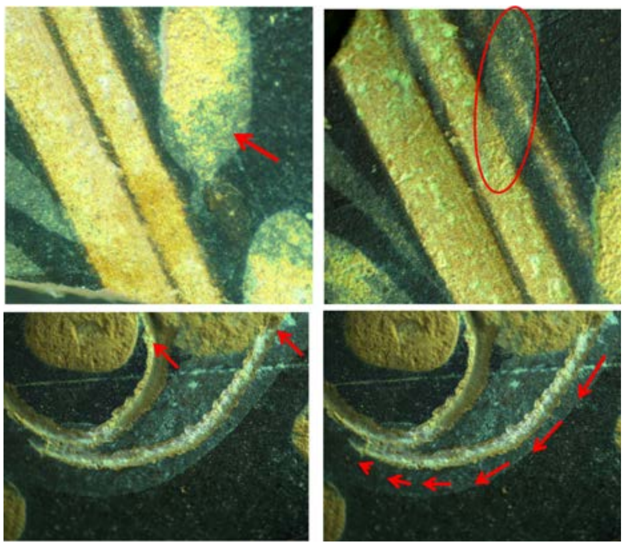


Figure 8: Fragments of the skyphos C3e (top) and cylix C3th II (bottom), microscopic RTI visualizations.

Figure 9: The red figure pelikes H2, height 0.030 m. (left) and C1, height 0.255 m. (right)

thick colour used in comparison to the white became obvious in the RTI visualizations, providing evidence for Cook's explanation about Greek vase painting (Cook 1997).

A few potters' tools are known, mainly the scraper, shaper, severing and measuring tools (Toby 1999).

Via RTI the shape and morphology of the different textures created by different tools and tool marks can be studied, providing evidence for the vase construction and decoration techniques used. Macroscopic RTI visualizations in our study revealed a number of incisions and emphasized adjacent concavities on the lower part of the body of the skyphos (black arrows). In the kylix RTI emphasized a surface anomaly under the rim, possibly created by the fingers of the potter or by a sponge (Figure 6).

Microscopic RTI enables the study of the shape of these incisions and can help defining tool marks and methodology. For example, the microphotograph revealed the presence of 4 lines in the central motif of the skyphos, whilst microscopic RTI visualizations enable the characterisation and further study of these lines. All of them were shown to be incised, with the three lower ones also painted. The salt efflorescence, the small pits and the remnants of colour from the leaves on the surface complicate the study of the tool marks in the two lower incisions. There is no doubt however that they were made with circular scrapers, possibly for the lower one a large tool was used. Both of them were rigid, considering that the width of the incisions is the same. From the horizontal lines on the inside of the incisions one can assume that the tools' surfaces were not very smooth. The upper incision is made with the edge of a sharp rigid scraper, while the second is larger and must have been made with a similar tool, but with stronger incision by the potter's hand (Figure 7).

The triangular tool used for the incisions of the kylix is easily detectable in the microscopic RTI visualizations and the orientation of the potter's hand direction can be identified. In both vessels the yellow leaves additions were the last step, usually covering other details of the decoration (Figure 8). Application of RTI and micro-RTI indicated that RTI is an excellent tool for the revelation of incised and painted decoration of West Slope wares, and also assisted in condition assessment. Scratches on the surface, flaking of the glaze and other alterations and damage were easily recognizable.

2.2. IR-RTI case studies

2.2.1. Red figure vases

RTI proved a valuable tool for the examination and documentation of the red figured technique vase painting technique, based on the observations of two pelikes - storage jars (C1, H2) from the Derveni Tombs, conserved from a number of fragments and restored with gypsum. The pelike H2 is attributed to the Painter of the Wedding Procession and dated to the last quarter of the 4th century BC. The detail depicts a young Goddess Aphrodite, seated, looking left at Eros. At the second level, there are two other woman standing looking right. At the right end the scene a Silenus is performing on a double aulos. Beneath the handles, the characteristic flower decoration motif is preserved (Themelis & Touratsoglou 1997). The surface presents discolouration, stain, salt efflorescence, colour loss, gaps and pits.

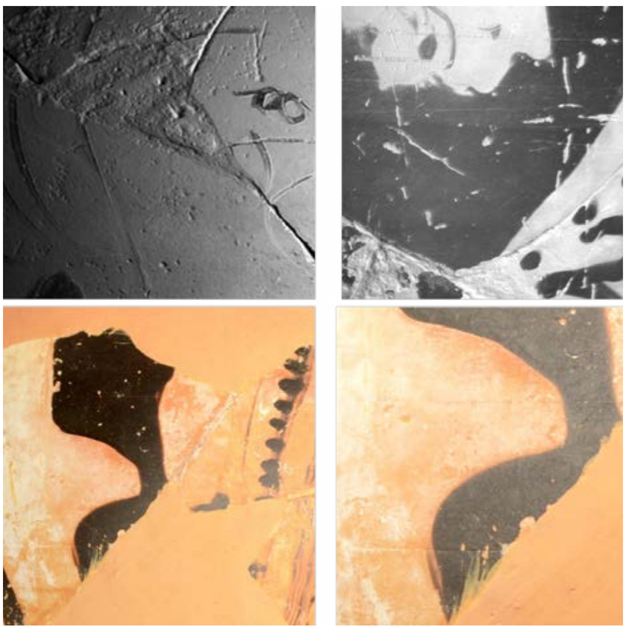


Figure 10: Red figure pelikes. Detail revealing surface topography and damage.

Figure 11: Red figure pelike H2. Detail, clockwise from top left, digital image and RTI visualizations emphasizing incisions developed during construction possible from the ceramic wheel and revealing the preparatory drawing

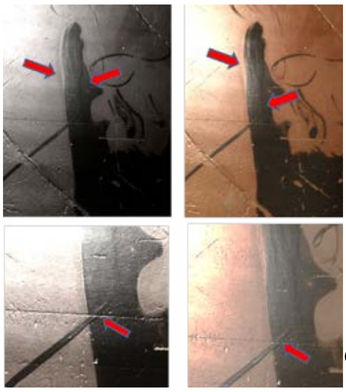


Figure 12: Red figure pelike, C1. Detail, RTI visualizations revealing the outline and black glaze brush strokes (top) and a relief line passing over the black background (below).

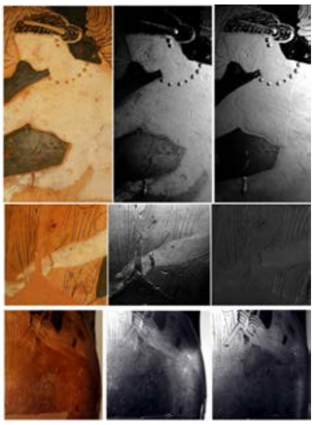


Figure 13: Details, red figure pelike H2. Digital images (left) and RTI visualizations (middle) and IR-RTI visualizations (below).

Figure 14: Icon, details, digital images (left) and RTI visualizations (right) emphasizing surface

The pelike C1, attributed to the group G, depicts a female with tiara, looking right, standing between a horse and a griffin. In front of the horse, there is a large acanthus leaf over a helix. The head of the figure (Arimasp or Amazon) and the griffin are painted white. On the other side two young men, wearing himation, are standing next to an altar (Themelis & Touratsoglou, 1997). The scene is indiscernible, because of the overall poor state of conservation, the large restored areas and the highly deteriorated surface with a large amount of colour loss (Figure 9).

In the red-figure style, decoration was orange-red and outlined in black, while the background outside was filled in with black. After forming the shape of the vase, following the known process of throwing, turning and joining, when the clay was relatively hard but not completely dry, the surface was rubbed in order to create a smooth, shiny surface, less susceptible to abrasion (Toby 1999). Then, a preliminary sketch was designed on the hard surface of the vase, at least including the basic elements of the decoration. Afterwards this was covered by the gloss. Preliminary sketches were essential for the red-figure technique because the lines define the forms. Harder impressions may leave traces on the surface of the vase after firing. Those lines are more recognizable in the reserved areas, but occasionally they extend into the black background, presenting a shininess which indicates that they were made before the final black glazing (Corbett 1965). The outlines were painted after the completion of the preliminary design, and played an important role in the painting process.

One of the most characteristic elements of the red-figure technique is the relief lines, which were used to outline figure and details. These were painted with thick clay slip and therefore resulted in visible raised lines. Apart from the relief lines, raised dots were used for details, such as hairs, ornaments on garments and other decoration elements. Also, worth mentioning features of the red-figure technique are the dilute glosses. Then the areas intended to be black was painted with clay mixed with alkaline water in order to produce the glossy black surface during the firing. Finally the firing followed a three step process of oxidising, reducing and re-oxidising atmospheres in the kiln, resulted in the known lustrous surface of the vases with the red reserved areas (Toby 1999). These characteristic features of red-figure vases are not easily detected using traditional documentation and recording methodologies. Cook noticed this deficiency and in particular the fact that the relief lines are not evident in usual archaeological reproductions (Cook, 1997). Corbett clearly states that 'it is impossible to light the vase so that all parts of the sketch show clearly at the same time' (1965). Specular and raking light photographing techniques have been proposed in order to successfully represent the surface variations (Clark, 2002; Corbett, 1965).

The most important conclusions from the RTI visualization of the red-figure vases were the following:

- The texture is underlined, including features of extreme importance, such as surface damage losses, scratches, pits, alterations, discontinuities of the paint and the glaze, craqueleure and depositions (Figure 10).

- Manufacture evidence can be investigated, such as incisions from the ceramic wheel and lines of the preparatory drawing (Figure 11)
- Painting techniques can be examined, such as outlines, relief lines, sketch lines, raised dots, and remnants of the black glaze (Figure 12).

RTI provides an enhanced view of the artefacts under examination, which was significantly improved by the implementation of IR-RTI, where the three dimensionality of the features of the inner layer are highlighted. While these characteristics appear partly hidden and obscured under visible light, they are emphasised in IR-RTI. The relief lines were significantly highlighted, providing an enhanced view of the preparatory drawing, as well as the three dimensionality of the painted pattern. Considering that they were used to outline figure and details, IR-RTI views of the painted scene are particularly informative. The outline is more differentiated in the IR-RTI renderings. Least but not last, restored areas appear highlighted, providing evidence for previous repairs. Hence, IR-RTI provides the opportunity not only for enhanced condition reporting and documentation, but also examination of painting techniques and stylistic comparison of artefacts (Figure 13).

2.2.2. Icons

An icon (or ikon from the Greek eikōn) is a representation of a sacred or sanctified Christian personage used in religious worship in the Russian or Greek Orthodox Church. The production of icons has been described by the monk Dionisiosek Fourna in 1728-1733. Icons are typically painted on a wooden panel using the egg tempera painting technique, over a preparation layer made of gypsum and glue. Sometimes, canvas is present between the wood and the preparation layer. The painter with the aid of a sharp tool or with diluted tempera colour either incised or painted the preparatory drawing. Afterwards the golden leaf was attached. The brush work initiates with a brownish colour (proplasmos), over which softer tones were applied (sarcoma) darker for shadows or lighter and reddish for highlights. The last step is the varnishing of the icon (Panselinou, 2010).

The application of RTI in icons is a prosperous area of research because icons are artworks of great historical and artistic value, well distributed around the world, which present interesting surface topography, difficult to capture and examine using conventional methods, due to their planar shape and their reduced three-dimensional characteristics, along with the gold or silver gilded areas. Notable is also that RTI can be an excellent addition to the icons examination protocol, which has been the matter of thorough research, either destructively, using analytical methods (Raman spectroscopy, SEM-EDX, GC-MS) or preferably non-destructively (optical microscopy, multi-spectral imaging, X-radiography, computer aided tomography, XRF).

FIGURE 15: ICON, DETAIL, RTI VISUALIZATIONS (LEFT) AND IR-RTI VISUALIZATIONS (RIGHT) REVEALING THE TEXTURE OF THE SUBSTRATE AND THE DEFORMATIONS OF THE PANEL.

FIGURE 16: ICON, DETAIL, RTI VISUALIZATION (LEFT) AND IR-RTI VISUALIZATION (RIGHT) EMPHASIZING LINES OF THE PREPARATORY DRAWING.

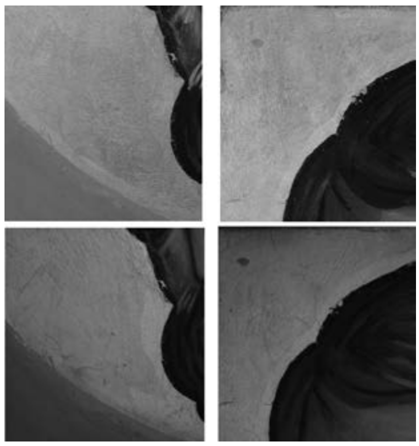


FIGURE 17: ICON, DETAIL, RTI VISUALIZATIONS (TOP) AND IR-RTI VISUALIZATIONS (BELOW) HIGHLIGHTING CRACKS FROM THE APPLICATION OF GOLD LEAF.

FIGURE 18: ICON, COMPARISON OF NORMAL MAPS, RTI (LEFT) AND IR-RTI (RIGHT).

Nevertheless icons consist of an artefact type, which has received minimum imaging attention, particularly in terms of RTI. Zányi *et al.,* underlined the inefficiency of photography or active 3D range scanning in capturing gold, silver leafs and tempera painting technique of icons. Their case study, of the apse mosaic at the Angeloktisti Church at Kiti, Cyprus, although did not include experimentation on icons, highlighted the efficiency of RTI in capturing gold, silver and glass tesserae, providing an advanced perception and enabling explorations of the various different views of the mosaic (Zányi *et al.,* 2007).

We undertook RTI visualization of modern replica icons, created using the traditional materials and techniques of byzantine egg-tempera painting, emphasises surface anomalies of low relief details. Characteristic features, such as brush strokes can be detected and areas with discontinuities can be easily documented. Thorough RTI examination of the surface topography reveals signs, which provide an insight into manufacture, such as gold gilding technique (Figure 14).

The IR-RTI of the icon, taking into consideration their planar shape and their reduced three-dimensional characteristics, gave interesting results, revealing hidden features and the preparatory drawing. IR-RTI by penetrating the varnish layer, the texture of the substrate can be observed. RTI reveals the texture of the varnish applied while in the IR-RTI renderings the texture of the substrate is visible, along with its deformations, such as the vertical crack on the area of the basket (Figure 15). The preparatory lines along with their texture can be examined. Final brushing of details disappears, emphasizing the three dimensional properties of the preparatory drawing (Figure 16). Areas where the colours have not been applied uniformly due to surface variation of the substrate become obvious, as well as crack in the application of the gold leaf (Figure 17). The best means of comparison between normal and IR RTI is the normal maps in the visible and infrared spectral (Figure 18).

This case study serves to show the potential of RTI in the study of painted panels. Apart from the interesting conclusions regarding methodological issues, the difference between authentic material and replicas should be noted. Alterations and damage present in the authentic artworks results in an advanced level of complexity in the study of icons. Yellowed varnishes, depositions, losses and cracks have a tremendous effect in the appearance of the artwork. It is expected that the application of the proposed methods in the study of original artworks would maximise the information obtained.

2.3. Transmitted RTI case study: Canvas paintings and testers

In trans illumination RTI the crucial parameter that governs the results is the thickness and the opacity of the painting layer. Its reflectance properties such as the gloss, affect the results but at a lesser extent. The direction, the shape, the thickness of the painting strokes can be observed. Moreover, conclusions about the materials (tools, mediums, pigments) and the painting techniques can be reached. As the surface is virtually relighted in transmitted mode the areas with the larger quantities of colour applied and the stronger strokes appear not to be differentiated. On the contrary the areas which present a certain degree of translucency appear more or less lit, in accordance to the direction and the location of the light beneath the surface under examination. In that way the visual result is differentiated in each trans illumination RTI snapshot and the viewer experiences the transmittance properties of the surface in a way superior to transmitted imaging, mainly due to the perception of three dimensionality of the painting (Figure 19). This feature can be characterised as a pseudo-three dimensional representation by transmitted RTI: the texture revealed does not relate to the surface, but is the result of the virtual relighting of the translucent areas of the painting. This pseudo-three dimensionality of transmitted RTI should not be confused with the ability of normal RTI to emphasize textures. Normal maps derived from TI-RTI do not represent the texture of the object as in reflected RTI.

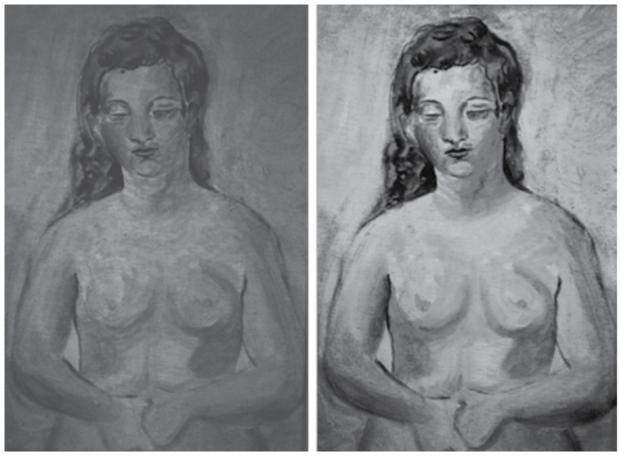


Figure 19: Canvas painting, transmitted digital image (left) and transmitted RTI visualization (right)

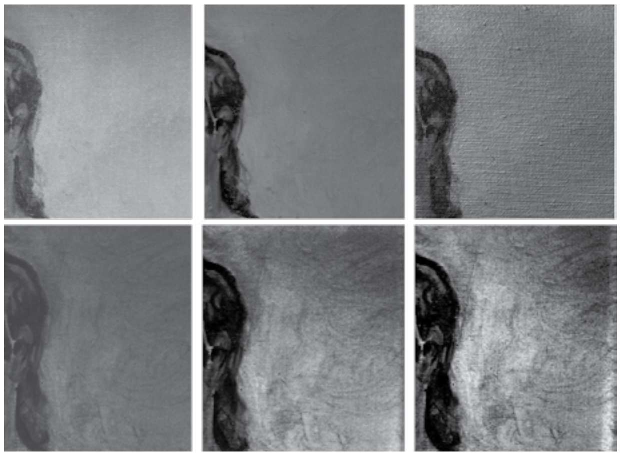


Figure 20: Detail, comparison of reflected and transmitted light digital images (left) and RTI visualizations (middle, right). Digital reflected image (up, left) and normal RTI visualizations (up, middle and right). Transmitted digital image (below, left) and transmitted RTI visualizations (bellow, middle and right).

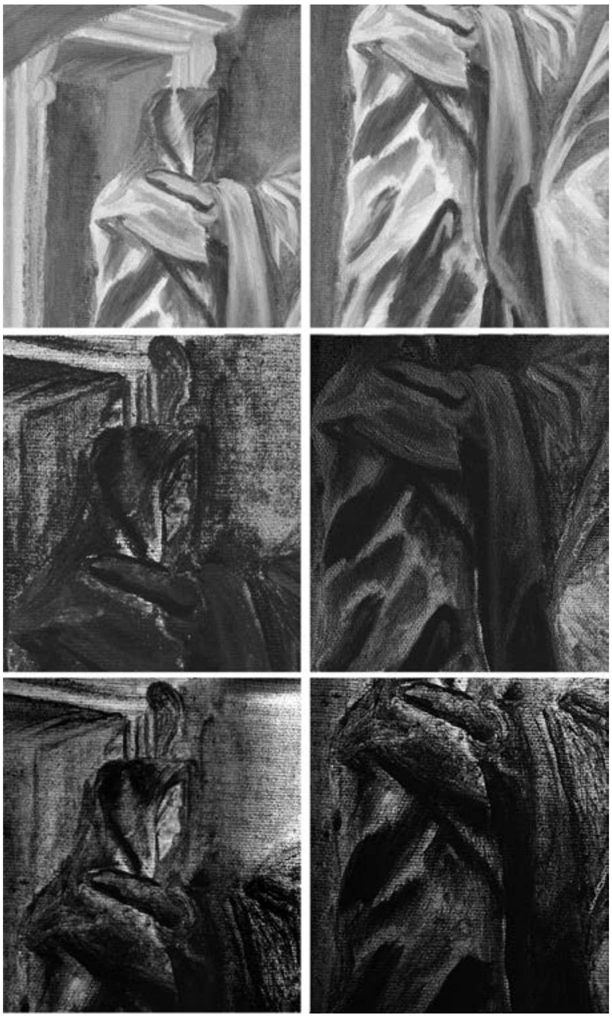


Figure 21: Detail, normal RTI (top) transmitted RTI (middle) and transmitted IR- RTI (below) visualization.

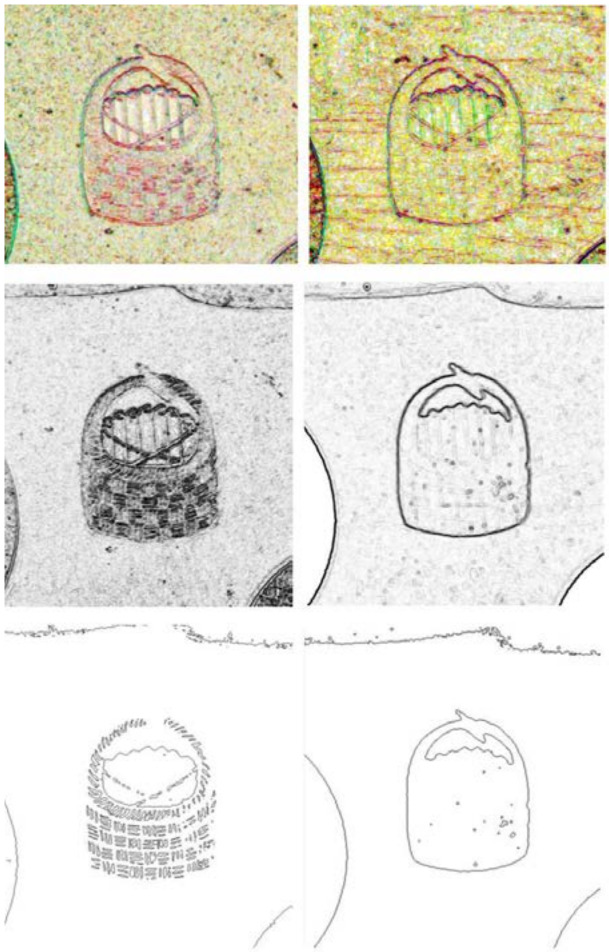


Figure 23: Icon, visible (left) and IR (right). Applying find edge and trace contours filters to normal maps and RTI visualizations

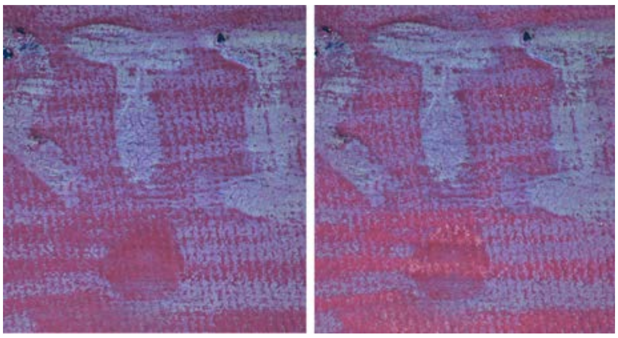


Figure 22: Canvas painting tester, comparison of reflective RTI visualizations captured without (left) and with (right) backlighting.

The comparison of the visualizations produced from the different RTI methodologies reveals the different features of the painting emphasized. TI-RTI techniques enable a thorough examination of painting strokes and surface deformations assisting condition reporting and conservation documentation. TI-RTI enables the advanced perception and the in depth examination of the application of the colour, providing evidence for the painting history of the artwork. TI-IR-RTI emphasizes the same features as the trans-illumination but furthermore is enriched with the penetration abilities of IR radiation. Hence, it is of great value in terms of diagnostic examination (Figure 20-21).

Capturing normal RTIs of canvas paintings placed on light boxes is another modification of the basic technique that can be used as complimentary methodology (Figure 22).

3. Digital Image processing of RTI visualizations

Filtering RTI visualizations using standard digital imaging filters provides interesting results, which can be useful for illustration and interpretation purposes. This on-going work has already demonstrated that moving from RTI outputs, including normal, to further enhanced imagery is a fruitful process. Furthermore it is clear that existing image processing frameworks and applications provide considerable potential. In particular, filtering RTI visualizations in the visible and infrared spectral area assists the comparative analysis of the artwork. For example the texture of the substrate in the IR RTI visualization after applying find edge and trace contours filters is emphasized by the horizontal lines as opposed to the smooth texture in the visible visualization (Figure 23). Using the same methodology different visualizations can be produced. The subtractive illustration approach may be useful for interpretation and explanation of the results of multispectral RTI. But at the same time the filtering of the RTI visualizations may result in confusing illustrations. Further experimentation is required so as lead to safe conclusion regarding the limitations and usefulness of this processing.

Conclusions

This paper has presented microscopic, multispectral and transmitted RTI methodologies. It has evaluated their application, with a particular emphasis on their potential for the study of painted surfaces, including painted ceramics and paintings on wood and canvas. The results of the proposed methods were promising. Microscopic RTI sheds light on minor details and is particularly useful for diagnostic examination. IR-RTI provides evidence about the texture of surfaces beyond the visible and proved to be a useful tool for the study of painted panels, canvas painting and painted ceramics. It enabled an advanced perception of IR information because it introduces the three-dimensionality, which is absent in infrared images. Multispectral RTI is a technique with great potential for a broad area of materials and artefacts types, including painted surfaces, such as pained panels, canvas, documentary artefacts, engraved stones in poor conservation condition, restored glass artefacts, painted or engraved ceramics with encrustations, mortars, paper and archival material. Trans illumination and trans irradiation RTI offers an insight into painting techniques and assisted the examination of canvas paintings.

Transmitted RTI's main potential is to document structure of materials (paper, canvas), their condition (repairs, abrasions, losses) and painting techniques and to reveal texture of hidden features, obscured by linings, mounts or inks and pigments, such as under drawings, inscriptions and watermarks, assisting in that way in the perception of three-dimensionality, an issue of great importance for the characterization of materials, diagnostic examination and visual analysis. Materials and artefact types that can primarily benefit are those which present a certain degree of translucency, such as canvas paintings, mounded papyri, photographic material, and works of art on paper or archival material.

The following list presents the basic points of interest and summarizes the contribution of integrated RTI methodologies in the examination of painted surfaces.

- RTI captures the texture of the surface and emphasizes the three dimensionality of the paint layer. When thin paint layers have been applied the texture of the substrate and the preparatory drawing can be revealed. It is particularly useful in visual analysis and condition reporting for the detection of deformations and other imperfections due to chemical, physical or biological damage.
- IR-RTI captures the texture beneath the visible due to the penetration ability of IR and emphasizes the three dimensionality of the paint layer according to the response of the pigments in IR irradiation. IR-RTI reveals preparatory drawings and other hidden features. It is useful in investigation, diagnostic examination and condition reporting for the detection of deformations and other imperfections in the layer beneath the surface.
- TI-RTI captures the texture of the substrate and the variability of the paint layer in case of colour applied to translucent substrates. It reveals painting techniques and emphasizes areas of an increased translucency. When thin paint layers have been applied the capture of the preparatory drawing is possible. It is useful in stylistic comparison of painted surfaces and in condition reporting for the detection of deformations and other imperfections, mainly mechanical shocks and physical damage, such as cracks, craquelere etc.
- TI-IR-RTI captures the texture of the translucent substrate and the preparatory drawing clearly and emphasizes the paint layer in case of pigments transparent or translucent in IR irradiation. It is useful in art diagnostics due to the successful detection of deformations, imperfections and painted history of artworks. It is particularly effective in the detection of over painting.

References

ARTAL-ISBRAND, P., KLAUSMEYER, P. & MURRAY, W., (2011). An evaluation of decorative techniques on a red-figure Attic vase from the Worcester Art Museum using reflectance transformation imaging (RTI) and confocal microscopy with a special focus on the 'relief line.' In P. B. Vandiver *et al.,* (eds.). *Materials issues in art and archaeology IX: symposium held November 29-December 3, Boston, Massachusetts, USA.* Materials Research Society symposium proceedings, 1319.Warrendale: Materials Research Society, pp. 3–33.

BAADE, B., KERR-ALLISON, A. & GIACCAI, J., (2012). Pursuit of the ideal effect: the materials and techniques of Henry Ossawa Tanner. In A. O. Marley (ed.). *Henry Ossawa Tanner: modern spirit*. Philadelphia: Pennsylvania Academy of the Fine Arts, pp. 157–166.

BAKIRTZIS, N. & GEORGIOU, R., (2014). *RTI results on El Greco Paintings. The Cyprus Institute*. Available at: http://www.cyi.ac.cy/component/k2/item/766-rti-results-on-el-greco-paintings.html [Accessed June 15, 2014].

BEALE, G., PAGI, H. & EARL, G., (2013). Roman painted statue head from Herculaneum. In S. M. Duffy, (ed.) *Multi-light Imaging for Heritage Applications*. London: English Heritage, pp. 19–20.

BRIDGMAN, R. & EARL, G., (2012).Experiencing lustre: polynomial texture mapping of medieval pottery at the Fitzwilliam Museum. In R. Matthews, R. Curtis, J., Symour, M. *et al.,* (eds.). *Proceedings of the 7th International Congress of the Archaeology of the Ancient Near East (7th ICAANE) 12 April - 16 April*

2010, British Museum and UCL, London. Wiesbaden: Harrasowitz, pp. 497–512.

CAINE, M. & MAGEN, M., (2011). *Pixels and Parchment: The Application of RTI and Infrared Imaging to the Dead Sea Scrolls*. In EVA London.

CASTRIOTA, B. & SEROTTA, A., (2011). Reflectance Transformation Imaging Training at the Conservation Center. *Newsgram: News in brief from the Conservation Center*, 4, pp.1–2.

CLARK, A.J., ELSTON, M. & HART, M.L., (2002). *Understanding Greek vases: a guide to terms, styles, and techniques.* The J. Paul Getty Museum, ed., Hudson, Christopher.

COOK, R.M. (1997). *Greek painted pottery*, New York: Routledge.

CORBETT, P.E. (1965). Preliminary Sketch in Greek Vase-Painting. *The Journal of Hellenic Studies* 85, pp.16–28. Available at: http://www.jstor.org/stable/628804.

CORNELIUS, F.D. (1977). Correspondence: Transmitted infrared photography. *Studies in conservation* 22(1), pp. 42.

CUCCI, C., PICOLLO, M. & VERVAT, M., (2012.) Trans-illumination and trans-irradiation with digital cameras: Potentials and limits of two imaging techniques used for the diagnostic investigation of paintings. Journal of Cultural Heritage, 13(1), pp.83–88. Available at: http://linkinghub.elsevier.com/retrieve/pii/S1296207411000653 [Accessed December 17, 2012].

CULTURAL HERITAGE IMAGING (2013a.) *Guide to highlight image capture v2.0. Cultural Heritage Imaging,* pp.32. Available at: http://culturalheritageimaging.org [Accessed January 19, 2014].

CULTURAL HERITAGE IMAGING (2011). *Guide to highlight image processing. Cultural Heritage Imaging*. [Online] Available at: http://culturalheritageimaging.org/What_We_Offer/Downloads/rtibuilder/RTI_hlt_Processing_Guide_v14_beta.pdf [Accessed April 3, 2012].

CULTURAL HERITAGE IMAGING 2013b. *Guide to RTIViewer, version 1.1. Cultural Heritage Imaging.* Available at: http://culturalheritageimaging.org/What_We_Offer/Downloads/rtiviewer/RTIViewer_Guide_v1_1.pdf [Accessed July 2, 2014].

DUFFY, S.M. (2013). *Multi-light imaging for heritage applications*, London: English Heritage.

EARL, G., MARTINEZ, K. & MALZBENDER, T.,(2010). Archaeological applications of polynomial texture mapping: analysis, conservation and representation. *Journal of Archaeological Science,* 37(8), pp.2040–2050. Available at: http://linkinghub.elsevier.com/retrieve/pii/S0305440310001093 [Accessed January 8, 2013].

GABOV, A. (2010). *IRR RTI using the Osiris. RTIiCAN blog.* Available at: http://rtiican.wordpress.com/2010/01/10/irr-rti-using-osiris/.

HEWLETT PACKARD, (2009). Hewlett Packard. Polynomial texture mapping (PTM).Hewlett Packard Available at: http://www.hpl.hp.com/research/ptm/downloads/download.html.

KLAUSMEYER, P., (2005). Polynomial texture mapping and conserving a fire damaged oil painting by Lovis Corinth. In: *Proceedings of the 2005 Seminar on Surface Metrology at WPI*. Available from: http://www.me.wpi.edu/Research/SurfMet/Seminar/Oct2005/philipklausmeyer.html (last accessed 01.05.09.).

KOTOULA, E. (2012). Application of RTI in Museum Conservation. In Graeme E. *et al.,* (eds.) *Archaeology in the Digital Era. Papers from the 40th Annual Conference of Computer Applications and Quantitative Methods in Archaeology (CAA)*, Southampton, 26-29 March 2012. Amsterdam University Press, Amsterdam, pp.

KOTOULA, E. (2013). Microscopic RTI of gilded silver discs from the Derveni tombs, Macedonia, Greece. In: S. Duffy (ed.). *Multi-light Imaging for Heritage Applications*. London: English Heritage, pp. 21–23.

KUSHEL, D.A. (1985). Applications of transmitted infrared radiation to the examination of artifacts. S*tudies in conservation* 30(1), pp.1–10.

MALZBENDER, T., GELB, D. & WOLTERS, H. (2001).Polynomial texture maps. In L. Pocock (ed.) *Proceedings of the 28th annual conference on Computer graphics and interactive techniques*, New York: ACM Press. pp. 519–528.

MANRIQUE TAMAYO, S.N., VALCARCEL ANDRES, J. & JULIA, O.P. (2013). Applications of Reflectance Transformation Imaging for documentation and surface analysis in conservation. *International Journal of Conservation Science* 4, pp.535–548. Available at: http://www.ijcs.uaic.ro/public/IJCS-SI-Manrique.pdf.

MUDGE, M. *et al.,* (2005). Reflection Transformation Imaging and Virtual Representations of Coins from the Hospice of the Grand St . Bernard. In M.Mudge, N. Ryan, and R.Scopigno (eds.) VAST 2005. *The 6th International Symposium on Virtual Reality, Archaeology and Cultural Heritage*. November 2005, Pisa: Eurographics Association, pp. 195–202.

PADFIELD, J., SAUNDERS, D. & MALZBENDER, T. (2005). Polynomial Texture Mapping: A New Tool for Examining the Surface of Paintings. In I. Verger (ed.)ICOM Committee for Conservation. London: Earthscan Ltd.

PANSELINOU, N., (2010). *Byzantine painting*, Athens: Kastanioti.

REDMAN, J., (2007). Advances in Digital Imaging for Fine Art and Cultural Heritage. In *NIP23: International*

Conference on Digital Printing Technologies and Digital Fabrication, Ancjhorage 2007. Springfield: Society for Imaging Science and Technology, pp. 355–363.

REDMAN, J. & MUDGE, M., (2007).The simultaneous capture of spectral and textural information. In S. A. Stovall, (ed.). *Archiving 2007: final program and proceedings, May 21-24, 2007, Arlington, Virginia.* Springfield: Society for Imaging Science & Technology, pp. 2–5.

SCHROER, C. & MUDGE, M., (2012). Advances in the computational photography tools: Reflectance Transformation Imaging (RTI) and Algorithmic Rendering (AR). In *Computer Applications and Quantitative Methods in Archaeology-CAA 2012*. [On line] Available at: https://www.ocs.soton.ac.uk/index.php/CAA/2012/paper/view/633.

THEMELIS, P. & TOURATSOGLOU, I. (1997). *Tafoitou Derveniou*, Athens: TAPA.

TOBY, S. (1999). *Athenian vase construction: a potter's analysis, Santa Monica*, California: Getty Publications.

WEBB, E.K. & WACHOWIAK, M. (2011). *Imaging Studio Technical Note : QuadriFlash Mantis for Reflectance Transformation Imaging (RTI)*, pp.1–7.

WILLIAMS, D. (1999). The Hellenistic Period. In *Greek Vases*. London: The British Museum Press, pp. 131–137.

ZÁNYI, E. *et al.,* (2007). Lighting and byzantine glass tesserae. In J. Bowen, S. Keene, & L. MacDonald (eds.) *Electronic Visualisation and the Arts.* London: EVA Conferences International, pp. 1–8. Available at: http://www.eva-london.org/past-eva-londons/2007/published-proceedings.

Survey, Documentation and Analysis of the Archeological Architecture: the House of the Knights of Rhodes in the Forum of Augustus

Carlo BIANCHINI, Gaia Lisa TACCHI

Department of History, Drawing and Restoration of Architecture,
Sapienza - University of Rome, Rome, Italy
Carlo.bianchini@uniroma1.it, gaialisa.tacchi@uniroma1.it

Abstract

The study and interpretation of the so-called Archeological Architecture is generally quite a difficult task: such elements or complexes nowadays show in fact very stratified and heavily restored configurations resulting from centuries of interventions.

Survey represents one of the major instruments for deeply investigating the intimate nature of such artifacts thanks to its ability to enlighten both the studied object and its context, their mutual relationships and the sequence of transformations.

Survey has to be intended though as an 'open' process aiming at improving the general Level of Knowledge of the studied object; a process in which three main phases can be recognized: Data Acquisition, Data Selection and Interpretation, Communication. Phase one includes all information-gathering activities coming mainly from measurements and, under certain conditions, it leads to datasets 'tending' to be objective. Conversely, during the Selection and Interpretation phase this database is 'intelligently' reviewed for a critical selection and interpretation leading thus to 'subjective' results. Finally, during the Communication phase results are codified in order to make them widely available for the scientific community.

The whole process always starts from getting acquainted with the artifact by direct inspection and by an historical investigation of archive sources (documents, drawings, pictures, etc.) able in outlining the timing and evolution of changes. All this information would in fact crucially guide the following operations of data capturing (surveying).

From this standpoint the House of the Knights of Rhodes in the Forum of Augustus is quite emblematic: no recent and systematic documentation is in fact available; present building is the result of many historical phases that have over time added or subtracted elements to the original roman building; the complex is actually a tangle of architectural and archaeological elements and for this reason it can be assumed as a 'showcase'; finally the researches on the House, as relevant part of the Forum of Augustus, have received a new significant impulse by some recent excavations campaigns.

The research on the House of the Knights of Rhodes we are presenting has been then focusing on all these issues, aiming at demonstrating both the inner 'coherence' of the Survey process and the potentials of the Integrated Survey procedures where many surveying techniques (3D scan, photomodeling, ortophotography, topography, GPS, direct survey) are used at the same time in order to optimize time, resources, models and results.

Models (2D, 3D) have been positively used to investigate this very fragmented portion of the Roman Forum (the building itself, the context of which it is part, the overall alignment of ancient structures, the consistency of buildings as well as the functions they used to host) not neglecting their role as means for communication of results and dissemination.

Keywords*:* Integrated Survey, Archaeological Architecture, House of the Knights of Rhodes, 3D Capturing

1. General Issues

Italy is undoubtedly characterized by the presence of a large number of architectures of the past: some of them testify the passage of time, showing a layering generated by the many structural changes from which they were concerned over the centuries. The study and interpretation of the stratified architecture, and more generally of the historical one, is a difficult to manage operation. The great importance assumed in Italy by the culture of conservation and restoration has provided a major boost to further research in the documentation and in the survey of these structures; the ultimate goal was often the return to the community, making them museums or containers of culture, preserving intact their architectural qualities.

The survey is one of the main tools for deep investigation of these artifacts thanks to its ability to understand the object of study and to put it in its context, as well as to decipher the relationships between the different components and the temporal sequence of the constructive transformations.

This particular contribution, regarding the experience of the Survey of the House of the Knights of Rhodes in the Forum of Augustus,[1] fits into continuity with the work done by the School of Survey of Sapienza University of Rome, which has focused for decades on the analysis and the arrangement of various advanced methodologies to be used in the activity of survey documentation of the architecture, in particular the complex and layered one, as well as the so-called archaeological architecture.

The methodological study on the House may be an exemplary and emblematic contribution in the context of layered architecture: the current building is the product of a centuries-old layering, and there is not a recent complete documentation about it, neither of the survey kind nor

[1] The contribution derives from the research carried out as part of the Doctoral Thesis in Sciences of Survey Representation - Sapienza University of Rome – by Gaia Lisa Tacchi entitled 'Documentazione e conoscenza di complessi architettonici stratificati: rilievo integrato e rappresentazione della Casa dei Cavalieri di Rodi al Foro di Augusto'. Supervisors professor C. Bianchini and professor M. Docci (http://padis.uniroma1.it/handle/10805/2069).

other. Furthermore, highly relevant architectural and archaeological elements coexist in it, as it is placed in the wider context of the Forum of Augustus, which, in the light of the recent excavations[2] (Fig. 1), is the subject of a significant renewal of the studies.

The graphic and iconographic documentation on the House of Rhodes is on the contrary rather significant, given its interconnection with the structures of the Imperial Forum, and given the long period of time in which the building has developed its phases of construction, from the late Republican Age until today. The morphology and appearance of the building have changed over time, as well as its function and naming.[3]

The overall survey of the building placed in its context may be a valuable tool for morphological reconstruction of the fragmented archaeological site. Parts no longer in contact with each other, even though coeval, may be connected again, establishing and verifying alignments that can lead to different interpretations of the planimetry, consistence and intended use of the old buildings.

2. Scientific Survey as implementable System of Knowledge

The investigation of the House of the Knights of Rhodes was conceived as an 'open' process, consisting of the creation of a knowledge system ever implementable, based on an integrated survey - carried out in a scientific manner - that uses verifiable existing data and can therefore become usable for the scientific community in the future.

The study had among its main objectives the demonstration of the internal consistency of the used methodological investigation process, built on the capabilities of the different methodologies of survey - 3D scanning, photomodeling, ortophotography, topography, GPS, direct survey - integrated with each other in order to optimize the processing time, the available resources, the proposed representative models and the results of the research.

Figure 1: The House of the Knights of Rhodes in the archaeological context of the Roman Forum. View from Via Alessandrina.

The stages that can be distinguished in a survey process are the *Data Acquisition*, the *Data Selection and Interpretation* and the last phase, which includes the *Representation and Communication of data*. With regard to the choice of different procedures, a significant role is played by the type of architecture to be detected, together with its characteristics, its morphology and the constraints that could make it difficult to survey. From these considerations have emerged some thoughts, or assumptions, which then guided the drafting of the survey project as well as the detection activity itself.

3. Preconditions for the drafting of the Survey project

The first considerations that have been carried out concern the *strong morphological heterogeneity* of the House of the Knights of Rhodes that, more than a building, takes shape as a complex architectural structure added to an archaeological site of which it is an integral part. Hence its double meaning of *architecture and archaeological structure*, which brings with it a great diversification of the structures to be detected, solvable with an integration of different data acquisition techniques and with subsequent different methods of constructing representative models (Fig. 2). The oldest part of the House and the whole part below the entrance walking level – on the 'Salita del Grillo' – and built on the Roman Forum is made up of complex and articulate structures with a layering of walls developed in three dimensions and in which it is often difficult to identify those geometries that can allow for a mental or hypothetical reconstruction useful to the detection of architectural elements through the acquisition of previously selected points. Not only that, the walls seem to configure alignments that vary, even if imperceptibly, at different elevations. It became clear that a campaign of laser scans could be useful in identifying the complex morphology of the walls, allowing to have, in the later stage of two-dimensional models processing, a huge series of sections at different heights to check the alignment of walls. In addition, the acquisition rate of the point clouds in the archaeological area would provide a vast pool of

[2] See: DELFINO, A. (2011). Le preesistenze del Foro di Augusto. In MENEGHINI R., SANTANGELI VALENZANI R. Scavi dei Fori Imperiali. Foro di Augusto (l'area centrale). Bullettino della Commissione Archeologica Comunale di Roma. p. 11-31; LAMBOGLIA, S., MUSOLINO, F. (1995). L'edificio romano della Casa dei Cavalieri di Rodi. In UNGARO, L., LA ROCCA, E., MENEGHINI R. I Luoghi del consenso imperiale. Il Foro di Augusto. Il Foro di Traiano. Roma: Progetti Museali Editore. p. 52-61.

[3] The historical iconography consists of several modes of expression, depending on the historical period in which it is produced, and this happens in parallel, over the time, with the gradual transformation of the architectural structures: from the early sixteenth notes (Sangallo, 1516) that focus the attention on the consistency of the emerging structures in the area of the Roman Forum and restitute part of the oldest structure of the House, to the seventeenth and eighteenth centuries views that represent it in its relationship with the context - at the time a dense urban fabric - when it was part of the subsequently demolished Convent of SS. Annunziata (Alò Giovannoli, 1618; Michel d'Overbeke, 1763), up to the twentieth century metric processing aimed at the discovery and valorization of the complex following the 'excavations of liberation' of the Forum made during the Fascism, the demolitions that spared the House making it de facto the symbol of the memory of the site. See: BUZZETTI, C., GISMONDI, I. (1985); FIORINI, G. (1951).

data useful in the control of wall structures adjacent to the House, also part of the same archaeological site, and of which we had an unverifiable graphic documentation. Also from the dimensional representation point of view, the point cloud could allow the realization of a three-dimensional model in a semi-automatic and sufficiently reliable and communicative way, at least for the parts with complex morphology.

The upper parts of the House were predominantly simple volumes, consisting of plans or however sufficiently regular surfaces: structures regulated by geometries identifiable from points that could be selected in the design phase of the survey; the interior could also be detected by points, while presenting decorative episodes (capitals and bas-relief) - and also a very large fresco - that deserved special attention in the process of detection. In this case the topographic surveying, that however would have been the basis and prerequisite of all survey activities - as an element of mediation between the different integrated detection techniques - could be enough to identify the main points of the architectural shapes, to detect with greater accuracy the edges (even in the presence of laser scans) and to form the basis for photo straightening - useful to the reading of the stratigraphy present on the masonry walls although flat - panoramic pictures - for the documentation of frescoed walls - and some test of photo modeling on capitals and small size relief structures.

Figure 2: The morphologically complex and heterogeneous structure of the House, which rises from the remains of the exedra of the Forum of Augustus.

Other considerations were carried out with regard to the *vastness of the complex* to be detected: having not as target a survey limited only to the building of the House of the Knights, but rather the investigation of the relationships and the changes that occurred within the complex system to get to reconstruct the building phases from which it was concerned over the time, it was necessary for the general survey not to arrive at a scale with a denominator too small. We then opted for the use of scale 1:100 in the two-dimensional representations, an unusual choice in architecture: this allowed us to have a pool of data not too large and therefore more manageable. From time to time, when necessary, starting from the main structure we reached a greater detail by increasing the data, the study and the representation of elements of particular interest.

The study aimed to understand the sequence of the structural transformations of the building and of its context from the creation of the first group – of which remains only a four sided travertine portico presumably of late Republican age - to its current configuration. In this sense, we would have data from the survey or from the direct ‘reading’ of the building and other types of information from literature and archival sources, as well as from historical iconography and most of all from the important survey drafted in 1930[4] by Italo Gismondi (Fig. 3). The next step would be to combine the data, that is construct two-dimensional geo-referenced models to put in relation with charts drawn up by the archaeologists that, in the past and even now, dedicated themselves to the archaeological site object of our interest. So another preliminary observation was made on the need to work on the field as a team in a multidisciplinary manner.

It was clear in this context that, after having synthetically derived the transformations occurred over time - on two-dimensional models – it was useful to make a three-dimensional representation of the sequence of construction phases along the line of time: a simplified model, neutral, constructed from the survey data, that would first check - even metrically - the reconstructive hypotheses carried out by Gismondi and other researchers, and that could also be a means for communication of the aforementioned modifications.

4. Survey Project and Data Acquisition

In order to optimize resources through the use of existing verifiable material, the survey project included the existing data of a survey campaign of the late Republican portico, now the Chapel of St. John.[5] The data consisted of a 3D laser scanning and a support topography used for the recording of

[4] The measurements of the Forum of Augustus taken by Italo Gismondi in 1930-31 for the *X Division of the Governorate of Rome* after the excavations of 1926-29, were published in 1985. They constitute an important contribution because the author has increased the knowledge of the building adding constructive data to the metric ones, assuming the dating for several layered and interconnected masonry apparatuses. See: Buzzetti ansd Gismondi 1985.

[5] The working experience of the Chapel of St. John, led by prof. Carlo Bianchini, started at the ‘arch_eo - Architecture for Archaeology, Archaeology of Architecture’ Master, held at the Sapienza - University of Rome and coordinated by dr. Claudia Cecamore.

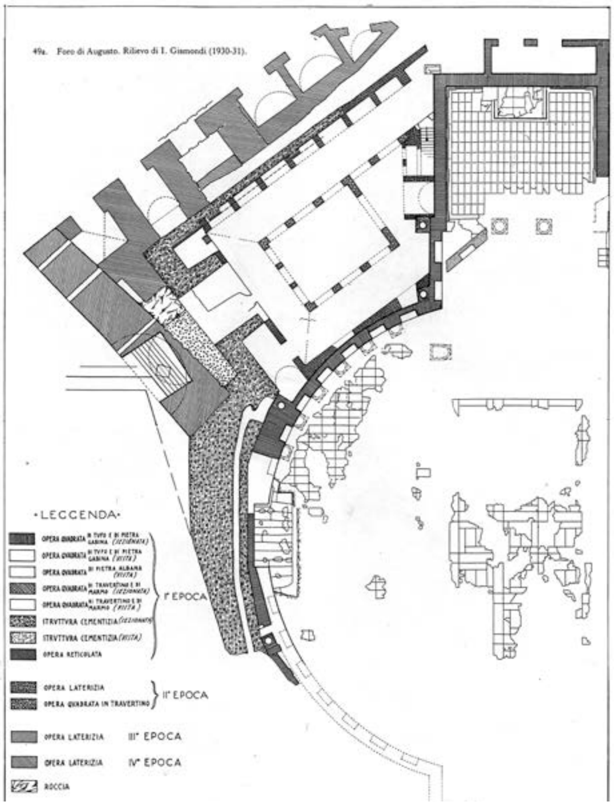


Figure 3. Plan at the level of the ancient four-sided portico drawn by Italo Gismondi in 1930 (from Buzzetti, Gismondi 1985).

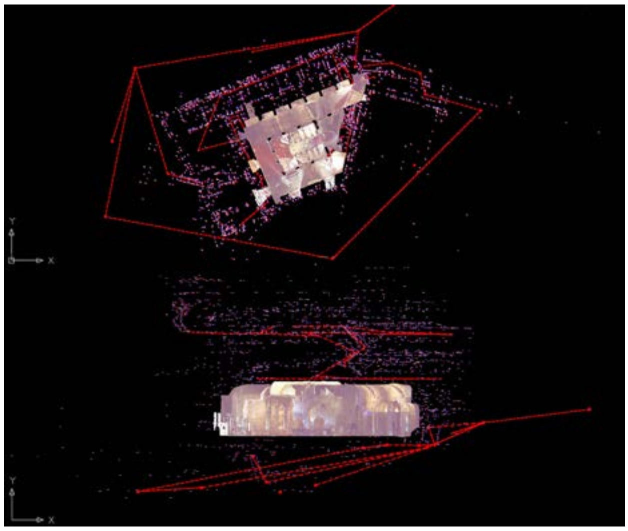


Figure 4. Construction of the general topography of the building from the first detection tests. The support topography of the first scans was the starting point of the overall polygonal.

the point cloud (Fig. 4). Starting from the points measured in this initial topography, a closed polygonal, external and compensated, was built to be the essential basis reference system for all the other detection techniques integrated to it.[6] The absolute coordinates spatial reference for the topographical reference points, useful to georeference the points of the polygonal and with it the survey as a whole, was determined using a GPS system.

The topographic survey, as planned, involved some points of the exterior of the building, selected at the beginning, the interior - consisting of simple geometries -, and provided the support for the realization of photo straightening of flat walls. The points of the station related to the external polygonal were used for the realization of a campaign of 3d laser scans,[7] concentrated mainly on the outside of the building: these data have allowed to study, measure, and reconstruct the extremely complex surface of the architectural work providing the level of precision and accuracy required in the study (average < 4 mm.). Furthermore, as expected, the large amount of data from fast acquisition scans has allowed the metric control of almost the whole site.

To facilitate and optimize the detection operations, the scanner was used in 'topographical mode', namely entering, one by one, the coordinates of the topographical landmarks inside the instrument, and directing it from time to time on the other cornerstones of known coordinates; this procedure has speeded processing operations of point clouds that are produced automatically recorded, that is to say joined and oriented between them. This method of scanning has allowed us to measure some new topographical landmarks that for logistical reasons had not been previously examined with the total station. The key points, known points, are a fixed reference system, to which you can couple every time you want to increase the data with survey campaigns to come, our own or of other scholars who wish to take up the work and carry it forward, in our view a fundamental assumption in research activities (Figs. 5 and 6).

The detection of the detail elements has been performed in different ways depending on the need to return them with a two-dimensional processing - in relation to the scale of return, which in this case is not a scale of detail, being the ratio 1:100 - or to obtain three-dimensional processing. Some essays of photo modeling were, however, carried out as a sort of experimentation and integration of the detection activity. These essays were performed using an *Autodesk open source software - 123D Catch* - on particularly interesting architectural details or on sculptural elements and decorations found in many parts of the building.

5. Data Selection and Interpretation: from the updated cartography to the model of the different phases of construction

With regard to the survey restitution, the above procedure has been realized organized in two phases: the creation of 2D models to be compared with the existing iconography and, therefore, the use of survey data to create a synthetic three dimensional model representing the results of this cross-referenced information analysis. This last phase

[6] The constraints in the archaeological site in order to close the external polygonal belong to different natures: the strong difference in height between the archaeological area and the entry level to the building on the 'Salita del Grillo', and the forced path to make a complete circuit around the complex are the main ones. As a result of this, the points of the station had to be numerous while the arms of the polygonal had inhomogeneous lengths. The polygonal has shown an error of closure of 2 cm on 9 points of the station, and while it were a more than acceptable error for the chosen scale of restitution - 1:100 - precisely because of the inhomogeneity of its components, we preferred to offset it.

[7] The type of laser scanner used is the Leica C10 model.

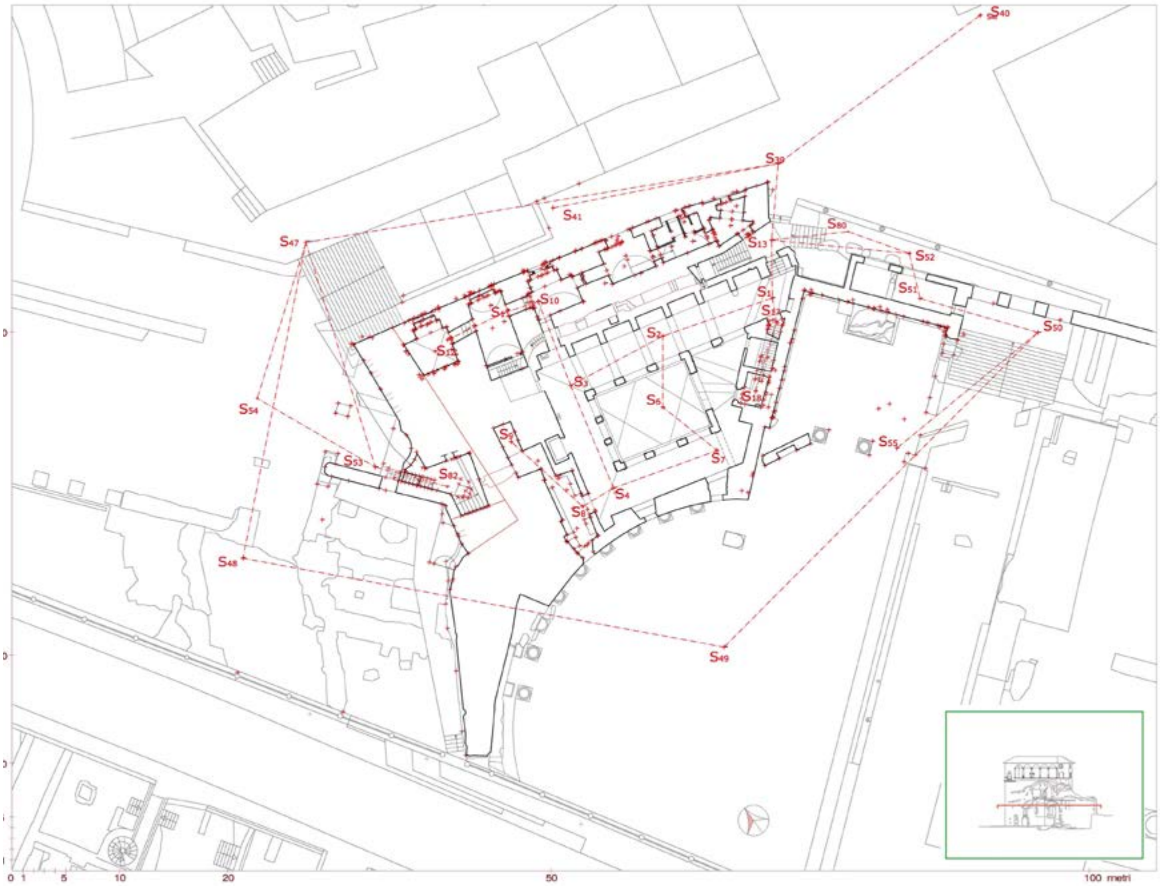


Figure 5. Geometric Survey. Restitution of the implemented survey methodology: topography, four-sided portico level plan.

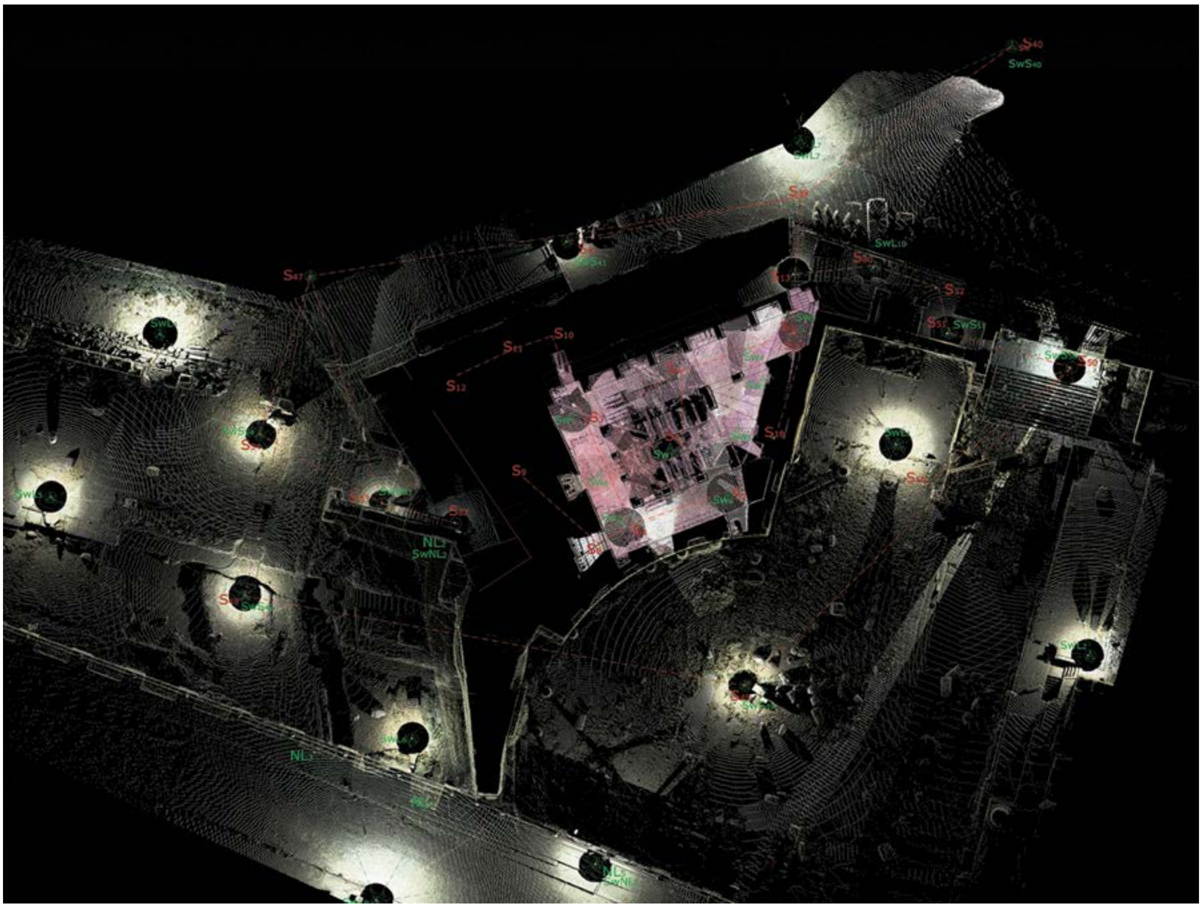


Figure 6. Geometric Survey. Restitution of the implemented survey methodology: laser scans, four-sided portico level plan.

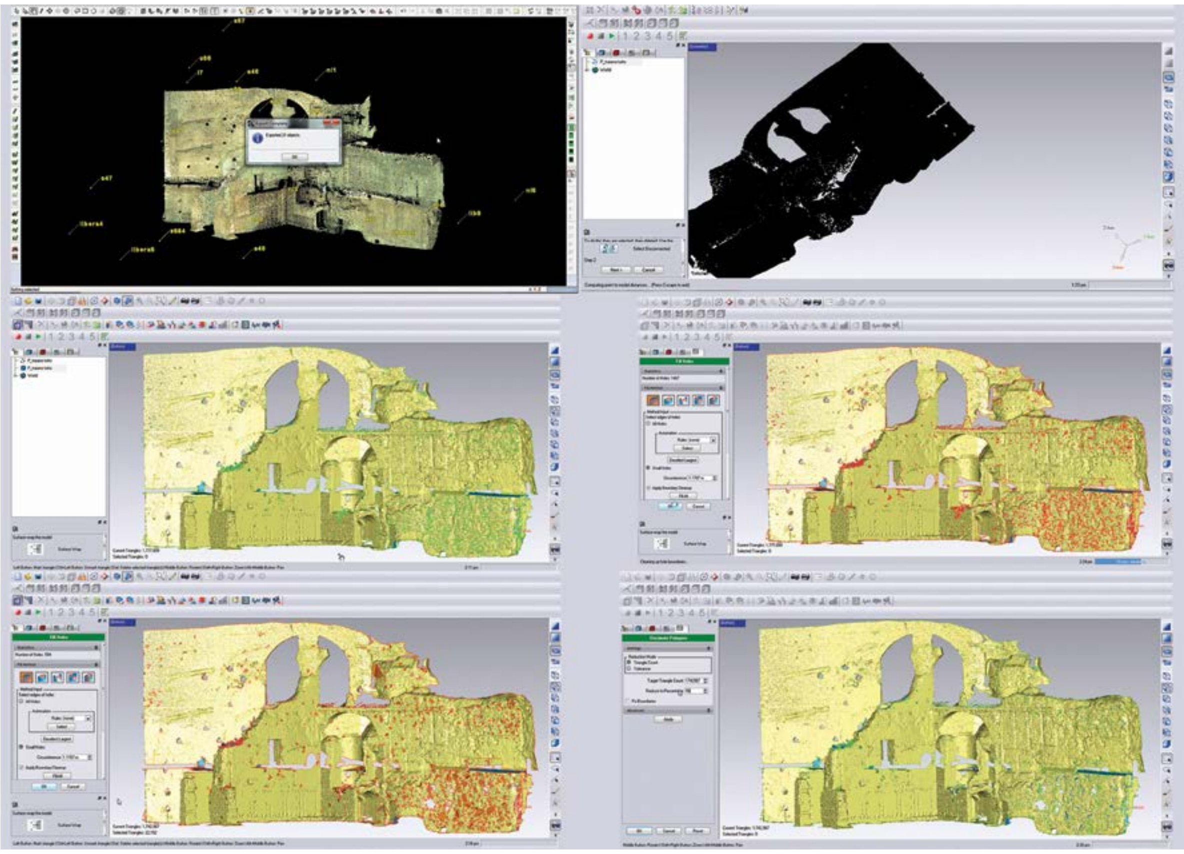


Figure 7. Modeling of the morphologically complex parts made from the point cloud to be assembled in the general model. The connecting structures between the building and the Forum of Trajan (Software Cyclone and Geomagic).

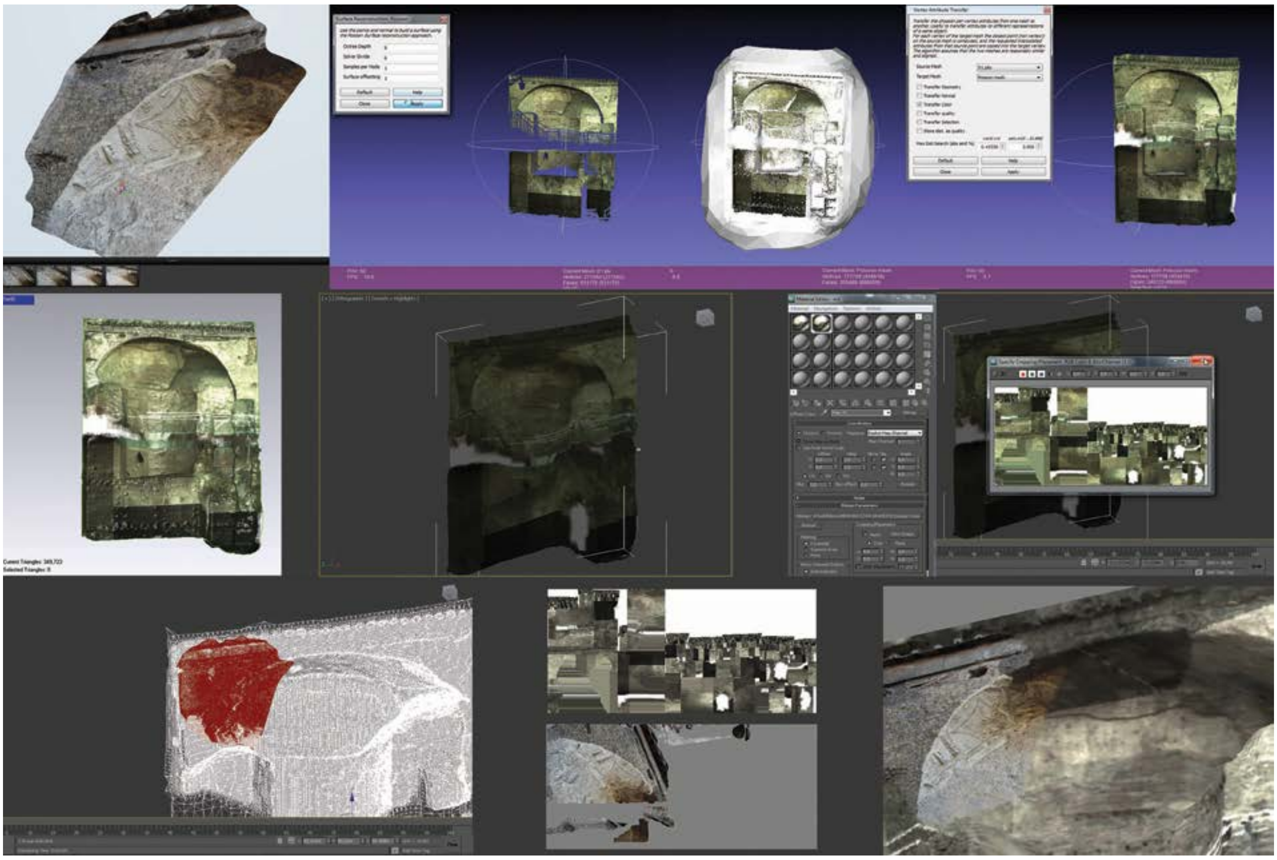


Figure 8. Textured model of an important element for the understanding of the historical stratification of the House. This is a portion of the vault with remains of decorated stucco against the wall on the Forum of Trajan. Processing of data from the point cloud (3D laser scans on the niche) and from the photo modeling of stucco (Software Cyclone, 123D Catch, 3DStudio Max with V-Ray).

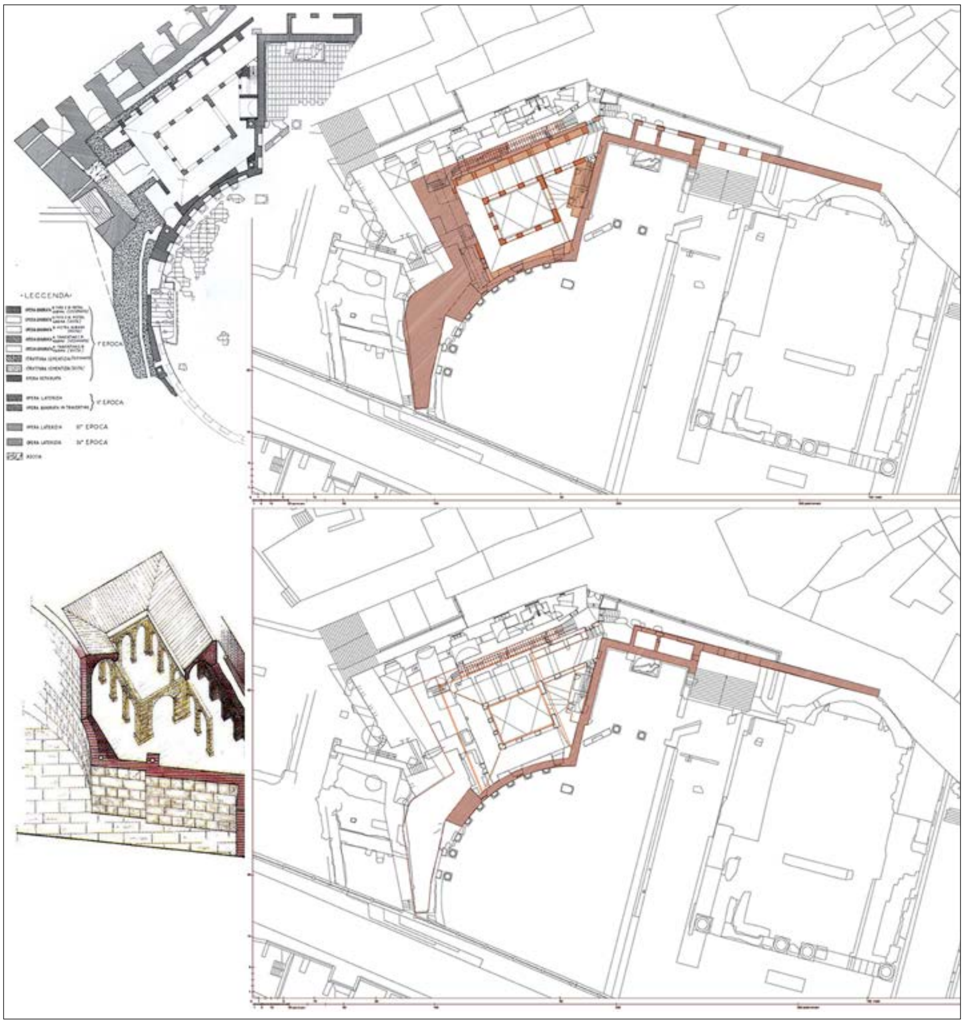


Figure 9. Late Republican / Augustan period. Translation on the present survey of the considerations made by scholars regarding the construction phases of the building (Buzzetti, Gismondi, 1985; Lamboglia, 1995).

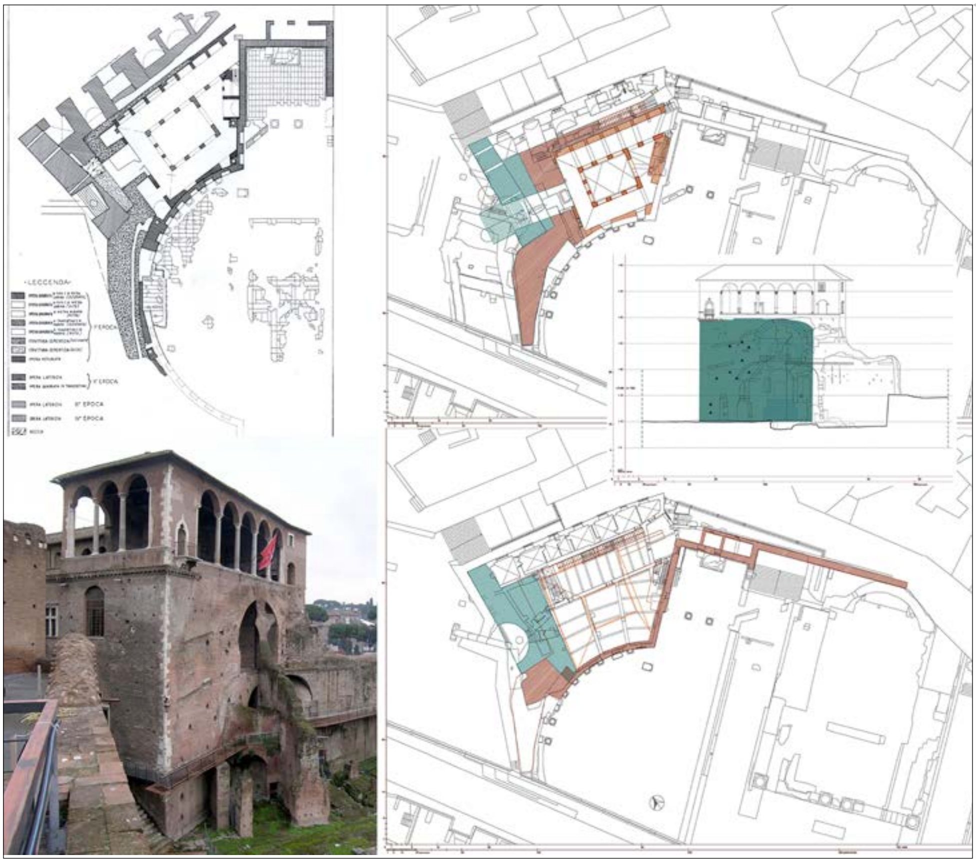


Figure 10. Domitian period. Some elements belonging to this stage are still visible, like the thick wall and the large exedras that probably looked out onto a square at the time of Domitian. At the top a reconstructive drawing of the staircase by Gismondi (Buzzetti, Gismondi, 1985; Meneghini, 2009).

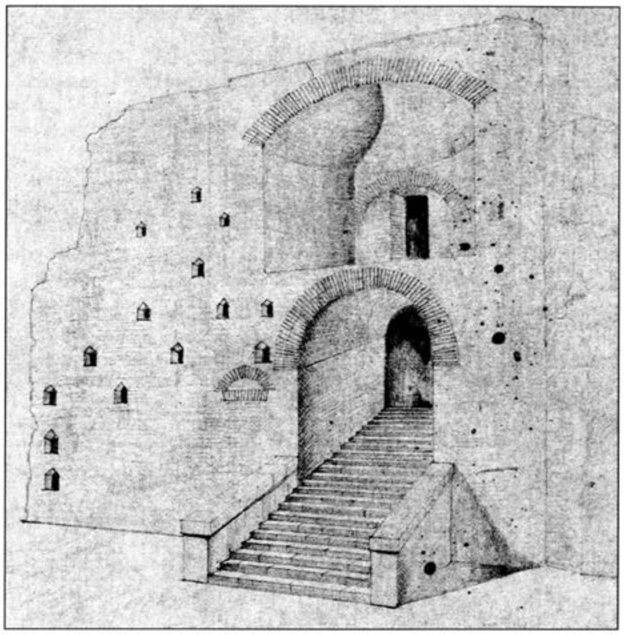


Figure 11. A 1930 three-dimensional reconstructive drawing by Italo Gismondi depicting the Domitian Wall with a hypothesis of the monumental staircase unwinding (Buzzetti, Gismondi, 1985; Meneghini, 2009).

has widely used numerical modelers that have allowed a more free definition of morphologically complex shapes, not based on geometric primitives, and that could easily include data from 3D scans (Fig. 7) or from photo modeling (this last mainly using Autodesk 123D Catch software)[8] (Fig. 8).

The major construction phases have been represented, in their evolution, through a schematic model of the building and of the immediately adjacent environment. Almost all of the structural changes represented are based on reconstructive studies prepared by scholars, published and known, but generally enclosed in particular studies and not really about the House of Rhodes, but about the buildings adjacent to it and that in one way or another have concerned and changed it (Figs. 9-13). In the selection of the evolutionary phases of the building to represent, we followed a chronological order and turned the attention to those which have seriously changed the morphology of the building making it progressively as we see it today, to those that have changed its relationships with the outside, and to the moments that have built wall alignments and defined spot elevations, obviously being shape generators. The deducted phases were called in this way: late Republican

[8] For the reconstruction of morphologically complex structures, we used the meshed point cloud, for the simple volumes we proceeded with two-dimensional graphics modeling; for the reconstructed parts we deliberately used simplified forms for the hypothetical quality that characterizes them. Reference software used: processing point cloud Cyclone 7.3.3, Geomagic 10; photo modeling 123D Catch; modeling and rendering 3DStudio Max with V Ray.

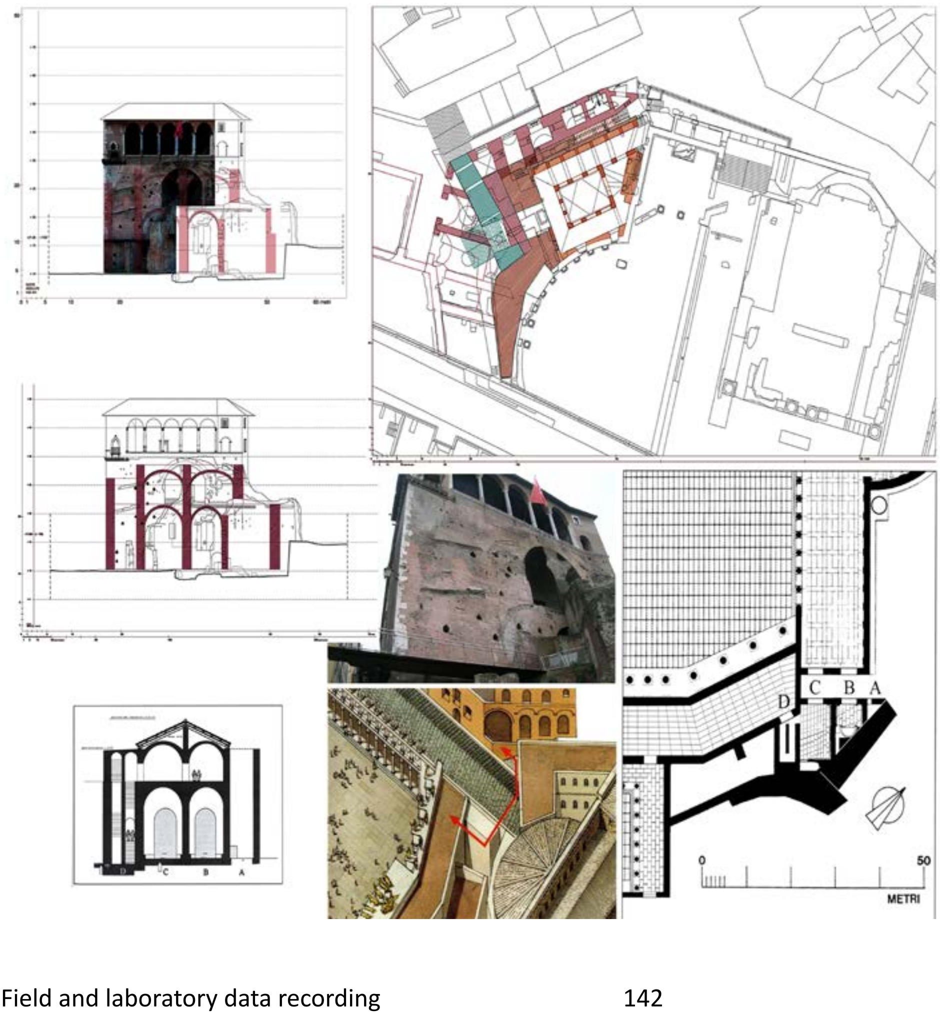


Figure 12. Trajan period. Roberto Meneghini produced some hypothetical reconstructions of the structures that at this stage leaned on the Domitian prospect. From a comparison of the current survey with this reconstruction, the mark of the structures that no longer exist on the current prospectus has been identified, confirming his hypothesis of reading (Meneghini, 2009).

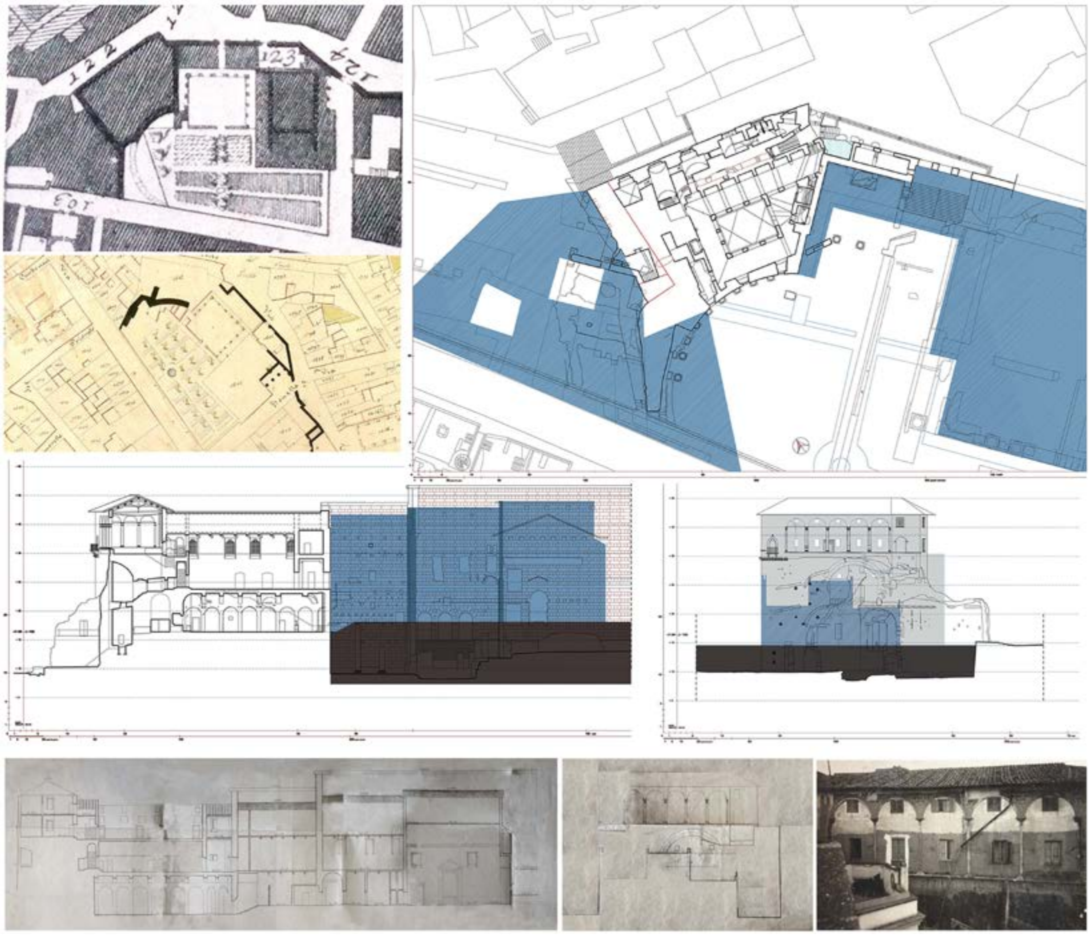
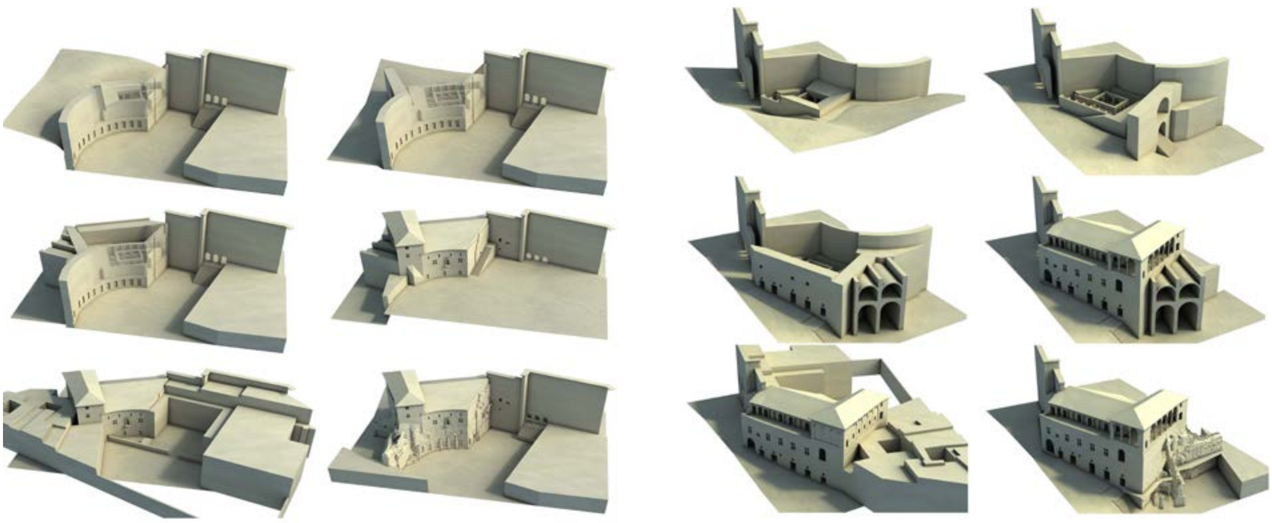
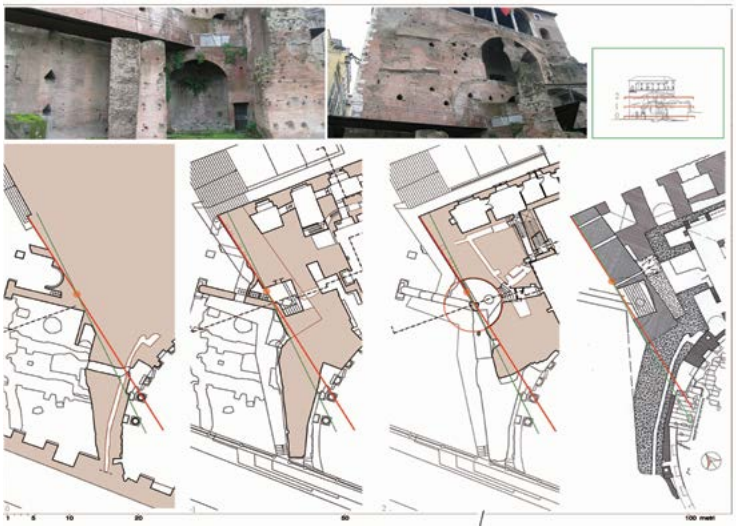


Figure 13. Post Renaissance and Modern period. At the top, the plan by Nolli (1748) and the Urban Land Registry of the Monti neighborhood (1818-1871) represent the Convent of SS. Annunziata. Other information comes from historical photos and a few drawings of the Convent from the Central Archive of the State (bottom). The images point out a thickening of the buildings adjacent to the House as the Convent planned for it an intensive use of space. The integration of deduced information with the current survey allows a simplified reconstruction of the constructive phases. The historical photos come mostly from the 1928 archive 'Foto Filippo Reale', and from the 1930 archive of Cesare Faraglia (from Leoni, Margiotta, 2007).



Figures 14-15. The reconstruction of the evolutionary phases of the building via a synthetic three-dimensional model of the acquired knowledge. The long history of the building was reduced to a few key phases, those that have clearly contributed to a substantial transformation of the architectural structure: late Republican and Augustan period; Domitian period; Trajan period; Renaissance period; post-Renaissance and modern period; current period.



Figures 16. Analysis of the Domitian wall in plan. Identification of the hinge and of the rotation of the wall. Comparison of the new survey's plans taken at different heights with the 1930 plan by Gismondi.

and Augustan period; Domitian period; Trajan period; Renaissance period; post Renaissance and modern period; current period[9] (Figs. 14 -15).

6. Conclusions

The main objective of this study was to lay the foundation for a system of knowledge concerning the House of the Knights of Rhodes based on an integrated survey, implementing all the available cognitive material on the building and opening up new prospects of research. The new study has in particular allowed, due to the potential expressed by the point clouds processing, to be aware of the inclination of a Domitian period wall, apparently quite flat, corresponding to the prospectus crowned by the Renaissance loggia, called 'Domitian Terrace'[10] (Fig. 16), and to evaluate its angle. Since these changes are not easily visible to the naked eye, they are difficult to include in a selective investigation because they can't be a subject of study. The interest in this kind of 'discoveries' is based on the amount of information that they can provide to the scholars who for many years have focused on the analysis of the layered walls of the building, and to those who, like us, investigates the methodological usefulness of new techniques for detection and the possibility of processing the collected data.

[9] During the Renaissance, the building of the House of the Knights looked similar to the current days, then it was gradually the subject of an intensive use of the area, being included in a monastery leaning against the thick retaining wall of the Forum of Augustus. The progressive building development of the site has stopped with the demolition of the 'Alessandrino' neighborhood to free the Forum, an operation that saved the House of the Knights of Rhodes as a symbol of remembrance of the site destining it to the Museum of the finds from the Roman Forum. For documentation on the historical phases beginning from the Renaissance, see: DANESI SQUARZINA, S. (1989); RICCI, C. (1925-26); RICCI, C. (1930); PIETRANGELI, C., PECCHIOLI, A. (1981); LEONI, R., MARGIOTTA, A. (2007).

[10] Much has been written about the so-called 'Domitian Terrace', the thick wall that, since the initial stage of its construction was crowned with a terrace - where the fifteenth-century loggia now stands. Studies on this wall are mainly dedicated to the understanding of its structure and its function. See TORTORICI, E. (1993) La Terrazza Domizianea, l'Acqua Marcia ed il Taglio della Sella tra Campidoglio e Quirinale. Bullettino Comunale 95. p. 7-24; MENEGHINI, R, SANTANGELI VALENZANI, R. (2009).

References

BIANCHINI, C., TACCHI, G.L. (2013). Il Rilievo come Sistema di Conoscenza: la Casa dei Cavalieri di Rodi - Survey as a Knowledge System: the House of the Knights of Rhodes. *Disegnare Idee Immagini - Drawing Ideas Images*, 44. Roma: Gangemi Editore. pp. 60-71.

GAIANI, M., BENEDETTI, B., APOLLONIO, F. (2011). Teorie per rappresentare e comunicare i siti archeologici attraverso modelli critici. In: *SCIRES-IT SCientific RESearch and Information Technology Ricerca Scientifica e Tecnologie dell'Informazione*. CASPUR - CIBER Publishing, vol. 1 (2). pp. 37-70.

MENEGHINI, R., SANTANGELI VALENZANI, R. (2009). La Terrazza Domizianea. In: *I Fori Imperiali e i Mercati di Traiano: storia e descrizione dei monumenti alla luce degli studi e degli scavi recenti*. Curated by the authors. Roma: L'Erma di Bretschneider. p. 111-116.

LEONI, R., MARGIOTTA, A. (2007). *Fori Imperiali. Demolizioni e Scavi. Fotografie (1924/1940)*. Milano: Electa, 2007.

DANESI SQUARZINA, S. (1989). La Casa dei Cavalieri di Rodi: architettura e decorazione. In *Roma, centro ideale della cultura dell'antico nei secoli XV e XVI*: *da Martino V al sacco di Roma 1417 - 1527*. Curated by the author. Milano: Electa. pp. 102-142.

BUZZETTI C., GISMONDI, I. (1985). Foro di Augusto. *Bollettino Comunale*. 90. pp. 341-343

PIETRANGELI, C., PECCHIOLI, A. (1981). *La Casa di Rodi e i Cavalieri di Malta a Roma*. Roma: Editalia.

FIORINI, G. (1951). *La Casa dei Cavalieri di Rodi al Foro di Augusto.* Roma: Libreria dello Stato.

RICCI, C. (1930) Il Foro di Augusto e la Casa dei Cavalieri di Rodi. *Capitolium.* VI. pp. 157-189.

RICCI, C. (1925-26). La liberazione dei resti del Foro di Augusto. *Capitolium*. I. pp. 3-7.

Digital Research Strategies for Ancient Papyri: A Case Study on Mounted Fragments of The Derveni Papyrus

Eleni KOTOULA, Graeme EARL

Archaeological Computing Research Group, University of Southampton (ACRG), UK
http://acrg.soton.ac.uk/ac.ukreview/acrg.soton.ac.uk
graeme.earl@soton.ac.uk

Abstract

The present paper discusses the problematic aspects of the study of papyri both form papyrological and conservation perspective. Common problems for examination, documentation and study of the writings are presented as well as conventional conservation treatments, such as unrolling, mounting and de-restoration. It addresses the interventive conservation operations and the advanced digital analogues available nowadays for virtual unrolling, as demonstrated by the recent successful applications of 3d Computed Tomography (CT) and terahertz imaging. Moreover, it evaluates virtual visual analysis, documentation and diagnostic examination based on reflectance transformation imaging (RTI) approaches in the visible and the infrared spectral area using both reflected and transmitted illumination and irradiation. The new RTI techniques introduced have the potential to enhance study and conservation of papyri. Results indicate that the combination of RTI, IR-RTI and Transmitted RTI is an enhanced methodological approach for non-destructive examination, documentation and presentation. A case study on one of the most important papyri ever found in Greece, considered as the oldest Greek text written on a papyrus discovered in 1962 in the ancient Mygdonian city of Lete, on the pass of Via Egnatia, demonstrates the potential of integrated RTI approaches for the study of ancient papyri.

Keywords: Papyrus, RTI, IR-RTI, TRANS-RTI

1. Study and conservation of papyri: conventional practices and common problems

1.1. Study of the writings

One of the main problematic aspects of the study of ancient papyri is the readability of the texts, not only because of the various handwriting styles and languages present in ancient papyri but also because of the bad preservation of the inks or the papyrus substrate. Ink may be faded, flaking off, rubbed off or lost making it almost impossible to decipher the writings. In case of charred, carbonised or semi-carbonised papyri reading texts is highly problematic because of the blackening of the papyrus material in addition to the blackened inks, which might originally have been black or blackened due to pyrolysis: the process that leads to carbonisation of the papyrus. Another problem is the buckling of the surface which distorts some letters and may lead to misreading of the text. Moreover, after unrolling it is highly likely for the different layers of papyrus to be stuck together, making the identification of the edges of the different layers extremely difficult and in some cases impossible (Janko, 2002).

1.2. Conservation Treatments

Unrolling papyri was a common conservation practice in the past and it is still in use nowadays. Several techniques have been applied, including the juice from the papyrus plant, static electricity, electromagnetic action of the bulb humidification and cellulose treatment (Shepherd, 2008; Frosen, 2009; Lau-Lamb, 2005; Leach & Tait, 2000; Owen & Danzing, 1993). These methodologies fall in a category of interventive conservation techniques similar to archaeological autopsy, micro-excavation and investigative cleaning. There is no doubt that archaeology has gained a vast amount of information from the execution of such processes, but at the same time their destructive, non-reversible interventionist character should not be neglected. Even when highly skilled professionals are employed following appropriate tools and techniques for these 'micro-archaeology' operations, including careful recording, mechanical controllable means and additional stabilization, this type of operations may endanger the materials integrity of fragile papyri.

Another common practice is the mounting of papyri fragments and their preservation sandwiched between two pieces of glass, Plexiglas or mylar (polyester) (Machteld van der Feltz, 2010; Lau-Lamb, 2005; Johnson & Horgan, 1979; Owen & Danzing, 1993). The main reason for mounting is to secure the fragile material and enable its study by researchers and its exhibition. As a result the material integrity is safeguarded and the undesirable material deterioration due to handling is limited. But, the presence of mounts is associated with blooming due to salts contamination (Banik & Stachelberger, 1987; Pollard *et al.,,* 2011; Stanley, 1994), biological growth (Cappitelli & Sorlini, 2005) and sometimes irreversible damage due to glass breaking. In the past the use of inappropriate materials resulted in discoloration of the backing due to acidification, fractures and cracks due to cockling, and also damage cause by slipping papyrus fragments. The direct contact of the glass with the papyrus is responsible for heat and moisture, both considered agents of deterioration.

Also, sometimes the back side of the mount is completely covered by the backing material. This is the case not only in mounted papyri but also in made up rolls, where fragments stacked together and shaped in the form of a roll often with linen tape bund around the outside (Leach & Tait, 2000). Such actions limit considerably the study of the material aspects of the papyrus.

Moreover, de- restoration of papyri, meaning the removal of previous repairs, especially for those papyri which have already been subjected to a significant chemical breakdown either due hydrolysis or oxidation, may be a greater danger. Although methods to deal with this problem have been presented in the literature (Wright, 1983; Owen & Danzing, 1993; Taylor *et al.,,* 2011) there is the risk to reduce the papyrus back to fragments or loose fibres and ink. This applies also in de-mounting operations, where the removal of the glass may lead to damage of the inks.

1.3. Examination & Documentation

The mount poses a barrier in the study of the papyrus and affects negatively the results of visual analysis, which is a milestone in artefacts studies, as it seeks to provide data relevant to structure, manufacture, damage, decay and use-wear of the object as well as materials identification. One of the most crucial features for the correct perception of information revealed from visual analysis is the perception of three-dimensionality, which is limited because of the glass mount. In particular in case of papyri conservation examination is a demanding operation. The most important features that need to be identified are the joins, the old repairs and the possible mounting problems, so as to define the treatments needs. Other features included in reports are the state of preservation of the fibres and the ink, the presence of dirt, accretions, stains, misalignments and missing fibres. The inks may have a damaging effect on the papyrus, as reported for copper acetate but also for black and red inks (Graf & Krutzsch, 2008). Also, of interest is the detection of evidence of papyrus re-use. In palimpsests the abraded surface usually becomes evident under microscopic examination (Leach, 2008). Close inspection under the microscope reveals patterns (vascular bundles) of the papyrus which can help in alignment of fragments. The detection of these features is crucial for the reconstruction of the object history, its investigation and future preservation. The fragility of papyrus material makes the execution of the already demanding operations even more difficult. Sometimes the proper photographic documentation becomes problematic, either due to scale issues or difficulties in handling.

2. The Derveni Papyrus Case study

2.1. Discovery, Preservation & Conservation

The Derveni tombs in Thessaloniki, Macedonia, discovered in 1962, are considered one of the most significant archaeological sites in northern Greece because of their numerous rich grave offerings and their important location in the ancient Mygdonian city of Lete, on the pass of Via Egnatia. The cemetery comprises seven graves, and according to the excavation publication dates to 320–290 BC. One of the most remarkable finds of the Derveni Tombs is a papyrus roll (DP), discovered in the remains of a funeral pyre above tomb A, a cist grave full of bronze and clay pots, vessels, jewellery and small objects. The DP is a unique case not only for the region of ancient Macedonia but also for the whole ancient Greek, if not European culture, as it is considered the oldest Greek text written on a papyrus. Its archaeological and philosophical significance as well as its extraordinary burial and preservation pose interesting research questions (Themelis & Touratsoglou 1997). Although papyri are the overwhelming writing material of ancient texts (Bulow-Jacobsen 2009), archaeological research in Greece has revealed a limited

Figure 1: Detail of Derveni Papyrus fragment, comparison of digital image (left) and RTI visualization (right).

number in comparison to Egypt and Middle East. One of the most crucial factors is the fragility of the organic matter which is significantly affected by climatic conditions. The absence of wet anaerobic conditions or dry conditions favours the disintegration of the vast majority of papyri in the Greek climate. The DP does not fall within these categories but it survived because of the burnt soil and the ash layer of the funeral pyre. It was the carbonization by the fire of the funeral pyre that protected the roll from degradation and enabled its partial survival.

Papyri in such a condition have been found in Petra in the Nile delta (Boubastos, Tanis and Thmouis), while the Herculaneoum papyri, although considered an exceptional case, present similarities to the Derveni papyrus roll. The upper part of the papyrus roll, around 7-8 cm, was preserved carbonized in two pieces, while the rest was burnt. After its discovery scholars wondered whether it will be possible to open it out and recover the text remains to be seen because of its powdering state (Eugene Vanderpool 1962: 390). Kapsomenos expressed his concerns regarding the condition of the papyrus stating that '(the papyrus) reduced to carbon dust by the slightest touch' (Kapsomenos 1963: 4). Mr. Anton Fackelmann, curator of the papyrus collection of the National Library of Vienna, unrolled the papyrus roll. Although the unrolling method used proved successful scholars reported the difficulties faced due to the carbonization of the papyrus, in particular in the outer part, where the thermal degradation had been more severe. 'The inevitable damage caused at the edges of the fragments by the conservation processes' was mentioned (Kapsomenos, 1963: 5). Moreover, notable is the fragmentation of the text, as a result of the unrolling, in addition to the changes in the physical state of the roll, such as applying forces that cause damage to the edges of the fragments.

2.2. The Derveni Papyrus RTI Visualization

RTI (Mudge *et al.,,* 2005), a non-destructive, affordable and easy imaging technique, was applied in order to overcome the already presented limitations and to acquire views of the papyrus which would be impossible by other means. Although RTI is increasingly being used in the cultural heritage sector, there have been limited case studies on papyri. A characteristic example is the ancient papyrus fragment from the Bancroft Library (UC Berkeley) visualised in RTI form (Cultural Heritage Imaging, 2012).

One glass mount was captured with reflected and transmitted illumination and infrared radiation, resulting in a series of RTI files. The recto (front side) was captured with visible reflected illumination and infrared radiation at 720nm, 850nm and 950nm. The verso (back side) was captured with reflected visible light and infrared radiation at 850nm, as well as with trans illumination and trans irradiation at 850nm. The initial hypothesis for the RTI visualization of the Derveni papyrus, based on previous experience in the field of reflectance transformation technologies was that RTI would enhance the study of the papyrus, due to its ability to emphasize texture and colour. However following the exact mainstream methodology as explained in the guide for H-RTI data capture (Cultural Heritage Imaging, 2013a) was not possible. The reflections of the flash on the glass mount result in unacceptable images. However, this phenomenon occurs only in wider angles, so the H-RTI procedure was applied but using only a subset of potential capture data, where the incident illumination angle was low. The reflectance transformation data were processed with the RTI builder software (Cultural Heritage Imaging, 2011) and examined with the RTI Viewer (Cultural Heritage Imaging, 2013b).

2.2.1. RTI

The comparison between the digital images and the RTI visualization demonstrates that RTI visualization is an improved way for the study of ancient texts (Figure 1). It is a common practice for scholars to move the papyrus in their hands so as for the letters to be revealed due to the different reflectance properties of the matte ink and the shiny papyrus background (Janko, 2002). RTI offers the opportunity to perform this examination virtually avoiding any human-object interaction.

Apart from the study of the texts, a complete visual analysis is possible via RTI, including examination of the manufacture techniques as well as determination of papyrus condition and conservation needs. As shown in figure 2, RTI visualization emphasizes damage and alterations on the surface of the papyrus material, such as racks, losses, flaking and distortions. In that sense RTI can be used as a tool, so as to monitor the extent of the on-going deterioration, and guide conservation treatment. Moreover, RTI can document areas in need of repair and twisted fibres, assisting the process of realignment in the correct position. Another advanced option for research provided by RTI is the examination of the papyrus fibres and the study of the criss-cross pattern.

Comparative RTI analysis of papyrus fragments visualised in RTI form can help researchers reach conclusions regarding different papyrus manufacture techniques. In addition, previous repairs such as joins can be documented. Considering that the vast majority of repairs are not documented, RTI can assist in understanding the papyrus's biography by revealing signs of treatments or operations executed in the past.

Virtual relighting of the glass mounted DP fragment emphasizes different features of the fragment. For example, using the specular enhancement rendering (Malzbender *et al.,,* 2004) at the edges of RTI space emphasize the texture of the papyrus, the criss-cross pattern and the irregularities introduced due to degradation and conservation treatment. The writings appear clearer in sharpening or in specular enhancement rendering mode at less extreme angles with maximum specularity and minimum diffuse colour and highlight size. The multi-light rendering mode successfully underlines both texture and writing (Figures 3).

Via RTI the text can be studied more easily and graphic characteristics can be analysed so as to detect the handwriting styles. Scholars discussed the style of

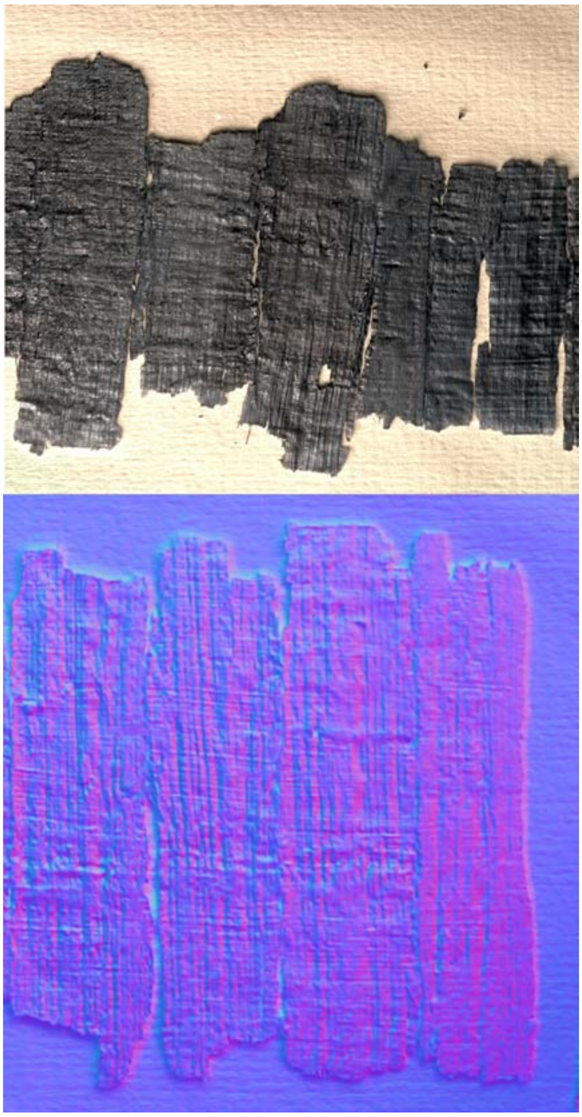


Figure 2: Detail of Derveni Papyrus fragment, RTI visualization emphasizing damage and alterations of the papyrus material

handwritings of the DP and draw attention to the alpha with horizontal crossbar, archaic epsilon and sigma, as well as the omega with convex central curve and the zeta with parallel outer horizontal strokes and theta with its central element reduced to a point (Cavallo, 2009). Undoubtedly RTI can provide an advanced alternative option for the graphic characteristics of handwriting and moreover can assist in the determination of the direction of the writer hand, by highlighting the three dimensional characteristics of letters. Consequently the use of RTI visualizations instead of digital images is highly recommended for automated and semi-automatic manuscript analysis. There is a strong potential in the development of systems able to assist the study of the papyri by incorporating the useful features such as those presented in the 'Toolbox for Manuscript analysis' (Gau *et al.,* 2009) with the ability of RTI to enhance readability obscured texts and assist surface analysis.

2.2.2. IR-RTI

Capturing data beyond the visible using hyper-multispectral imaging signals the most valuable contribution in papyrology studies so far. There are numerous examples of successful applications of multispectral and FCIR Imaging because of the ability of multispectral imaging to enhance readability and assist material characterization. Multispectral and false colour imaging, especially under near infrared radiation, of the highly deteriorated and fragile Daphne papyrus improved the readability of the text. Also, it made possible the differentiation of ink and damage, which were invisible to the naked eye (Alexopoulou *et al.,* 2013). Material differentiation is also possible based on the properties of the inks. In the case of the Egyptian papyrus examined in the Brooklyn Museum, the carbon black text absorbs infrared radiation and appears intense. On the contrary, the iron oxide red pigment is faded due to the low absorption of iron in the infrared spectral region (Kapetanakis, 2010). Especially in the case of charred papyri the application of multispectral

Figure 3: Detail of Derveni Papyrus fragment, RTI visualizations in multi-light (left) and specular enhancement rendering mode (middle, right).

imaging gave impressive results and enables the reading of black letters on the equally black papyrus material, as demonstrated by the Herculaneum and the Petra Papyri (Ware *et al.,* 2000), as well as the earlier application on the Derveni papyrus.

Infrared RTI meets the need for experiencing the third dimension beyond the visible. It is complementary to normal RTI technique with great potential for a broad area of materials and artefacts types, including painted surfaces, such as pained panels, canvas, documentary artefacts, engraved stones in poor conservation condition, restored glass artefacts, painted or engraved ceramics with encrustations, mortars, paper and archival material. The experimentation for the simultaneous capture of both spectral and textural by Hewlet-Packard and Cultural Heritage Imaging (CHI) researchers reached the conclusion that the technique is feasible (Redman, 2007; Redman & Mudge, 2007). A combined RTI and infrared methodology was used for the study of the Dead Sea Scrolls, in order to reveal the text and examine the condition of the parchment. This successful project demonstrated a promising methodological approach for the study of documentary artefacts but also the necessity for further experimentation (Caine & Magen, 2011).

In the present case study Infrared Imaging within the abilities of digital cameras (circa 700-1000 nm) was tested in order to obtain a record of how the papyrus reflects varying amounts of the infrared radiation falling on it. Images were captured with a UV-VIS-IR modified digital SLR camera, along with a series of filters that block visible and transmits infrared above 720, 760, 850 and 950 nm. Taking into consideration the artefact's deterioration due to thermal ageing infrared light emitting diodes (LED) are recommended for infrared RTI acquisition of papyri. The deterioration by IR radiation, discussed in literature, is strongly connected to the raise in temperature of the

Figure 4: Details of Derveni Papyrus fragment, RTI visualizations, comparison of renderings in visible (left) and IR spectral area (right).

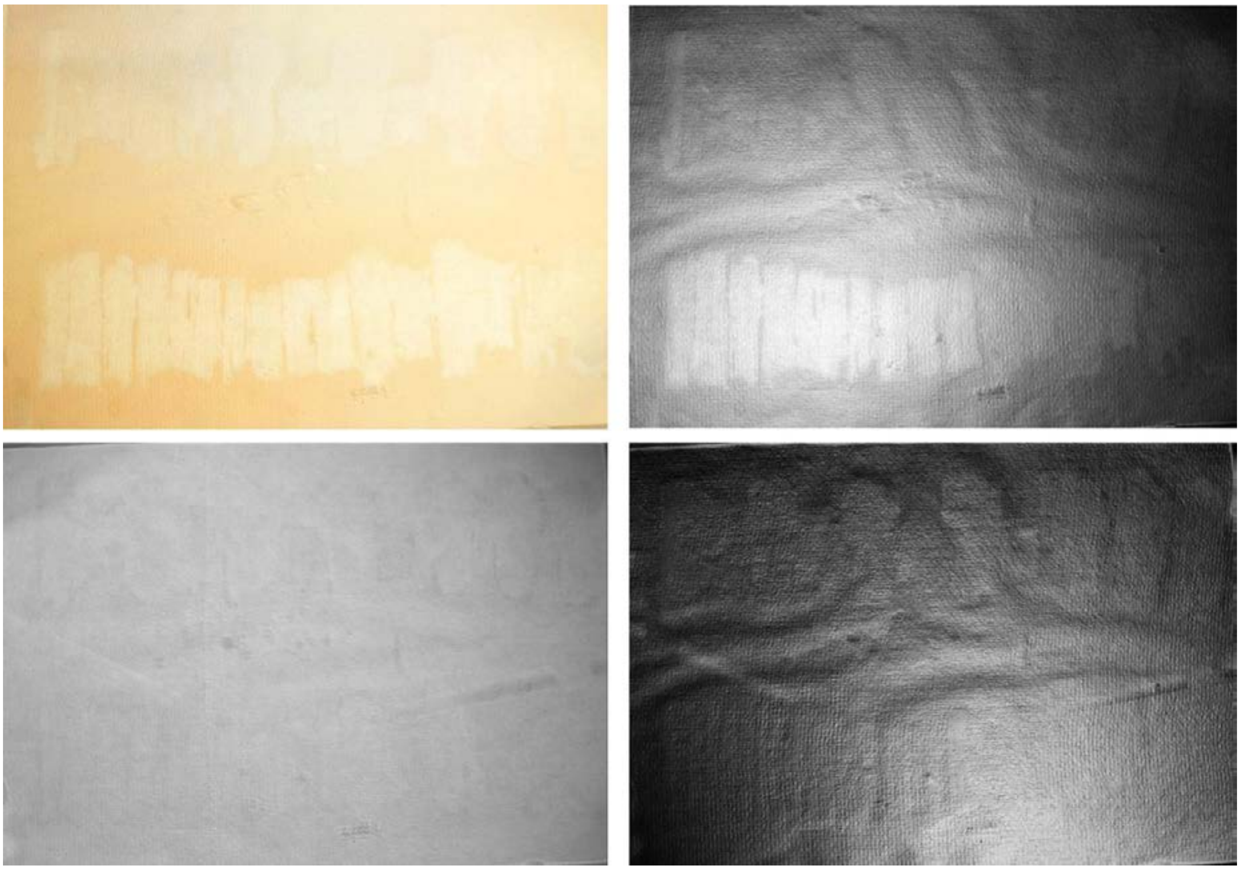


Figure 5: Derveni Papyrus fragment, comparison of digital images (left) and RTI visualizations (right) in visible (top) and IR spectral area (below).

artefact under examination. According to the Canadian Conservation Institute, IR radiation heating properties of high intensity incandescent lamps and direct sunlight can lead to temperature above 40°C. This results in a raise of thermal decay factor of 20, when chemical and physical damage as well as biological deterioration phenomena can occur. It is particularly harmful for collections of textiles, paper, and leather, photographic materials, rubber, and plastics (Michalski, 2010). Consequently, we omitted the use of such IR sources. The LED flashes were equipped with 36 and 47 LEDs each.

In the Derveni Papyrus RTI case study the IR-RTI emphasises the three dimensional characteristics of the object and consequently assists the Derveni papyrus examination, while the texts can be examined more efficiently. The enhanced view of the papyrus fragment under examination, provided by RTI and IR imaging, was significantly improved by the implementation of IR-RTI. The legibility of the writings is particularly enhanced as shown in figure 4. Also, the three dimensionality of the papyrus is emphasized. Via IR-RTI, the distortion of the both papyrus and its backing on the verso can be documented and examined (Figure 5). This distortion caused probably by the backing material should be monitored so as to avoid further fractures of the papyrus fragment. The evaluation of IR-RTI visualizations was based on the visual comparison between visible and infrared spectral regions. The latest release of RTI Viewer software includes useful features for comparative visual inspection of RTI. The numeric parameters values are displayed in the screen, enabling the accurate description and reproduction of a specific view of the visualised object. The visible images were converted to monochrome and re-processed so as to perform the comparisons without the confusing colour information, which may had an impact on the conclusions reached.

2.2.3. Trans illumination and irradiation RTI

Transmitted imaging (TI) techniques, proposed for the first time in the 1970s and further developed during the 1980s, are considered as standard photographic techniques for conservation documentation examination and condition reporting of translucent materials, canvas paintings, mounded papyri, photographic material, and works of art on paper or archival material. In TI the camera records the radiation transmitted through the investigated surface (Cornelius, 1977; Kushel, 1985; Cucci *et al.,*, 2012). The use of light boxes is common in conservation of paper and archival material. For the current study transmitted RTI was tested in an attempt to provide an enhanced RTI visualization, complementary to reflected RTI. This was the first known implementation of such an approach.

The transmitted RTI is created from a series of images captured with transmitted illumination and/or radiation and a lighting positions file. The latter can be obtained by pre-capturing a sphere from specific lighting positions, which can be re-used for capturing the RTI dataset images. This process requires accurate marking of the lighting positions. It is very time-consuming and mismarking the lighting positions is very easy to occur. The use of auxiliary devices for RTI data capture can be helpful. In the case of artefacts where only low angle lights are employed, such as glass mounded works of arts (most commonly found in museums collections are the papyri and paintings), placing the object on a Plexiglas base and moving the light around the object at the height of the object and slightly beneath enables the acquisition of a set of transmitted photographs.

Trans illumination RTI provides an enhanced RTI visualization, complementary to reflected RTI (Figure 6). Particularly it proved useful for the visualization of the back side of the papyrus fragments which was covered by the backing material. The features highlighted were the fractures and incorrectly aligned fibres, as well as fibre structure and joins between sheets of papyrus.

3. Virtual unrolling

Although Kapsomenos expressed his admiration to Fackelmann's work and commented on the 'good fortune' of Prof. Hunger's visit in the Archaeological Museum of Thessaloniki, after which the papyrus had been unrolled (Kapsomenos 1963, 4), the methodological approach followed may have been criticised negatively if executed by a contemporary conservator. Nowadays digital technology overcomes the limitations of manual conservation and provides an alternative to traditional interventive operations. Successful applications of virtual unrolling using micro tomography have been presented in the bibliography. By using high resolution micro-CT scanning (micro focus or nanofocus X-ray sources or synchrotron light), scientists can overcome problems associated with the physical manipulation of vulnerable finds and enable virtual unrolling and non-invasive study of the composition of inks (Baumann *et al.,,* 2008). X-ray tubes with low focal spots operate at low current and the voltage does not exceed 150 kV (Bradley & Creagh, 2006). Apart from documentary artefacts CT scanning - 3D tomography has been applied to pots filled with coins (Casali, 2006; Miles *et al.,*, 2012), soil blocks (Ebinger-Rist, N. *et al.,,* 2011) and mummies (Baumann *et al.,,* 2008).

3D tomography enables better preservation of archaeological material, avoiding actions which may threaten their material integrity, so it is considered as a preventive conservation measure. The conservator may choose not to execute the physical remedial conservation operation, if the data revealed from 3D scanning point out that the material is in bad conservation state and cannot withstand physical manipulation. This is of great significance in the case of papyri, because of their vulnerability. Another crucial point is that it enables evidence-based conservation decision making, avoiding judgements based on general knowledge. Moreover, from an economic point of view, 3D tomography is cost-effective, compared to the interventive conservation approach. According to the findings of the virtual unrolling of the roll, museum professionals may take appropriate actions. Among them there is the option not to devote the amount of time and skills in cases where digital micro excavation revealed the absence of valuable contents.

4. Conclusions and Future Work

According to Lehoux in the not-too-distant future RTI would be the standard means of consulting ancient inscriptions of all sorts (Lehoux, 2013).The present paper strengthens such an approach by providing a new RTI methodology for non-destructive diagnostic examination and study of mounted papyri. RTI enhances the readability of the texts and overcomes problems such as buckling and edges identification. RTI in the infrared spectral region is particularly useful in case of blackened papyri in carbonised state, as demonstrated by the IR-RTI renderings of the DP fragments. Trans irradiation RTI approaches can assist the detection of writings covered by backing materials. Also, RTI visualization enables the study of papyri without any interaction with the fragile papyrus material. This is of great importance because of the vulnerability of the papyri and the necessary manipulation involved in conventional artefacts examination. Virtual study of the papyrus not only protects its material integrity but is also more effective than physical examination because of the ability to virtually relight and apply different filters using the available rendering modes so as to create the appropriate visualizations.

The potential of RTI applications in papyri is not limited in the enhanced legibility. RTI can successfully overcome the problems caused by the presence of the mounts regarding the visual analysis. Moreover, it is a powerful condition-reporting and documentation tool. For example joins, repairs, the state of preservation of the inks, misalignments, missing fibres and patterns can be emphasized in RTI visualizations. The application of microscopic RTI is a promising technique for the study of the papyri. It is expected based on research in other materials that the results of microscopic examination of abraded surfaces of palimpsests and corrosive inks would be enhanced by micro RTI.

The synergy of RTI, Infrared imaging and trans illumination-irradiation and enhances the readability of texts and assists the overall diagnostic examination of the papyrus, the examination of the manufacture techniques, the determination of its condition and its possible conservation needs. The resulting different views of the Derveni papyrus increase our understanding of its artistic, historic and cultural heritage significance and are useful for documentation, presentation, and communication and research purposes. Considering the low cost of the necessary equipment, the freely available software for processing and viewing *rti and/or *ptm files, the low level of expertise necessary for data capture in addition to the flexibility of RTI technology, which can adopt different set ups in order to cover the visualization needs of artefacts

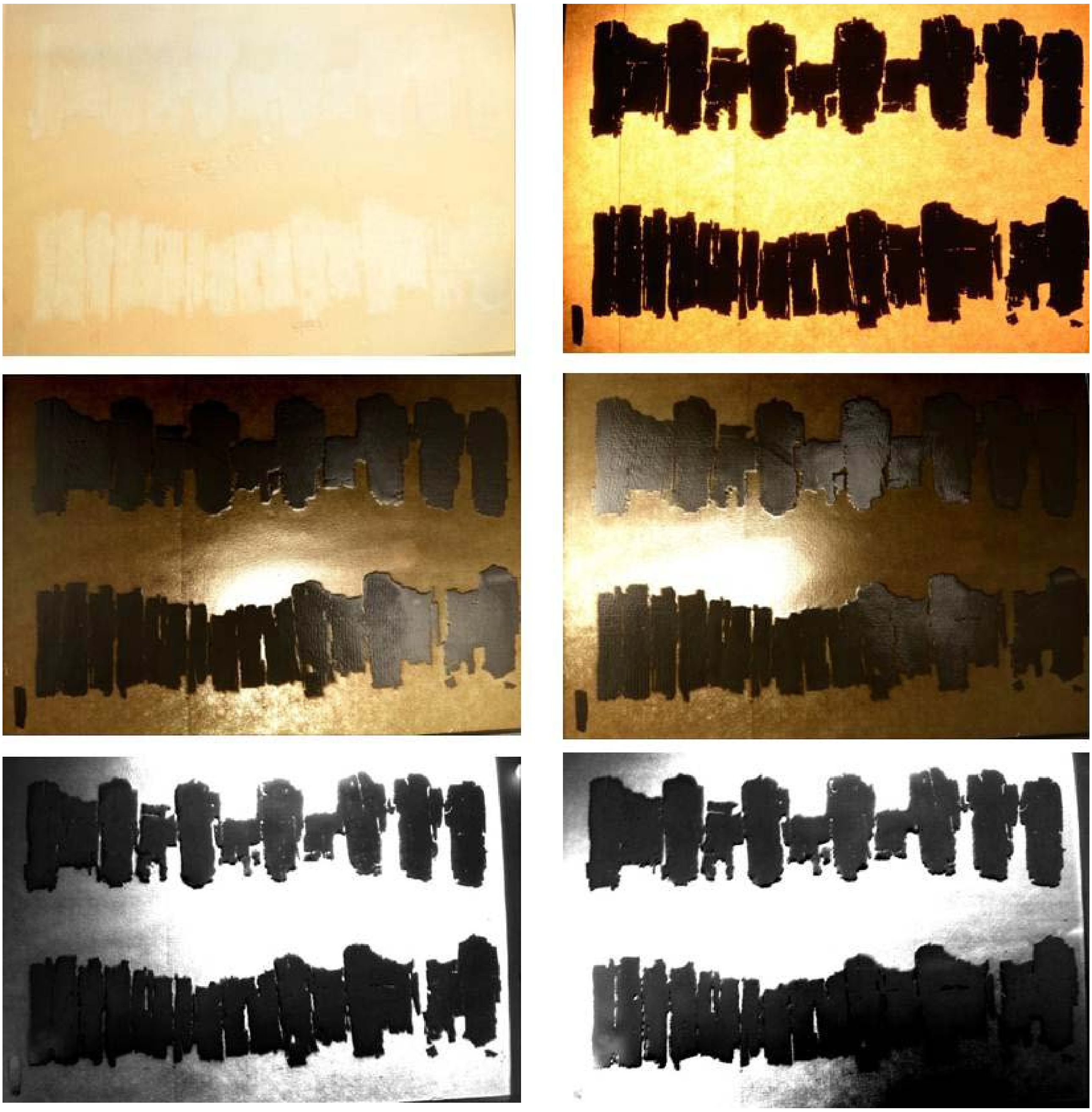


Figure 6: Derveni Papyrus fragment, digital image (top left) and trans illumination image (top right), trans illumination RTI renderings (middle) and trans irradiation RTI renderings (below).

of different type, there is no doubt that RTI is a very useful methodology for documentation and recording. The application of the technique in the infrared spectral area demonstrate RTI's potential in diagnostic documentation and condition reporting, while trans illumination and irradiation RTI makes possible the visualization of features hidden behind backing materials.

The enhanced readability of RTI snapshots compared to digital images offers the opportunity for enhanced optical character recognition of ancient texts, especially those on bad preservation state. The increased interest of the scientific community for the development of OCR software capable of reading ancient Greek and the modern text mining approaches used in combination with the advanced RTI visualizations can assist the study of ancient texts.

From this case study a series of research questions evoked, such as the impact of the glass mount on the final visualization, the degree of blackening that can be visualised appropriately using near infrared imaging equipment, the perfect combination between ambient and transmitted light in trans RTI. An on-going research project in the Archaeological Computing Research Group in co-operation with V. Valergas, MSc student, addresses the application of the proposed integrated RTI methodologies using a series of on purpose build replicas as a case study. The papyrus replicas have undergone different levels of thermal ageing resulting in different degrees of blackening as well as deformations of the material. The application of the proposed methodologies on these replicas will enable us to define their effectiveness in the reading of the texts and gain a better understanding of the limitations and potential of the techniques. In the meanwhile there is a strong interest of the academic community in testing the methods introduced in this paper in other papyri, whose study is problematic using conventional methodology. A related area of research funded by the AHRC RTI FOF project is exploring means to display and annotate multiple RTI datasets in order that the comparison and integration of approaches can be made simpler, more effective and easier to share and publish.

Acknowledgments

The authors gratefully acknowledge the Archaeological Museum of Thessaloniki (Hellenic Republic-Ministry of Culture and Sports) and would like to express many thanks to Dr. Despina Ignatiadou.

References

ALEXOPOULOU, A. A., KAMINARI A.-A., PANAGOPOULOS A., PÖHLMANN E. (2013). Multispectral documentation and image processing analysis of the papyrus of tomb II at Daphne, Greece. *Journal of Archaeological Science* 40(2), pp.1242–1249. Available at: http://linkinghub.elsevier.com/retrieve/pii/S0305440312004153 [Accessed January 16, 2013].

BANIK, G. & STACHELBERGER, H. (1987). Salt-Migration in papyrus fragments. In: *Recent advances in the conservation and analysis of artifacts: Jubilee Conservation Conference, London 6-10 July 1987*. London: University of London. Institute of Archaeology. Summer Schools Press, pp. 199–202.

BAUMANN, R., PORTER, D.C. & SEALES, W.B. (2008). The use of micro-CT in the study of archaeological artifacts. In *9th International Conference on NDT of Art.* Jerusalem Israel, pp. 25–30.

BRADLEY, D. & CREAGH, D., (2006). *Physical Techniques in the Study of Art, Archaeology and Cultural Heritage* (1st ed.), Amsterdam: Elsevier.

BULOW-JACOBSEN, A. (2009). Writing materials in the ancient world. In: *The Oxford Handbook of Papyrology.* New York: Oxford University Press, pp. 3–29.

CAINE, M. & MAGEN, M. (2011). Pixels and Parchment: The Application of RTI and Infrared Imaging to the Dead Sea Scrolls. In *EVA London*. Available at: http://www.eva-london.org/past-eva-londons/2011. [Accessed July 7, 2014]

CAPPITELLI, F. & SORLINI, C. (2005). From papyrus to compact disc: the microbial deterioration of documentary heritage. *Critical reviews in microbiology,* 31(1), pp.1–10. Available at: http://www.ncbi.nlm.nih.gov/pubmed/15839400 [Accessed July 7, 2014].

CASALI, F. (2006). X-ray digital radiography and computed tomography for cultural heritage. *Archeometriai Mühely* 1(1), pp.24–28.

CAVALLO, G. (2009). Greek and latin writting in papyri. In *The Oxford Handbook of Papyrology.* New York: Oxford University Press, pp. 101–148.

CORNELIUS, F.D. (1977). Correspondence: Transmitted infrared photography. *Studies in conservation* 22(1), p.42.

CUCCI, C., PICOLLO, M. & VERVAT, M. (2012). Trans-illumination and trans-irradiation with digital cameras: Potentials and limits of two imaging techniques used for the diagnostic investigation of paintings. *Journal of Cultural Heritage* 13(1), pp.83–88. Available at: http://linkinghub.elsevier.com/retrieve/pii/S1296207411000653 [Accessed December 17, 2012].

Cultural Heritage Imaging, (2011) *Guide to highlight image processing. Cultural Heritage Imaging*. Available at: http://culturalheritageimaging.org/What_We_Offer/Downloads/rtibuilder/RTI_hlt_Processing_Guide_v14_beta.pdf [Accessed April 3, 2012].

Cultural Heritage Imaging, (2012) *Cultural Heritage Imaging.* Available at: http://culturalheritageimaging.org/Technologies/RTI/ [Accessed January 18, 2013].

Cultural Heritage Imaging, (2013a) *Guide to highlight image capture v2.0. Cultural Heritage Imaging,* p.32. Available at: http://culturalheritageimaging.org [Accessed January 19, 2014].

Cultural Heritage Imaging, (2013b). *Guide to RTIViewer, version 1.1. Cultural Heritage Imaging.* Available at: http://culturalheritageimaging.org/What_We_Offer/Downloads/rtiviewer/RTIViewer_Guide_v1_1.pdf [Accessed July 2, 2014].

EBINGER-RIST, N., C., PEEK, J. & STELZNER, F.G. (2011) Computed tomography: a powerful tool for non-destructive mass documentation of archaeological metals. In P. Mardikian *et al.,*, (eds.) *Metal 2010-Interim Meeting of the ICOM-CC Metal Working Group.* Charleston, pp. 342–347.

FROSEN, J. (2009). Conservation of ancient papyrus material. In: R. S. Bagnall, (ed.) *The Oxford Handbook of Papyrology*. New York: Oxford University Press, pp. 79–100.

GAU, M., VILL, M., KLEBER,F., DIEM, M., MIKLAS, H. & SABLATNIG R. (2009). A Toolbox for Manuscript Analysis. In B. Frischer, J. W. Crawford, & D. Koller, (eds.) *37th International Conference on Computer Applications and Quantitative Methods in Archaeology*. Williamsburg, Virginia, USA, Oxford: Archaeopress, pp. 86–92.

GRAF, J. & KRUTZSCH, M. (2008). *Ägypten lesbar machen - die klassische Konservierung/Restaurierung von Papyri und neuere Verfahren*, Berlin: Walter de Gruyter GmbH.

JANKO, R. (2002). The Herculaneum Library: Some recent developments. *Estudios Clasicos*, 121, pp.25–41.

JOHNSON, E.V. & HORGAN, C.J. (1979). *Museum collection storage No 2*, Paris: Unesco.

KAPETANAKIS, P. (2010) *Brooklyn Museum: IR and UV Examination of Egyptian Papyrus. Brooklyn Museum.* Available at: http://www.brooklynmuseum.org/community/blogosphere/2010/12/03/ir-and-uv-examination-of-egyptian-papyrus/ [Accessed January 16, 2013].

KAPSOMENOS, S.G. (1963). The Orphic Papyrus Roll of Thessalonica. *The Bulletin of the American Society of Papyrologists* 2(1), pp.3–22.

KUSHEL, D.A. (1985). Applications of transmitted infrared radiation to the examination of artifacts. *Studies in conservation* 30(1), pp.1–10.

LAU-LAMB, L. (2005) Advanced Papyrological Information Systems - Guidelines for Conservation of Papyrus, Available at: http://www.lib.umich.edu/papyrus-collection/advanced-papyrological-information-system-guidelines-conservation-papyrus#_ftn1.

LEACH, B. (2008). Papyrus conservation at the British Museum. In: J. Graf & M. Krutzsch, (eds.) *Ägypten lesbar machen - die klassische Konservierung/Restaurierung von Papyri und neuere Verfahren.* Berlin: Walter de Gruyter GmbH.

LEACH, B.. & TAIT, I. (2000). Papyrus. In: P. T. Nicholson & I. Shaw, (eds.). *Ancient Egyptian Materials and Technology.* Cambridge, UK: Cambridge University Press, pp. 227–253.

LEHOUX, D. (2013) Ancient Science in a Digital Age. *Isis,* 104(1), pp.111–118. Available at: http://www.jstor.org/stable/10.1086/669892.

MACHTELD VAN DER FELTZ (2010). De conservering van de papyruscollectie in de Groningse Universiteitsbibliotheek. *RN M@gazine*.

MALZBENDER, T., GELB, D., WOLTERS, H., ZUCKERMAN, B. (2004) Enhancement of Shape Perception by Surface Reflectance Transformation. In B. Girod, Marcus A. Magnor & H. P. Seidel (eds.). *Proceedings of the Vision, Modeling, and Visualization Conference 2004 (VMV 2004), Stanford, California, USA, November 16-18, 2004, Aka GmbH*. pp. 183–186.

MICHALSKI, S., (2010) *Agent of Deterioration: Light, Ultraviolet and Infrared*. Canadian Conservation Institute. Available at: http://www.cci-icc.gc.ca/resources-ressources/agentsofdeterioration-agentsdedeterioration/chap08-eng.aspx [Accessed July 3, 2014].

MILES, J., (2012) Computed Tomography scanning of Roman Coins. *Archaeological Computing Research Group Blog*. Available at: http://acrg.soton.ac.uk/blog/1168/ [Accessed November 14, 2014].

OWEN, A. & DANZING, R., (1993). The History and Treatment of the Papyrus Collection at The Brooklyn Museum. In *Book and Paper specialty group session, AIC 21st Annual Meeting*. Denver, Colorado: AIC. Available at: http://cool.conservation-us.org/coolaic/sg/bpg/annual/v12/bp12-10.html.

POLLARD, M., NEATE, S. & DECOUX, S. (2011). An Investigation of the Extent and Causes of Salt Migration in the Papyrus and Related Collections in the Bodleian Library. Restaurator. *International Journal for the Preservation of Library and Archival Material*, 32(2), pp.135–139.

REDMAN, J. (2007). Advances in Digital Imaging for Fine Art and Cultural Heritage. In *NIP23: International Conference on Digital Printing Technologies and Digital Fabrication 2007,* Springfield: Society for Imaging Science and Technology, pp. 355–363.

REDMAN, J. & MUDGE, M. (2007). The simultaneous capture of spectral and textural information. In S. A. Stovall, (eds.) *Archiving 2007: final program and proceedings,* May 21-24, 2007, Arlington, Virginia. Springfield: Society for Imaging Science &Technology, pp. 2–5.

SHEPHERD, W.H. (2008). The Preservation and Conservation of Papyri: *MLIS 7710, Archival Theory and Issues*, pp.1–28. Available at: https://www.yumpu.com/en/document/view/11403310/preservation-and-conservation-of-papyri-comcastnet.

STANLEY, T. (1994). Papyrus Storage at Princeton University. In *The Book and Paper Group Annual 13*. Washington, D.C.: American Institute for Conservation of Historic and Artistic Works (AIC).

TAYLOR, J., LEACH, B. & SHARP, H., (2011). The history and conservation of the papyrus of Tuy. T*echnical Research Bulletin*, 5, pp.95–104.

THEMELIS, P. & TOURATSOGLOU, I., (1997). Tafoi tou Derveniou, Athens: TAPA.

VANDERPOOL, E., (1962). News letter from Greece. *American Journal of Archaeology,* 66(4), pp.389–391. Available at: http://www.jstor.org/stable/502026.

WARE, G. A., CHABRIES, D.M. ,CHRISTIANSEN, R.W., MARTIN, C.E. (2000) Multispectral document enhancement: ancient carbonized scrolls. In *Proceedings of the Geoscience and Remote Sensing Symposium. IGARSS 2000. IEEE 2000 International*, 6, IEEE pp.2486–2488. Available at: http://ieeexplore.ieee.org/lpdocs/epic03/wrapper.htm?arnumber=859615.

WRIGHT, M. (1983) A Method of Extracting Papyri from Cartonnage. *Studies in Conservation*, 28(3), pp.122–126.

Chapter 3. Ontologies and Standards

Towards Linked-Data in Numismatics: How the DIANA Approach can Improve the Diachrony Integrating Heterogeneous Pieces of Data

Maria CALTABIANO,[1] Mariangela PUGLISI,[1] Antonio CELESTI,[2] Grazia SALAMONE[1]

[1] DICAM, University of Messina, Italy
[2] DICIEAMA, University of Messina, Italy
caltabia@unime.it, mpuglisi@unime.it, gsalamone@unime.it, acelesti@unime.it

Abstract
The Digital Iconographic Atlas of Numismatics in Antiquity (D.I.A.N.A.) aims to perform an in-depth analysis of the coin iconography according to time and space through digital maps. The project has been designed with an eye toward the possibility of a future integration with other digital archives. The basic idea is to enable DIANA to analyze, besides ancient mints and coin iconographies, even other related data. An objective of DIANA is to link and integrate its datasets with others coming from different digital archives in order to improve a particular study aggregating different pieces of information to better understand the diachrony and the cultural context. In this paper, we analyze the main aspect involved for the integration of DIANA with other digital archives in a scenario of linked data.

Keywords: ancient coins, iconography; diachrony, digital maps, linked data

1. Introduction

Nowadays, in the field of science of antiquity, there are a lot of digital archives and web applications that allow displaying ancient entities and artifacts on digital maps. Entities may include, e.g., gods, abstract personification, historical personages, etc. Artifacts may include coins, weapons, architecture, etc. In particular, ancient coins represent also out-and-out documents that need to be properly studied and analyzed. From the point of view of numismatics, there are not so many web applications enabling researchers to perform an in-depth analysis of the coin iconography.

The *Digital Iconographic Atlas of Numismatics in Antiquity* (DIANA) aims to fulfill such a gap. DIANA is part of the LIN project, aimed at compiling the *Lexicon Iconographicum Numismaticae Classicae et Mediae Aetatis*, a dictionary whose entries record all the principal and secondary images found on ancient and medieval coins. It is a web platform that allows researchers to analyze the 'coin iconography' according to time and space through digital maps. The DIANA's digital archive is based on a relational Data Base Management System (DBMS). The web application is developed combining both server-side and client-side programming languages. The server-side is developed using the PHP language, whereas the client-side is developed using JavaScript. In order to provide users a good degree of reactiveness, DIANA has been developed adopting the Asynchronous JavaScript and XML (AJAX) programming technique. In order to build digital maps, the system uses the Cloud Computing Google Maps Platform as a Service (PaaS). A mint can be searched on DIANA considering a target coin iconography through a web form. By means of AJAX requests, data are retrieved on the DIANA's digital archive and they are sent in XML format to the user's web browser. After that, the web browser processes the received data and it forwards a second AJAX request to the Google Map PaaS that return a digital map displaying the mint and ancient coins.

With DIANA it is possible to study the 'diachrony' with a new innovative approach starting from ancient coins. The project has been designed with an eye toward the possibility of a future integration with other digital archives. The basic idea is to enable DIANA to analyze, besides ancient mints and coin iconographies, even other related data. Currently, in DIANA, places are located according to ancient mints. In other digital archives places can be located according to ancient artefacts, locations, and names, as well as in Pleiades (http://pleiades.stoa.org). An objective of DIANA is to link and integrate its datasets with other ones coming from different digital archives in order to improve a particular study aggregating different pieces of information. For example, an artefact such as a 'phiale' (a pot used to make sacrifices) depicted on a coin coming from DIANA, can be related to artifacts coming from other digital archives. Such a data linking enables researchers to better understand the diachrony and the cultural context. Linking data of different digital archives it is possible to exploit different data integration techniques, e.g., importing datasets in XML, JSON, CSV, KML or RDF format, or by means of web service (e.g., REST, SOAP, etc) interactions.

In this paper, we discuss the main aspects involved in the design of a software system able link heterogeneous data coming from different digital archives. In particular, subsequently we have highlighted the advantages of integrating heterogeneous data related to ancient artifacts, we will discuss how such integration can be possible from a technical point of view considering the DIANA system. The rest of the paper is organized as follows. Section 2 discuss related works. A discussion regarding the approach used in DIANA to classify the iconography of ancient artifacts is presented in Section 3. A few technical details about how the DIANA system works are discussed in Section 4. In Section 5, we discuss the advantages of linking data related to ancient artifacts coming from different digital archives. In particular, we will discuss how it is possible to design a system for linked data using DIANA, considering as a study case a possible integration with Pleiades (http://pleiades.stoa.org/), a digital archives storing data on ancient places, locations, and names along with Geographical Information System (GIS) metadata. Section 6 concludes the paper with lights to the future.

2. Related Works

In the recent years, many efforts have been conducted to represent and distribute knowledge about archeological artifacts through the Web. In the field of Numismatics, several sites exist, that allow to visualize, query and share data about ancient coins. In this Section, we briefly describe the most significant web applications that provide access to ancient coin collections, in order to give a background to our work.

The American Numismatic Society (ANS, 2014) at present publishes more than 600 thousands objects from its collection, giving access to their information through a project named MANTIS (MANTIS, 2014). Accessing the website of MANTIS, it is possible to query a money database, specifying search words for all the attributes associated to the object, e.g., authority, mint, locality, material and so on. For many of the coins represented, two images for obverse and reverse types are given, otherwise only a brief description is displayed. An interesting feature of the site is the possibility of comparing coins (or money in general) on the basis of their attributes. MANTIS also provides a map representation of coin distribution, by visualizing the result of queries on a map provided by the Google Maps service (Google Maps, 2014): data related to the coins are not displayed directly on the map, but the user can follow a link that opens a description frame on the bottom of the page. Moreover, the online numismatic catalogue of Staatliche Museum (Staatliche Museum Numismatic Catalogue, 2014) in Berlin has about 20,000 objects (coins and medals), and it is possible to search coins on maps, although with a very simple interface and limited interactivity. Other websites aim to sotre data about coins of specific historical periods, as the Roman Provincial Coinage Project (Roman Provincial Coinage Project, 2014). Also in this case, the maps are used to give a simple representation of the cities related to the coins. The same happens for OCRE (Online Coins of the Roman Empire) (OCRE, 2014), a joint project of the American Numismatic Society and the Institute for the Study of the Ancient World at New York University to create a digital corpus of the coinage of the Roman Empire. OCRE is a numismatic tool based on the stable numismatic identities established by the Nomisma project (NOMISMA, 2014) which intends to host the Numismatic Description Standard (NUDS), a set of suggested field names for recording numismatic information in a column-oriented database (NUDS, 2014). NUDS has not yet expanded to define the contents of fields and does not suggest how to describe coin types. The space dedicated to the coin Obverse or Reverse types is in fact a free-text container, holder for undifferentiated description.

Other web sites use a map representation (Bracey, 2009; Amandry & Bateson, 2009) to display money data, even if the location of ancient places is often imprecise and uncertain, since modern cities are not always located in the same place of their ancient corresponding place. A very significant effort to face this issue is represented by Pleiades (Pleiades, 2014). Pleiades gives the possibility to create, query, share and represent historical geographic data in digital form, with an extensive coverage of Greek and Roman worlds. This is indeed a helpful tool for researchers, since this project promotes an open community with the intent to share information of historical geographic data.

Related to this, the Nomisma project (NOMISMA, 2014) is a very interesting effort in the direction of providing stable digital representations for numismatics concept and entities, both for geographical and money-related data. The use of Semantic Web principles (Berners & Hendler & Lassila, 2001), KML (KML, 2013) language for geographical data and XML for representing the resources of the project, has the explicit intent to encourage querying, sharing and reusing of these data.

3. The DIANA Approach

Coins, as public and official documents of a homogeneous nature, are ideal for identifying the subjects and themes of the figurative culture of antiquity. Through coin images the issuer achieves a process of clear and concise communication, directed at a wide public. The presence of one or the concentration of more iconographic subjects, helps us to view coin images as historical documents of the cultural environment which produced them, and to identify culturally homogeneous geographical areas.

The wide-ranging collection of data, performed on documents which cover vast periods of time and geographical areas, makes it possible to verify phenomena of continuity or irregularity in the meaning of the coin iconographies and their belonging to categories which are also defined in terms of cultural ‘continuity’ and ‘discontinuity’. The Digital Iconographic Atlas of Numismatics in Antiquity (DIANA, 2014) was designed to reconstruct ‘the history of the coin type’, or the ‘stratigraphic representation’ of its meaning, which strips the structure of iconic bare language and, using a multidisciplinary method, retrieves the relationship between the image, the realities which they represent, and that between the images and the cultural context in which they are used. Moreover, DIANA is part of the LIN project, aimed at compiling the Lexicon Iconographicum Numismaticae Classicae et Mediae Aetatis (Universities of Messina, Bologna, Genoa and Milan), a dictionary whose entries record all the principal and secondary images found on ancient and medieval coins.

DIANA allows researchers to analyze ‘coin iconography’ according to time and space through digital maps. Our archive highlights the geographical distribution of the issuing cities and their diachronical activity in striking coins in order to:

- create a digital INDEX of the ancient Mediterranean mints (Greek and Roman Periods);
- codify their NAMES;
- display their TOPOGRAPHICAL DISTRIBUTION on digital maps;
- identify the PERIODS OF MINTING ACTIVITY;
- simplify the INTERPRETATION OF COIN ICONOGRAPHY;
- reconstruct the RELATIONS AND INFLUENCES BETWEEN PEOPLES in the Mediterranean.

Compared to other existing web applications, strength is the standardization of all the 'voices' to be used for the description of the coin iconographies. The entries are the types surveyed within the Lexicon Iconographicum Numismaticae (Universities of Messina, Bologna, Genoa and Milan), including Personages, Animals/Monsters, Flora and Objects. All the other search options are standardized.

The creation of the digital Atlas DIANA, referring to the figurative culture of ancient coins presented on diachronic maps, is useful for the following reasons:

- it simplifies the interpretation of meanings of coin iconography to promote and increase awareness of the numismatic heritage;
- it views coin images as historical document of the cultural environment which produced them and identify culturally homogeneous geographical areas;
- it reconstructs the relations, influences and relationships between peoples on the basis of the 'journeys' of coin images in the Mediterranean;
- it offers to students and researchers of other fields of the Ancient Age an overview of the distribution of the iconic subjects and of cultural areas to which they belong.

The DIANA's iconographic classification is based on a General Index (Salamone & Caltabiano, 2007) created by consulting the indexes of the main catalogs of Greek and Roman coinage, and for mythological figures by considering the documents already collected in Lexicon Iconographicum Mythologiae Classicae (LIMC). The coin types have been organized into four large macro-categories or groups. Group I includes human figures: individual mythical figures (Dei, heroes, heroines, etc), personifications of abstract concepts (Nike, Victoria, Tyche, different Virtutes as Aequitas, Iustitia, Pietas, etc), personifications of geographical entities (Cities or Rivers), or legal bodies (Demos, Koinoboulion, etc), but also common types (athlete, heros equitans, etc), considered as generic because they lack a specific proper name or because they have not yet been identified. For each subgroup it is important to highlight the denotative iconic elements, which are indispensable for recognizing each deity (e.g., which elements identify Aphrodite? and which aspects, found in the connotative iconic elements, does Aphrodite have in common with Hera, Demeter or another personage?). Group II includes animals and fantastic creatures. Group III includes flora (trees, flowers, fruits). Group IV contains res: architectural types, arms, astronomic images, honoris et imperii signa, fishing and hunting tools, musical instruments, ships, vases, etc.

There is also the macro-category Themata that includes coin types, mainly Roman ones, which do not fall into any of the four afore mentioned groups due to their complexity. They are often identified by their coin legend, e.g., adventus, consecratio, gaudium, ludi, nuptiae, princeps iuventutis, propago imperi etc, and in most cases these are complex scenes.

Figure 1 shows the possible codification of the mythical figure of Apollo. For further details about the codification of the other groups see (Salamone & Caltabiano, 2007).

Figures 2 and 3 show examples of output on digital maps. Figures 2 and 3 depict the mints that have coined coins whose iconography present a female personage, standing, with a 'phiale' (a pot used to make sacrifices). The common iconographic scheme is related to the standing figure almost always at an altar, dressed, holding the 'phiale' in

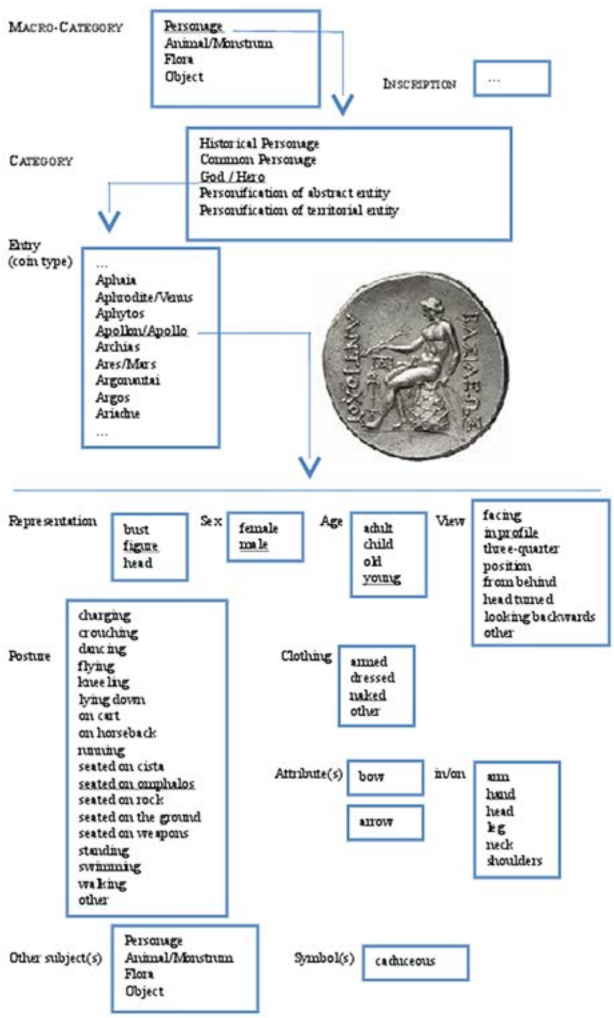


Figure 1: Example of possible 'codification' of the mythical figure of Apollo.

her hand. It is the iconography of libation, well codified and documented in ancient imagery also on other media (ceramics, reliefs, statues) in relation to Nikai figures or figures of uncertain identity. In the case of money it is possible, however, to reconstruct the identity of these figures through the name that sometimes accompanies them and thanks to the typological context in which they appear.

The diachronic distribution of the data, highlighted by the colors of the markers (from white to blue), shows that invention of the type of the female figure performing a libation belongs to Western Sicily and to the category of 'nymphs' eponymous (Simon, 1998; Salamone,, 2013): at Himera the 'nymphs' appears around 465 BC; shortly after, around 460 BC, the scheme switches to Eryx, to represent the great local goddess, Aphrodite; from the mid-fifth century BC the iconography spreads in the Elime area in order to represent the 'nymphs' of Entella and, finally, that of Segesta (ca 410 BC). The branch held in her hand by the 'nymphs' Segesta, in addition to the 'phiale', appears later in the hands of the goddess Aphrodite represented performing a libation on a stater of Paphos in Cyprus (ca 385 BC). To these Sicilian figures are added the 'nymphs' Terina (Magna Graecia, 440-425 BC) and the Thessalian Trikka (second half of the 5th century BC), depicted with the same iconography, although in arms. In the view of a figurative and semantic continuity, at the end of the 3rd century BC, the representation of the Tyche of Thermai Himeraiai, the site founded after the destruction of Himera in 409 BC, is significant: the goddess is full

Figure 2: Map showing the Mint that have coined coins whose iconography presents a female personage, standing, with a 'phiale'. South Italy zone.

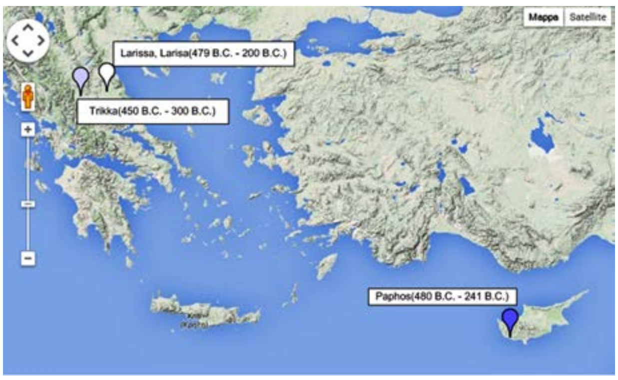


Figure 3: Map showing the Mint that have coined coins whose iconography presents a female personage, standing, with a 'phiale'. Greek zone.

length, standing, veiled, turreted and with the cornucopia of Tyche, but with the 'phiale' in her right hand, as already in the 'nymphs' type of two centuries before.

4. Design Principles of the DIANA System

New emerging web technologies are rapidly changing the way of conceiving web applications. At the same time, more and more services are available over the Internet that can be used to build new innovative applications. This paradigm is known with the term 'Cloud Computing'. Cloud Computing provides three different services, i.e., Infrastructure as a Service (IaaS), Platform as a Service (PaaS) and Software as a Service (SaaS). Services over the Cloud Computing can be build integrating other services available over the Internet whose data can be exchanged through the Dataweb[10]. According to such a philosophy, sophisticated web applications can be defined as SaaS. DIANA is a SaaS built using the Google Maps PaaS [3].
As previously discussed, our approach consists in studying the diachrony starting from the concept of ancient coin iconography. More specifically, from a particular iconography, we find the coins in which it is drawn, the mints where the coins were minted, and finally their geographical locations.

The DIANA database has been designed according to a hierarchical data model. The first layer of the hierarchy is the 'Authority'. It can be a king, a city, a State, etc. Each authority is responsible for one or more 'Mint(s)', i.e., the entity responsible for issuing coins. Such an entity also stores the data related to the geographical location in terms of latitude and longitude. Each Mint is related to one or more ancient 'Coin(s)'. In turn, each 'Coin' can be bound to one or more 'Bibliography' entities in which the 'Coin' is cited. In addition, each 'Coin' entity can be bound to [0,2] iconography entities. We defined four types of iconography, i.e., 'Personage', 'Animal/Monster', 'Flora', and 'Object'.

The system includes four main components: the user, the DIANA web application, the DIANA database, and the Google Maps PaaS. Figure 4 summarizes the components involved in the DIANA system. The user is the subject who wants to access the DIANA services through his/her web browser via the Internet querying the DIANA web application. The latter is responsible for processing the user' requests and for providing output data related to authorities, mints, and coins in form of digital geographical maps. In order to do so, the web application interacts with both the DIANA's database and Google Maps PaaS.

In particular, when the web application receives a request from a user, by means of server-side mechanisms developed in PHP language, it queries the DIANA's database and it retrieves the needed data. After that it parses the data formatting them in an XML document describing the retrieved data. In the end, using the Asynchronous JavaScript and Xml (AJAX) programming technique, the XML document is sent to the client-side JavaScript code on users' web browser. The latter processes the XML document, it extracts data, and it builds a request form map drawing to the Google Maps PaaS through the Internet. In the end, the output map, displaying mints and ancient coins is sent to the user's web browser.

In the following we report the main steps required to generate a digital map. In step 1, the user requests the PHP page for querying the system. So, he/she types the DIANA web address in address bar of his/her web browser and push the 'go' button. In step 2, the DIANA web application sends to the user's web browser the web content (including both HTML and JavaScript codes) dynamically build processing the PHP code. We remark that the JavaScript code implements the client-side logic of DIANA. In step 3, the JavaScript client-side code is processed by the browser and sends a request for maps service to the Google Maps PaaS. In step 4, the google maps services and software libraries are loaded on the web browser. In step 5, the user submits a web form and an asynchronous request is sent querying the system according to a given iconography. In step 6, the system processes the request and queries the database. In step 7 the system builds an XML document according to the retrieved data. Figure [XML] show how data are formatted using an XML document. According to the XML scheme, each ancient mint corresponds to a marker in the digital map, and each mint includes several coins, each one identified by a Unique Identifier (UID), used to retrieve additional information when required.

In step 8, the DIANA server-side system sends the XML document to the DIANA client-side code on the user's web browser. In step 9, the Javascript interpreter of the web browser, reads the received XML document and build a map request according to the received data and it sends a map request to Google Maps PaaS. Finally, in step 10, the DIANA client-side code displays the map on the user web browser. The mints are represented by markers displayed from the oldest one to the latest one according to

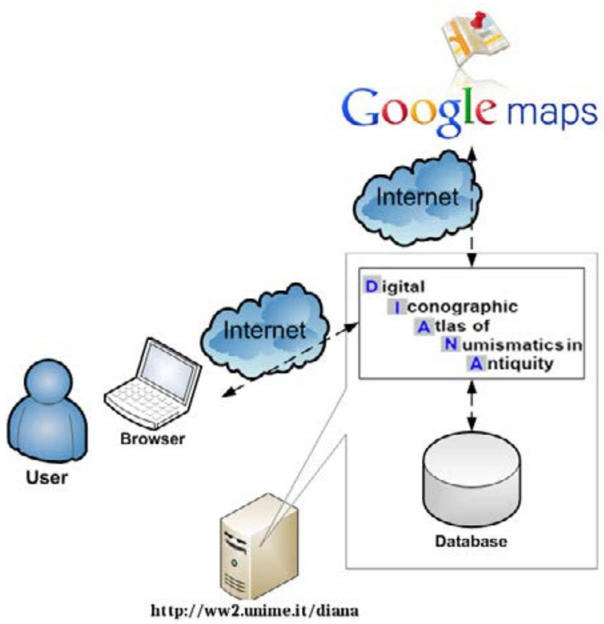


FIGURE 4: COMPONENT INVOLVED IN THE DIANA SYSTEM.

```
<markers>
  <marker mintId="518bc41d4a455" name="Zankle Messana" fromYear="525 B.C." toYear="200 B.C." authName="Zankle-Messana" authFromYear="525 B.C." authToYear="200 B.C." lat="38.1938137" lng="15.5540152">
    <coins>
      <coin coinId="518ca0e34f729" denomination="Tetradrachm" metal="silver" fromYear="420 B.C." toYear="413 B.C." approx="1" obverseImage="518c9dda4f739.jpg" reverseImage="518c9dea9ebca.jpg" uid="gsalamone"></coin>
      <coin coinId="518ca739db2f0" denomination="Tetradrachm" metal="silver" fromYear="420 B.C." toYear="408 B.C." approx="1" obverseImage="518ca4bb54730.jpg" reverseImage="518ca4c57d3c3.jpg" uid="gsalamone"></coin>
      <coin coinId="518caad6e20a2" denomination="Litra" metal="bronze" fromYear="317 B.C." toYear="311 B.C." approx="1" obverseImage="518ca94dd1aae.jpg" reverseImage="518ca95793813.jpg" uid="gsalamone"></coin>
      <coin coinId="522aa55aab33e" denomination="Tetradrachm" metal="silver" fromYear="420 B.C." toYear="413 B.C." approx="1" obverseImage="522aa4c907df0.jpg" reverseImage="522aa4ce80ec0.jpg" uid="mpuglisi"></coin>
    </coins>
  </marker>
  <marker mintId="518e483adbc0f" name="Zankle-Messana" fromYear="264 B.C." toYear="190 B.C." authName="Mamertini" authFromYear="264 B.C." authToYear="190 B.C." lat="38.1938137" lng="15.5540152">
    <coins>
      <coin coinId="518edac8687cf" denomination="Litra" metal="bronze" fromYear="264 B.C." toYear="241 B.C." approx="1" obverseImage="518ed9f6160a3.jpg" reverseImage="518eda107725e.jpg" uid="bcarroccio"></coin>
      <coin coinId="5228a217204cb" denomination metal="bronze" fromYear="211 B.C." toYear="190 B.C." approx="0" obverseImage="5228a12ade611.jpg" reverseImage="5228a1383f35f.jpg" uid="gsalamone"></coin>
    </coins>
  </marker>
</markers>
```


FIGURE 5: EXAMPLE OF XML DOCUMENT INCLUDING A LIST OF MINTS AND COINS RETRIEVED ACCORDING TO PARTICULAR SEARCH CRITERIA.

their colors. Colors are assigned varying the gradient: the oldest white and the latest one is blue. The other markers are colored varying the gradient progressively from the clearest color (i.e., white) to the darkest color (i.e., blue) according to the mint chronology. Such a mechanism allows the user to easily recognize the evolution of an iconography according to both time and space. By clicking on a marker it is possible to display both data concerning the mint (authority, chronology, and so on) and the ancient coins of such a mint according to the iconography looked for in step 1.

5. Towards Linked Data to Improve the Diachrony in Numismatics

The DIANA project responds to a real need, since there is no general work providing organized documentation in the field of coin iconography and also to the well-known need in the archaeological field for a concise and immediate picture of the figurative subjects and themes adopted in official contexts, to verify their appurtenance and coherence with the cultural heritage of a country. As previously discussed, DIANA makes it possible to study the 'diachrony' with a new innovative approach starting from ancient coins. The project has been designed with an eye toward the possibility of integration with other digital archives for the support to linked data. The basic idea is to enable DIANA to analyse, besides ancient mints and coin iconographies, even other related data coming from other digital archives. However, this process of integration is not trivial at all, because the current digital archives are based on different technologies that often are not compatible. Moreover, several digital archives are often updated to the last Information Technologies, others are based on old systems. Figure 6 shows a list of the major digital archives related to ancient coins, artefacts, and places

We can observe that the only web digital archive that allows third party system to link its data is Pleiades that provides data export features in CVS, KML, RDF format. Moreover, web digital archive has offered public cloud computing services yet. This means that currently promoting a web environment where communities, organizations, universities, and museums can cooperate

Digital Archive	Web Site	Support to Linked Data	Support to Cloud Computing Services
Mantis	http://numismatics.org/	No	No
Staatliche museum online numismatic catalogue.	http://www.smb.museum/	No	No
Roman Provincial Coinage Online	http://rpc.ashmus.ox.ac.uk/	No	No
Numismatic Museum Athens	http://www.nma.gr/	No	No
Department of Archaeology and Museums, Govt. of Rajasthan	http://www.ancientcoins.rajasthan.gov.in/	No	No
Pleadies	http://pleiades.stoa.org/	Data export in CSV, KML, RDF format	No

Figure 6: Support to data export and cloud computing services of the major coin archives.

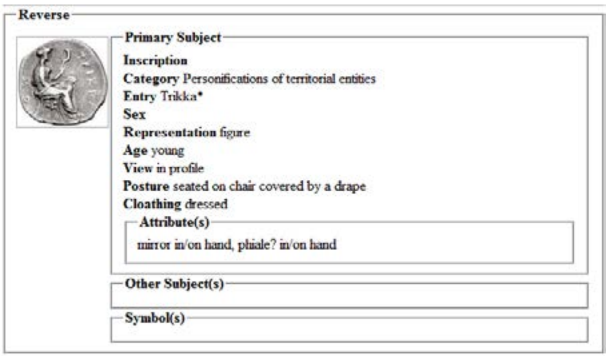


Figure 7: Iconography of the reverse side of a coin belonging to Trikka.

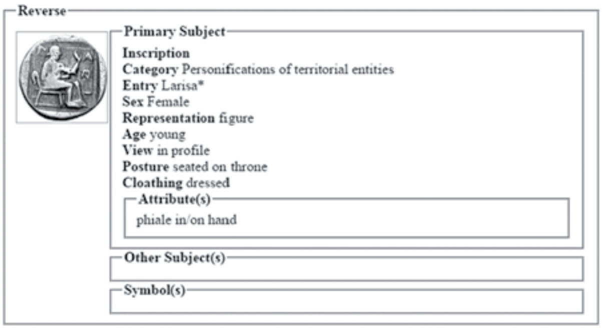


Figure 8: Iconography of the reverse side of a coin belonging to Larissa.

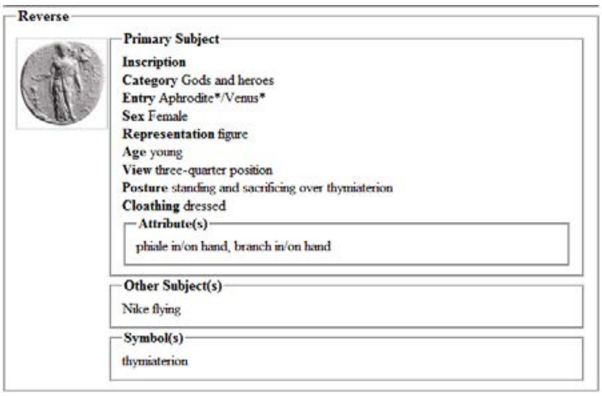


Figure 9: Iconography of the reverse side of a coin belonging to Paphos.

integrating their data is very hard. In addition, from a technological point of view, the major digital archives are quite antiquated. In fact, no web digital archive reported in Figure 6 supports any form of cloud computing PaaS and SaaS.

In order to better explain the advantages of linking heterogeneous data to improve the diachronic study in Numismatics, in the following we will discuss a possible integration between DIANA and Pleiades (i.e., the only digital archives listed in Figure 6 supporting public data export). Currently, in DIANA, places are located according to ancient mints, but in other digital archives places can be located according to ancient artefacts and names, as well as in Pleiades.

Pleiades gives scholars, students, and enthusiasts worldwide the ability to use, create, and share historical geographic information about the ancient world in digital form. At present, Pleiades has extensive coverage for the Greek and Roman world, and is expanding into Ancient Near Eastern, Byzantine, Celtic, and Early Medieval geography. Currently Pleiades stores over than 30.000 places, locations, and names.

Considering the example described in Section 3, if we try to look for the word 'phiale' in Pleiades, we can observe two results in a geographical location near to the mint 'Trikka', 'Larissa', and 'Paphos' found in DIANA (see Figure 3). Considering DIANA, the 'phiale' attribute is depicted in the reverse side of three ancient coins (see Figures 7, 8 and 9) respectively coined in 'Trikka', 'Larissa', and 'Paphos' from 470 B.C. to 380 B.C.

Instead, looking at the two occurrences found in Pleiades (see Figure 10), the term 'phiale' corresponds to a geographic name (Phiela/Phiale), dated in late antique (300 A.D. - 640 A.D.), accurate, certain, and complete (see Figure 11) and to a geographic name (Phiale Limne) related to a crater lake located in the Golan Heights, dated in the early Roman Empire (30 B.C. - 300 A.D.).

Observing the results found in DIANA and in Pleiades, from the iconography of ancient coins we can speculate that the term 'Phiale' was originated in a period from 480 B.C. to 200 B.C. Moreover, from Pleiades, we have evidence that the term was used in a period from 30 B.C. to 640 A.D.

Since Pleiades does not provide any form of real time interactive query (for example, by means of web services), the integration with DIANA presents several issues. In fact, in order to link data between the two systems, the dataset of Pleiades should be periodically downloaded, parsed, processed, and stored in the DIANA system causing an evident overhead. This problem would be easily mitigated using web service mechanisms provided by cloud computing services. In addition, in order to allows DIANA to display, besides data related to mints, also related data

Figure 10: Digital map displayed in Pleiades looking for the term 'Phiale'.

Romanized Name(s):
Phiale

Name as attested:

Language:

Name type:
geographic name

Accuracy of transcription:
accurate

Level of transcription completeness:
complete

Level of certainty in association between name and the place:
Certain

Temporal attestations:
- Late Antique (AD 300-AD 640) (confident)

References:

Initial Provenance:
Barrington Atlas: BAtlas 53 B2 Phiela/Phiale

Figure 11: *Description of the marker displayed at north of the map depicted in Figure 10.*

Romanized Name(s):
Phiale Limne

Name as attested:

Language:

Name type:
geographic name

Accuracy of transcription:
accurate

Level of transcription completeness:
complete

Level of certainty in association between name and the place:
Certain

Temporal attestations:
- Roman, early Empire (30 BC-AD 300) (confident)

Figure 12: *Description of the marker displayed at south of the map depicted in Figure 10.*

coming from Pleiades, it is require to develop a software adapter in order to build 'on-fly' XML documents such as the ones used by DIANA (see Figure 5) for ancient mints and coins. Such XML documents are indispensable to build a digital map according to the Google Maps PaaS.

4. Conclusion and Future Work

Thanks to the new information and communication technologies, DIANA allows researchers to study diacrony according to an approach to classify the iconography of ancient coins. The objective of the DIANA project is to extend its scope supporting the integration with data coming from other digital archives. In fact, the objective of DIANA is to become a very useful tool to represent the diachrony of coin types stored according to the *Lexicon Iconographicum Numismaticae* and link this data with other related information coming from other digital archives. Nevertheless, there is a long way to go for a complete integration between different digital archives. In fact, the major digital archives do not provide any data export mechanism and any interactive query system. Even though cloud computing technologies are emerging embracing different scientific fields, currently this is not true also for digital heritage. We hope we will succeed in stimulating the scientific community regarding the need to update digital archives for the support to the possibility to link their data. In future work, we plan to integrate the DIANA system with other digital coin archives starting from a dialogue with researchers responsible for other digital heritage projects.

References

MANTIS. (2014) [Online] http://numismatics/search/. [Accessed: 1st July 2014]

ANS. (2014) [Online] http://www.numismatics.org. [Accessed: 1st July 2014]

MANTIS (2014). [Online] http://numismatics.org/search/. [Accessed: 1st July 2014]

GOOGLE MAPS (2014). [Online] http://maps.google.com. [Accessed: 1st July 2014]

SMNC. (2014) Staatliche Museum Numismatic Catalogue [Online].http://www.smb.museum/ikmk/index.php?lang=en. [Accessed: 1st July 2014]

OCRE. (2014) Online Coins of the Roman Empire [Online]. http://numismatics.org/ocre/.[Accessed: 1st July 2014]

RPCP. (2014) Roman Provincial Coinage Project [Online]. http://rpc.ashmus.ox.ac.uk/. [Accessed: 1st July 2014]

NOMISMA. (2014). [Online] . [Accessed: 1st July 2014]

NUD. (2013) *Numismatic description standard* (NUDS). [Online] http://nomisma.org/nuds/numismatic_database_ standard. [Accessed: 1st July 2014]

BRACEY R. (2013) *Online numismatic databases a review*. [Online] http://www.academia.edu/2335397/Online_Numismatic_Databases_-_A_Review. [Accessed: 1st July 2014]

AMANDRY M., BATESON D. (2009) A Survey of Numismatic Research 2002-2007 in *International Association of Professional Numismatists, Special Publication 15.* Glasgow.

PLEIADES. (2014). [Online] http://pleiades.stoa.organization. [Accessed: 1st July 2014]

BERNERS-LEE T., HENDLER J., LASSILA O. (2001) *The semantic web. In Scientific American* 284(5), pp. 28–37, 2001.

KML. (2014). [Online]. Available: https://developers.google.com/kml/. [Accessed: 1st July 2014]

DIANA. (2014). *The Digital Iconographic Atlas of Numismatics in Antiquity*. [Online] http://ww2.unime.it/diana. [Accessed: 1st July 2014]

SALAMONE, G. (2013). *'Una' e 'molteplice': la Ninfa eponima di città. Iconografie monetali e semantica*, Reggio Calabria 20132 (Semata e Signa 6).

SALAMONE, G., CALTABIANO, M. (2007). *Il significato delle immagini. Codice e immaginario della moneta antica.* Reggio Calabria, Italy: Falzea, pp. 129–174.

SIMON, E. (1998). *Archäeologisches zu Spende und Gebet in Griechenland und Rom*. In: F. Graf, (ed.). *Ansichten griechischer Rituale. Geburtstags-Symposion für Walter Burkert*. Stuttgart – Leipzig.

Celtic Coins in Context, a New Database

Katherine GRUEL,[1] Agnes TRICOCHE,[2] Philippe CHARNOTET[3]

[1] Archéologie et Philologie d'Orient et d'Occident (AOROC) CNRS : UMR8546, École normale supérieure [ENS], Paris, France
[2] Pays germaniques , histoire - culture - philosophie (PG,HCP) - CNRS : UMR8547, Université Paris VIII - Vincennes Saint-Denis, École normale supérieure [ENS] – Paris, France[3] Archéologie et Philologie d'Orient et d'Occident (AOROC), Paris, France
katherine.gruel@ens.fr, agnes.tricoche@ens.fr, Philippe.charnotet@free.fr

Abstract
The object of this data base is to show the links between the coins and the settlement. The entry of data can be done by the individual coins or by the archaeological context. In that case, you can introduce only the number of coins in each typological serie. The central object of this base is the Celtic period, but all the coins from Roman and Greek worlds can also be registered. Coins are classified by groups, series and classes.

Keywords: Celtic, coins, archaeological context, Collaborative Database, GIS, Iron Age

Introduction

We would like to present a collective database on coinages in Celtic context: 'Base faciès monétaires antiques'. Actually, it contains nearly ten thousands coins and three hundred settlements. All the students, which have worked on Celtic coinages for ten years in France, have been federated around this project. The object of this database is to study the links between the coins and the settlement. It permits also to generate easily coins distribution maps. Data can be entered through the individual coins or through the archaeological context. In that case, you can introduce only the number of coins of each typological serie. The central object of this database is the Celtic period, but all the coins from Roman and Greek worlds founded in the same site can also be registered. The archaeological context is essential to precise the datation of the coins circulation and to define the first apparition of the type in the circulation. The connexion between these two informations, given by the digging, is essential to valid the chronology of the issues and the use the coins. It permits to graduate the monetarisation of the antic economy at different periods.

1. The data structuring

Coins are classified into groups, series and classes. The 'Facies monétaires' database is divided in 6 Tables, including 3 Levels of Cataloguing with 2 layers each:

Archéolocalis : The Geographical index:

1. Sites
2. Contexts

The Coins inventory :

1. Individual.
2. Facies (by typological groups)

The Typological index:

1. Series
2. Classes

Then there are four additional Tables:

– Séries_facies,
– Recorded Coins Photography,
– Type drawing,
– Coordinated Participants Addresses.

Several models and count tables permit statistics and exports into excel tables to perform mapping and create numismatic catalogue.

1.1. Archéolocalis : The Geographical index

Geolocation is based on town centroid in WGS84 geographic coordinate system or in IGN Lambert II étendu, with additional data about four levels of administrative units, from town up to country, based on European NUTS classification (Eurostat, 2012).

The context can be precise until the stratigraphical unit. This table is use for all the archaeological database of our laboratory. That permits to connect all the specialis's database to the general database on the Iron Age: BaseFer.

1.2. Typological description

The type catalogue is potentially one of the most difficult problems for the Celtic coinages and we have made big progress by using this same database. We have distinguished monetary Series and division of each Serie in classes (Gruel, 1989 p.8). Eventually, we can indicate a die number. We don't use the term of 'type' because of its ambiguity. In this table, it is standard coin types which are described in more detail. It should be pointed out that this is not a description of an actual individual coin, nor

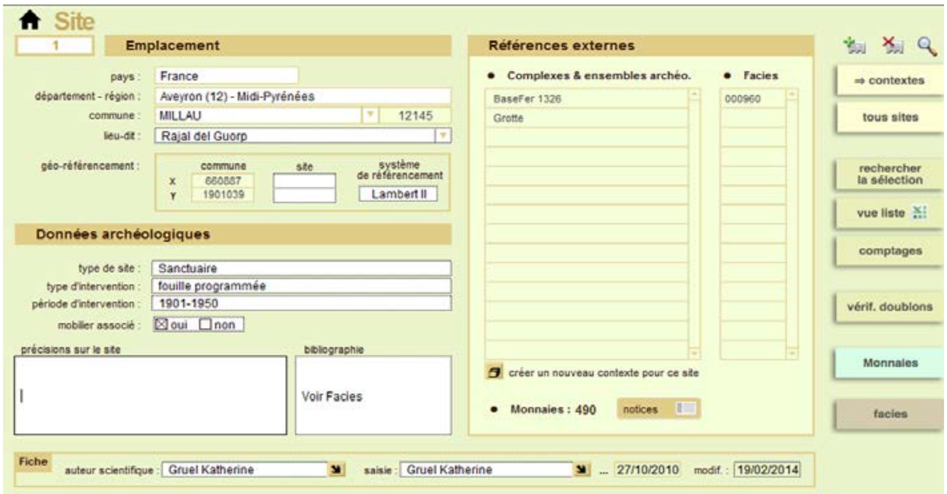


Figure 1: The data entry form to create a new site.

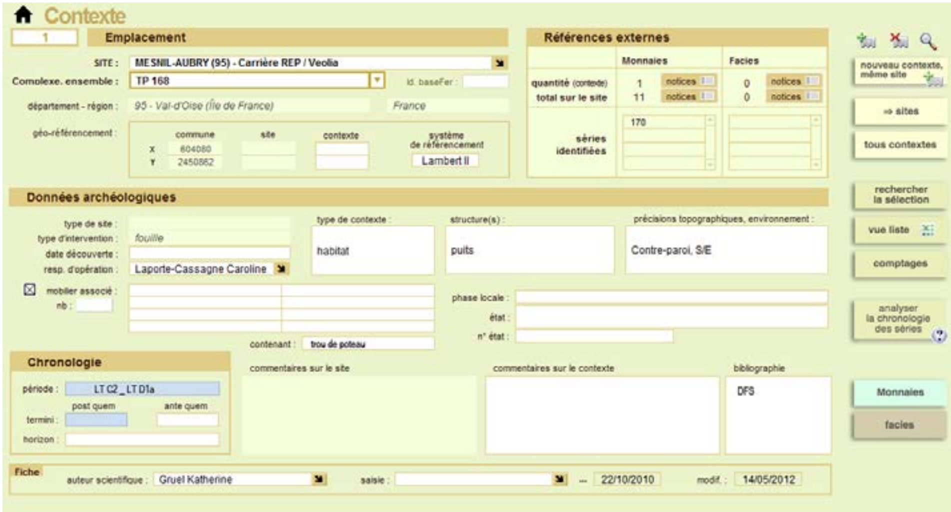


Figure 2: The data entry form to create a new context in the site.

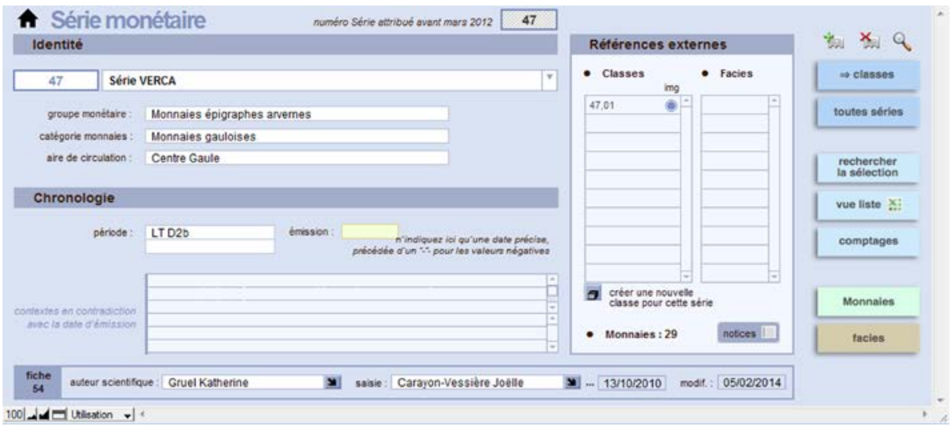


Figure 3: The data entry form to create a new typological serie.

of what you can actually see on it, but of an ideal type (i.e. what you would see if you had a perfectly struck and perfectly preserved coin). It send back to several coins referenced in catalogue available from the great museums or other public collections. That means that coins used in reference can always be consult by anybody.

For the roman coins present in the same context, the coins are arranged according to the RRC1 and RIC2. We have integrated the number of the RRC and of the RIC in the identification number. 20000 contains all the provincial roman coinages, 30000 + RRC number, the roman republican series, 40000 and following + RIC number , the roman imperial coinages. By example, 30413,01 is equal to RRC 413,1 from L. CASSI LONGINVS.

1.3. Facies Registration only

For a lot of archaeological sites and for the big hoards, we have the possibility to integrate only a coins list by series without individual description. A special table compiles the bibliographic data. It is quite useful, to compare several settlements and to map the series distribution.

1.4. Record by coins

The 'Monnaie' database records information on the individual coins. It imports details of the site from 'Archeolocalis' and 'Site'. To facilitate data entry, it uses an additional base to enter the basic information on standard coin types, using the type code.

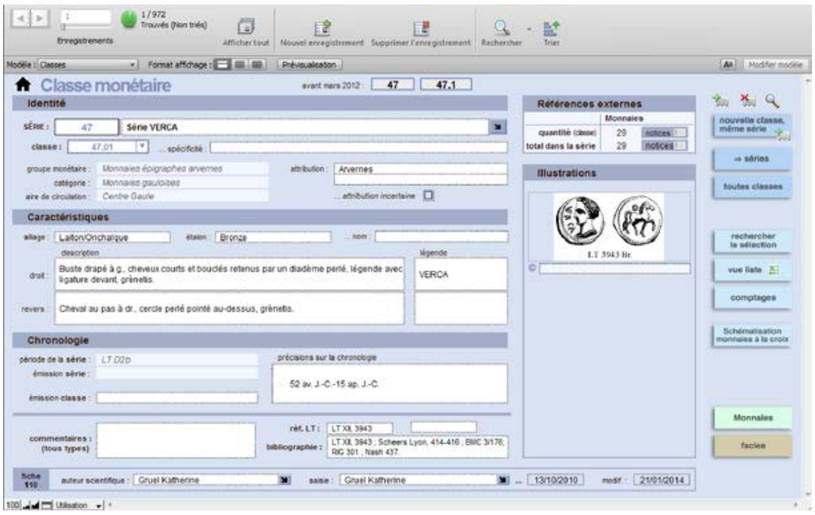


Figure 4: The Data Entry form to create a new typological class.

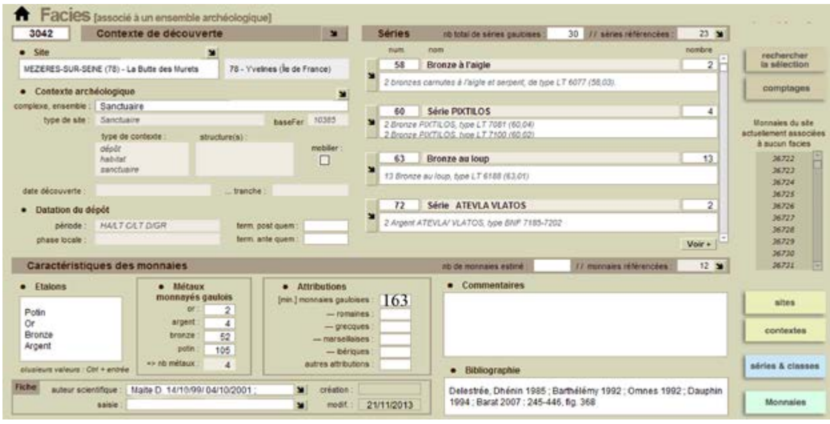


Figure 5: The Data Entry form to create the monetary facies of the site.

Figure 6: The Data Entry form to record each coin.

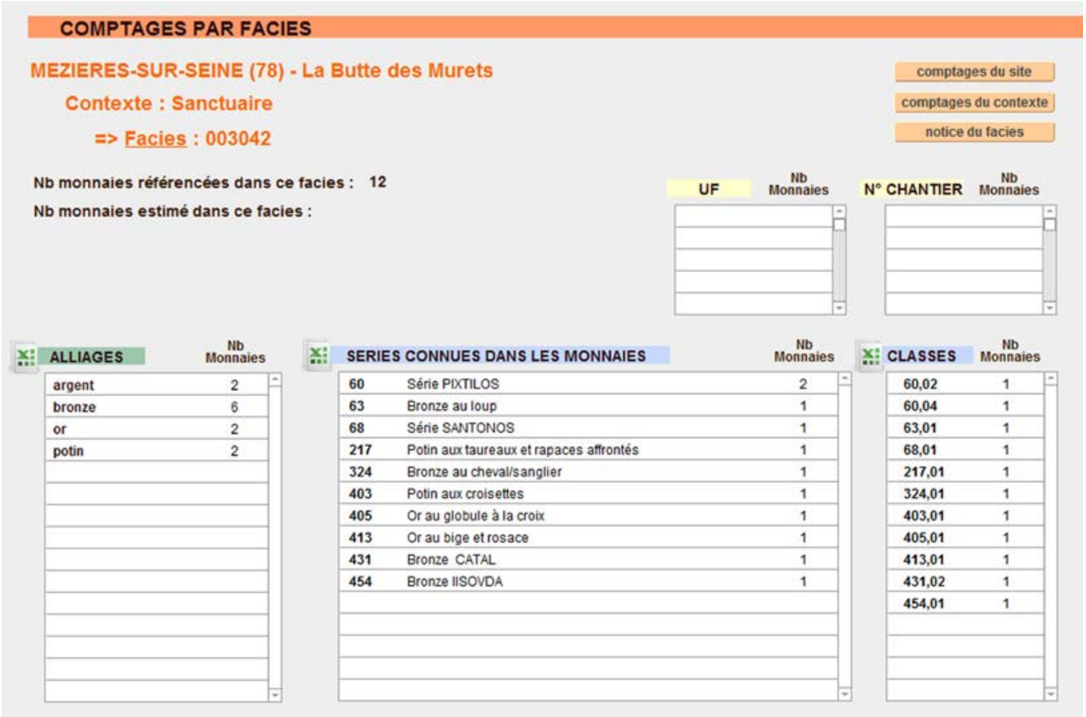


FIGURE 7: COUNT BY FACIES.

2. Exports and statistics

We can use several models to compare the sites, the series, the alloys, the contexts and the dating obtained by the context and by the coinage chronology.

- Models counts are accessible from the main models (buttons), they display counts automatically generated and used to move from model to model as scale required counting.
- Export Models Excel are direct extensions of models of counts, they take all tables counting individual way, to export selected information to Excel.
- Models are dedicated to the Web : graphical interface dedicated to the online with restricted access permissions (view only, limiting available models, activation of import, export, printing, etc.), scripts.
- Links are foreseen with a GIS for mapping and spatial analyses. It could be use for the alimentation of the Iron Age Atlas. Standardised search and print facilities are being worked on. But anybody who knowns FileMaker Pro can of course already print and search using the standard programme menus.

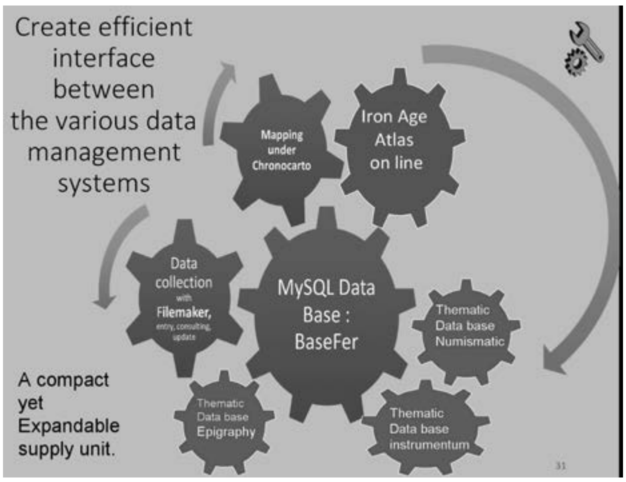


FIGURE 8: AOROC DATA MANAGEMENT SYSTEM.

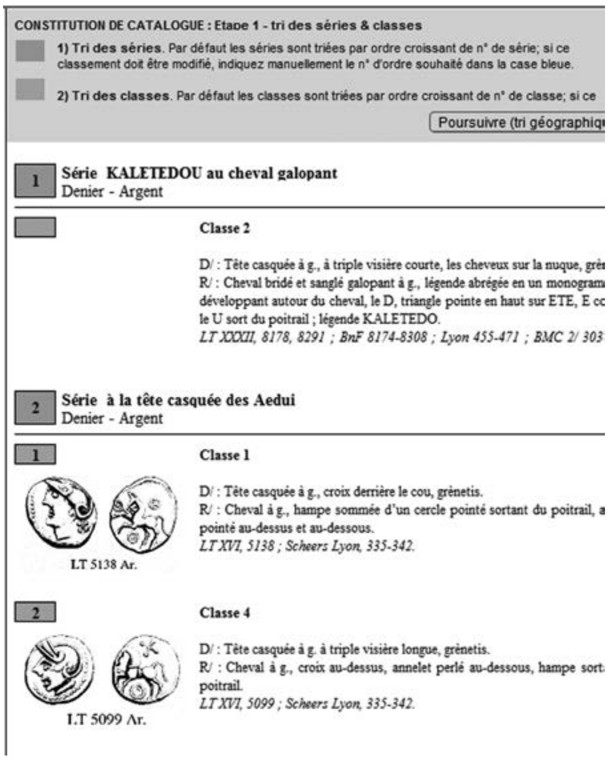


FIGURE 9: EXAMPLE OF THE GREAT FACILITY TO MANAGE THE STORE REFERENCES WITHIN DIFFERENTS CATALOGS.

- Models for tables give access to the main headings in tables form of individual records - Models in list include the major headings of each table presented in lists
- Models are dedicated to images from the records coins and typological classes, with access to zoom in metadata and scripts to import new images.

To summarise, with this database, all the data can be automatically count and use for statistical analyses. It is possible to determine the coinage function in different contexts, to show the existence of specific coins for special purpose, the evolution of their use. Around the qualitative Iron Age database, the numismatic database refers to BaseFer3 for the generalities of the sites. It supplies the atlas on line4 on request to establish coins distribution maps. For the Iron Age period, during a special meeting with Colin Haselgrove and David Wigg-Wolf, in Bibracte, 20 years ago, we have established some general European topics we need to register Celtic coins in archaeological contexts, Its finality would be to pass easily from one database to another. We hope to put this database on line in 2016 and connect it with The Oxford Celtic Coin Index, l'Inventaire des trouvailles monétaires suisses and NUMIDAT-WEB.

References

BATARDY, C., BUCHSENSCHUTZ, O., GRUEL, K. & LEVERY M., (2008). ' Analyse spatiale et statistique de l'âge du fer en france. l'exemple de la ' basefer ' ', *Revue Archéologique du Centre de la France* 47. available from : url : http://racf.revues.org/1159 [on line 16 july 2014]

CRAWFORD, M. (1974). *Roman Republican Coinage.* Cambridge: Cambridge University Press, RRC

GRUEL, K.(1989). *La monnaie chez les Gaulois*, Paris: Errance

SUTHERLAND, C.H.V. CARSON, R.A.G. (1999). *Roman Imperial Coinage Vol. I: From 31 BC to AD 69 - Augustus to Vitellius*. First published in 1923, last reprinted in 1999, in RIC, Roman Imperial Coinage series, 10 vol. London: Spink.

EUROSTAT (2012). NUTS - Nomenclature of territorial units for statistics, European Commission. Available from: url http://epp.eurostat.ec.europa.eu/portal/page/portal/nuts_nomenclature/introduction [16 july 2014]

THE OXFORD CELTIC COIN INDEX, [on line] Available from : URL : http://www.finds.org.uk/CCI/index.php [16 july 2014]

Inventaire des trouvailles monétaires suisses [on line] Available from : url : http://www.trouvailles-monetaires.ch/ [16 july 2014]

NUMIDAT-WEB / Fundmünzen den Antique (FDA) [on line] Available from : url : http://www.adwmainz.de/index.php?id=338[16 july 2014]Blank, G. (2013) Who Creates Content? *Information, Communication & Society* 16(4), pp. 590–612.

Uncertainty Handling for Ancient Coinage

Karsten TOLLE,[1] David WIGG-WOLF[2]

[1] Databases and Information Systems (DBIS), Johann Wolfgang Goethe-University of Frankfurt, Germany
[2] Römisch-Germanische Kommission (RGK), Frankfurt, Germany
tolle@dbis.cs.uni-frankfurt.de, David.Wigg-Wolf@dainst.de

Abstract
For archaeologists uncertainty is part of daily business. When data are recorded it is important to concentrate on facts, yet in many cases the recording of uncertain data are mandatory in order to answer research questions. As long as we stay within one system and the number of users is limited, the handling of such uncertain data is up to the individual users.
But when the data are to be exchanged, the question arises as to how to model uncertainty without risking that others fail to notice this uncertainty and build hypothesis based on uncertain data without knowing it.
In this paper we show how we store our uncertain data internally and discuss different approaches to modelling uncertainty based on ontologies with different drawbacks and benefits.

Keywords: Uncertainty, Ontology, Modelling

1. Introduction

Coinage is a rich source for the study of the ancient world, and the study of Roman Imperial coinage in particular is now well established in the print medium. The basic structure of Imperial coinage has been the focus of a type corpus known as *Roman Imperial Coinage* (RIC) providing a basic description of each of the 40,000+ recorded varieties of the coinage. This allows the date of production of coins which are found, for example in archaeological excavations, to be dated quite exactly. If these coins are found together with other artefacts, then this contextual information helps to infer the dating of the latter.

However, due to corrosion and wear coins, in particular coin finds may be badly preserved and even experts can often hardly identify the exact type of a given coin. In our database *Antike Fundmünzen in Europa* (AFE), where finds of ancient coins from within Germany are recorded, we want to preserve as much information as possible. This also includes alternatives and combinations of possible coin types, or marking attributes of the coin as uncertain if the exact value can not be assured. Together with other contextual information and reasoning techniques the existing possibilities can then be narrowed down.

As long as data remain within the original system, we are free and flexible as to what and how we model uncertainty or imperfect data. The data in AFE per se are based on a relational model. However, we plan to link AFE with other databases in Europe, and have already generated some promising results within the framework of the *European Coin Find Network* (ECFN). Utilising the concept of linked open data, the goal would be to use existing ontologies to allow us also to represent uncertain or incomplete data from AFE and so link them to other resources.

The first part of this paper describes in more detailed our present situation, including how we deal with uncertain and incomplete data within the different process steps during entering and storing them in a relational database. We will then concentrate on the exchange of data and explain existing approaches to modelling uncertainty and incomplete data, especially in the archaeological domain. Finally, we propose our solution and how we plan to implement it, and discuss the pros and cons of the different approaches.

2. Our situation

Uncertainty, imperfect information and imprecision have different reasons and are not always easy to distinguish. A nice overview about these areas and their relation to knowledge bases was provided by (Parsons 1996; Parsons 1998) as early as the end of the 90s. Here, rather than entering a deep theoretical discussion, we try to explain our situation and process of our system, AFE, based on examples.

At present we are using a Web form in order to enter information on finds of coins into a relational database. The form is used by different experts in the domain of coins (numismatists). The coins need to be described with certain attributes that are specific for coins. Some of these attributes can be measured, such as weight or size. Others cannot be measured such as: *Who is shown on the coin*?

In our database we deal with ancient coin finds from within Germany. However, for many archaeological analyses a broader regional scope is required so we are therefore collaborating with other parties in the ECFN (ECFN, 2014) framework, and also with Nomisma.org, in order to exchange data across countries. The goal of ECFN and Nomisma.org is to harmonise code lists and thesauri, and to build ontologies and best practices for describing coins and coin finds for exchanging information and benefit from linked open data concepts.

As illustrated in the figure 1 this means that the real world needs to be described by numismatists based on the model provided by the Web front end. In our case this description is then stored in a relational database, but the data need to be transformed into an ontology in order to be exchanged with other parties.

The single parts of the process:

1. **Real World**: Due to wear and corrosion coins may be poorly preserved and therefore clearly identifying the coin type can be impossible. Archaeological context may also not be clear, and although additional information might be encountered in future, uncertainty will remain (epistemic uncertainty).
2. **Expert**: Each expert is different, with a different background that also changes over time. Thus two experts might each derive different interpretations out of given facts (truth vs belief). The expert also needs to understand the Front-End Form for entering the data and this understanding should be equal (or at least similar) for all experts using the system.
3. **Front-End Form**: A form for entering data is built on top of a model that lies underneath. The expressiveness is therefore limited to the underlying model (or at least the additional information will be cut off during mapping to the underlying model – structural uncertainty or model inadequacy). On the other hand the form itself may include additional restrictions, so that even if the underlying model can represent more complex situations than the form, they cannot be entered. Realistically, the form and entry process must be feasible within reasonable time and the information entered must be clear enough in order to be mapped into the model. This means there will always be a compromise between how detailed the data are and the level of certainty that can be entered and the time needed to do so. The focus here must be on avoiding errors during entry by means of a clear structure including a clear definition of the attributes and possible values. Supporting the user (expert) during data entry is clearly beneficial.
4. **Backend-Model**: A model based on natural language (unstructured data) could be used, letting the experts express whatever they wish this way. However, current existing tools for natural language processing still have their limits. The backend-models currently used to store data persistently use structured or at least semi-structured data. This allows fast and meaningful analysis on top. However, each model has its limits; it cuts off parts of the real world. Anything that the model is not capable of representing will not be.
5. **Mapping for exchange**: For the exchange of data in most cases the target model differs from the backend-model. These differences can be classified by syntax, expressiveness and semantics. Syntax itself is not problematic, while different expressiveness will result in the lowest common denominator for the mapping

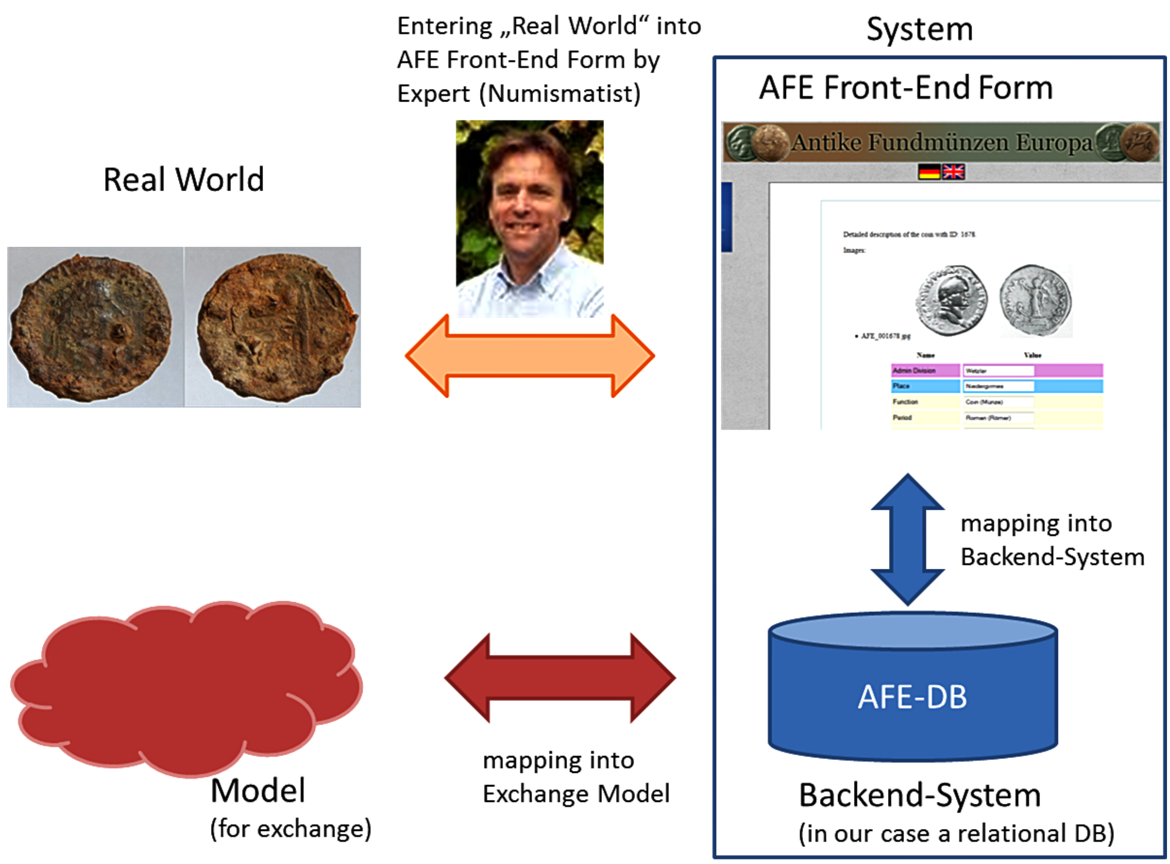


FIGURE 1: THE PROCESS FROM ENTERING COINS TO PUBLISHING DATA IN THE WEB.

Case Name	Description
Certain	The person portrayed on the coin is clearly identifiable.
One Uncertain	The person portrayed on the coin is NOT clearly identifiable. The expert has one candidate he favours.
Alternative	The portrayed person on the coin is NOT clearly identifiable. The expert can name more than one candidate for the person portrayed (and nobody else).
Alternative Uncertain	The portrayed person on the coin is NOT clearly identifiable. The expert can name more than one candidate for the portrayed person, but it could also be somebody else.
Uncertain	The portrayed person on the coin is NOT clearly identifiable and the expert is not capable of reducing the possibilities in a definite way.

Table 1: Uncertainty cases supported by AFE.

result. However, differences in the semantics are the most challenging. They need to be clear to the users and considered during mapping.

As with any chain, the weakest link determines its maximum load.

2.1. Portrait on a coin

Let us get more specific by concentrating on an attribute that cannot be measured. Assume we have a coin that shows a portrait of a man. The main question would be: is the numismatist able to clearly recognize who is shown on the coin? If the answer is 'yes', this is positive and reduces the number of possible coin types dramatically. If the answer is 'no', it gets more difficult.

By interviewing some experts for setting up our system, we encountered the following cases that happened while identifying a portrait (Table 1).

These general cases of course could be combined and extended with values for the likelihood set by the expert, for example: *I am 80% sure that the portrayed person is Titus*, or *the likelihood is 60% Titus and 40% Nero*. This approach can be helpful and there are various tools from the area of fuzzy logic or probabilistic models that could make use of such numbers. However, if you have various experts entering those numbers, you would need adjust the figures accordingly since experts might rate coins differently and therefore the numbers will not be 100% comparable. Reducing it to a limited number of cases reduces this problem (though still it remains). However, the main drawback of entering numbers is the increasing complexity this brings for entering data. It is important to note that for every value that cannot be measured automatically such cases occur. Asking users to qualify likelihood for each such field would annoy them. In fact in AFE we currently only do this for the issuer (mostly the person portrayed). For all other attributes we can only:

• mark it as uncertain by not entering a value and just checking the uncertain box,
• enter a value as certain by not checking the uncertain box
• or entering it as uncertain by entering the value and checking the uncertain box.

In fact there is another case (that we currently do not use in AFE) which is 'Derived'. There might be other criteria the numismatist can use for identification and reducing possible coin types. These criteria can range from other attributes of the coin to context information from the findspot that mean it might be possible to determine who is shown on the coin. However, to some extent this is done anyway by the expert as he has his own process based on his knowledge and experience. However, it would be very difficult for the expert to reengineer his own process (which runs automatically) in order to enter which information he derived.

We are in fact currently working on the opposite. We try to model the identification process of a coin (based on

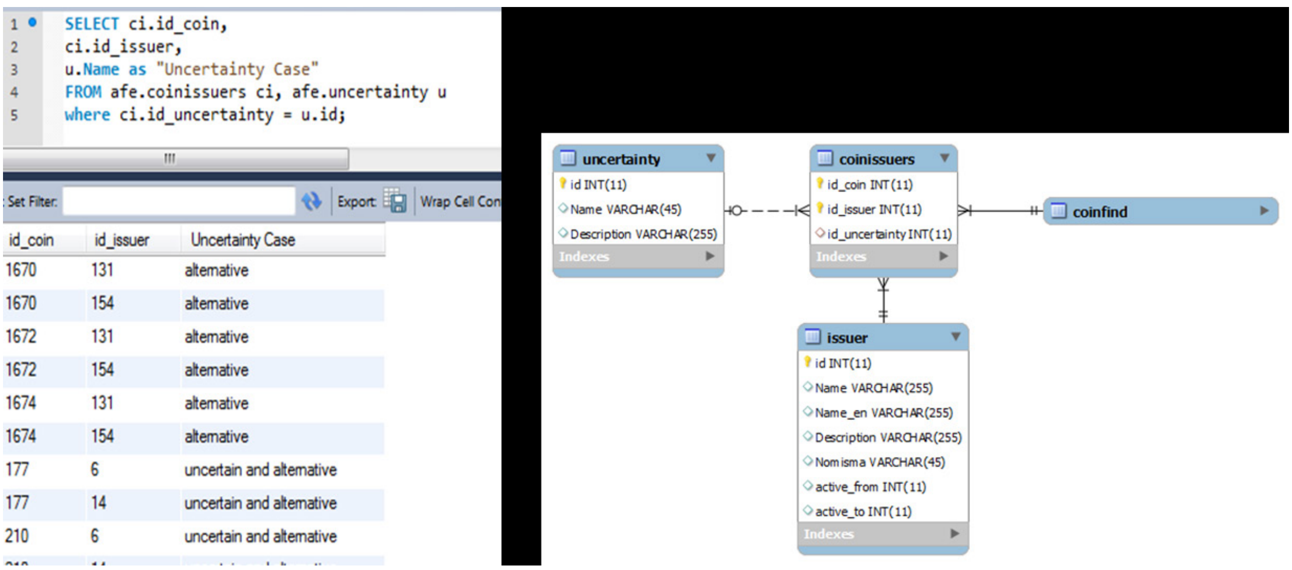


Figure 2: Internal relational representation of different uncertainty cases in AFE.

the experience of experts) inside the Front-End, and use existing available information sources (such as Nomisma.org or Online Coins of the Roman Empire (OCRE 2014)) to run this derivation of information during entry into the form.

2.2. Representation in the Relational World

How should we model our cases in the backend system? For AFE (and probably for most comparable systems) there is a relational database used for persistent storage. In the relational world only perfect information are assumed. There are tables with attributes and values entered. The relational model itself does not foresee multiple values with different likelihoods as an entry for one attribute in one row. There are approaches to extending relational databases to handle this (Beaubouef & Petry, 2009; Beaubouef & Petry, 1993), and the benefit would be that the database itself could understand and handle the uncertainty both during data entry and while handling requests. However, support and acceptance of this is still extremely limited.

For this reason, rather than extending the relational database, we included the uncertainty cases into our relational database model. This is in fact quite easy: we defined a table with all supported uncertainty cases as a reference table. For the coin-issuer-relation we defined a table which can hold multiple entries for the same coin referring to different issuers (representing the alternatives). Each row in this table has a value representing the case. Figure 2 shows the table structure we use on the right side and how it can be joined with an SQL-statement on the left. Further details can be found in (Herborn, 2013).

This is a simple and straight forward solution (and much better than adding different symbols like '?' at the end of data entries). Of course there are alternative ways of modelling uncertainty. The problem of modelling it in the model rather than it being supported by the system is more on the side of maintenance. You need to hard code how the data are entered, and manual changes tend to break the system. Just think of multiple rows for one coin with different uncertainty cases defined (which should not happen). You need to know the model when performing answers to requests; if you are asked for coins issued by Titus, do you include those where Titus is marked as uncertain or where he is one of the alternatives?

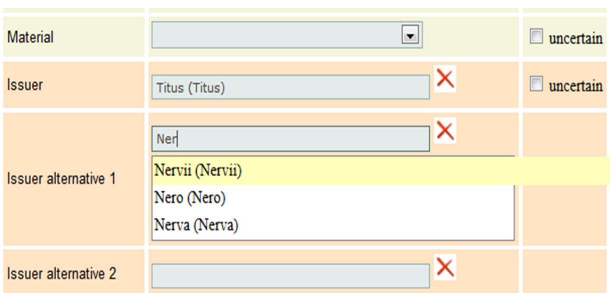


Figure 3: Front-End of AFE for entry of different uncertainty cases.

2.3. Representation at the Front-End

On the front-end there is a form that needs to be filled. In our case it is a table-based HTML form as shown in figure 3. To enter the issuer (or list of issuers for alternatives), the user starts typing and via an AJAX request values are proposed fitting what he types (based on a code list).

We decided to limit the number of alternative entries in the front end to three, as to date we have had no cases requiring more. Nevertheless, the backend model could handle more. As you can see in the figure above the uncertainty flag can only be selected once for all issuer entries (in order to represent the cases: 'One Uncertain' with just one issuer entered, or 'Alternative Uncertain' if more issuers are entered) and this flag can also be checked without entering any issuer at all (representing the case 'Uncertain'). Where it is not checked, we could have one entry ('Certain') or multiple entries ('Alternative').

Since the backend does not distinguish on the order of entries, this has no effect. However, rather than entering likelihood numbers, rating the first entry higher than subsequent ones would be a possible and easy way to extend our approach (although numismatic convention is to enter issuers in chronological or geographical order).

3. Representation for exchange

The front- and back-ends are both parts of the same system. The users of this system can be trained in order to understand it. This way they get accustomed to it and know the meaning of existing forms and how the data are presented. All work on the same common basis.

When we turn to the Open Data approach, where your data are provided to anybody, the situation changes dramatically. There is no common basis and things can easily be misinterpreted. From the technical point of view there are different approaches to avoiding problems on the syntactic level, e.g. using XML. In fact XML ensures that your data form a tree. Where you know and understand the underlying XML-Schema, you can also interpret the data. Without the underlying XML-Schema, it is just a tree.

By using the Semantic Web stack of standards such as RDF and OWL, you tackle the syntactic level, at the same time going one step further than with XML. The data are described in a graph where the directed connections (called properties) reflect a meaning. The meanings of the properties are defined in so-called ontologies, in which not just the properties but also classes and class hierarchies can be defined.

But again, without knowledge of the ontologies the data simply form a graph. However, RDF and OWL include rules on how to infer statements from existing ones. Even if you do not understand everything, you might still

Namespace	Shortcut	URL
British Museum	bmo	http://collection.britishmuseum.org/id/ontology/
CIDOC-CRM	crm	http://erlangen-crm.org/current/
Nomisma ID	nm	http://nomisma.org/id/
Nomisma Ontology	nmo	http://nomisma.org/ontology#
Resource Description Framework	rdf	http://www.w3.org/1999/02/22-rdf-syntax-ns#
W3C Uncertainty Ontology	un	http://www.w3.org/2005/Incubator/urw3/XGR-urw3-20080331/Uncertainty.owl

Table 2: Used namespaces and shortcuts for the examples.

understand some parts and they remain valid. With XML and XML-Schema this is different.

Various ontologies exist that are well known and highly reused, like Dublin Core. With Simple Knowledge Organization System (SKOS) (Miles & Bechhofer, 2009) relations between different ontologies as well as inside an ontology can be defined. This way you can define properties or classes in a range of ways, for example from being related (a kind of un-sharp relation), to being exactly the same. These relations can be used in an automated manner by programs. The final goal would be to automate the mapping between different data sources and in this way to allow searching across various sources as if they were just one database. To some extent this is already possible with currently manual mapping definitions, as we showed in (Lehmann & Varughese, 2008).

Unfortunately, not all problems can be solved with these tools alone. Even if they can explain the individual elements of syntax and semantics, they still do not explain how they all come together to give the model. There is still room for misunderstandings due to different modelling approaches that define how you build your graph in order to express certain things. This paper raises the question: *How can we model uncertainty*?

4. Modelling uncertainty

We explored different models within the archaeological and cultural heritage context such as *CIDOC-CRM* and *EDM*, looking explicitly at how to model the uncertainty we defined in our relational model. In addition we explored the *Uncertainty Ontology* of the *W3C* and how it could be adopted for AFE.

The W3C report Uncertainty Reasoning for the World Wide Web (Laskey *et al.*, 2008) includes the description of an Uncertainty Ontology. The report itself concentrates on a wider problem; it deals with the problem of inferring information from different sources by autonomous agents. It provides a classification of uncertainty types and lists different existing theories that deal with these uncertainty types. The focus of this paper is different; we aim to explore how to model uncertainty. However, the Uncertainty Ontology as described can still function as our vocabulary by using the *hasUncertainty* property defined in (Laskey *et al.,* 2008).

The URLs of the namespaces used in the examples below are listed in the following table (Table 2).

In order to explain the different approaches we take a simple situation where we have two coins (A and B). One (A) is well preserved and the portrait of Titus can easily be identified, while the second (B) is badly preserved and it is uncertain if the portrait shows Titus or not. In order to express this scenario, we will use the property nmo:hasPortrait from the Nomisma Ontology Namespace and nm:titus from the Nomisma ID Namespace. The graph in figure 6 would express that the coins A and B both show the portrait of Titus (Figure 4).

One standard way to include the uncertainty we want to express in this example would be to use the reification that comes with RDF. This means using standard RDF

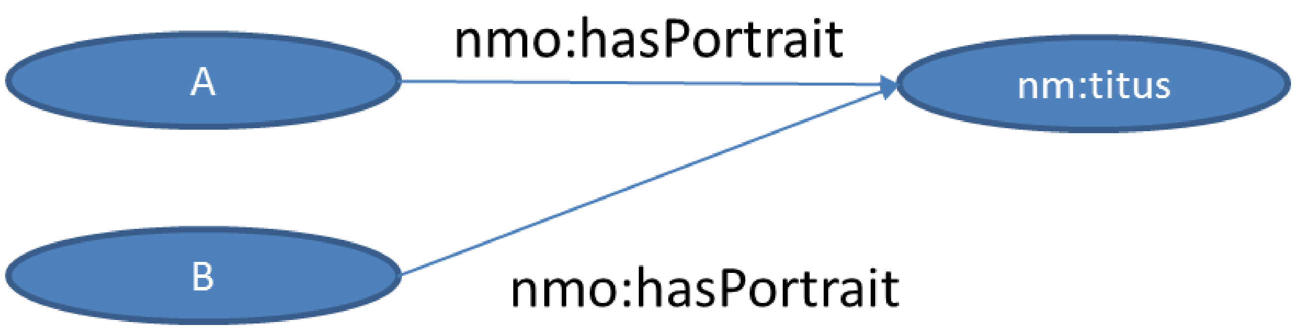


Figure 4: RDF graph representing two coins (A and B) with a portrait of Titus.

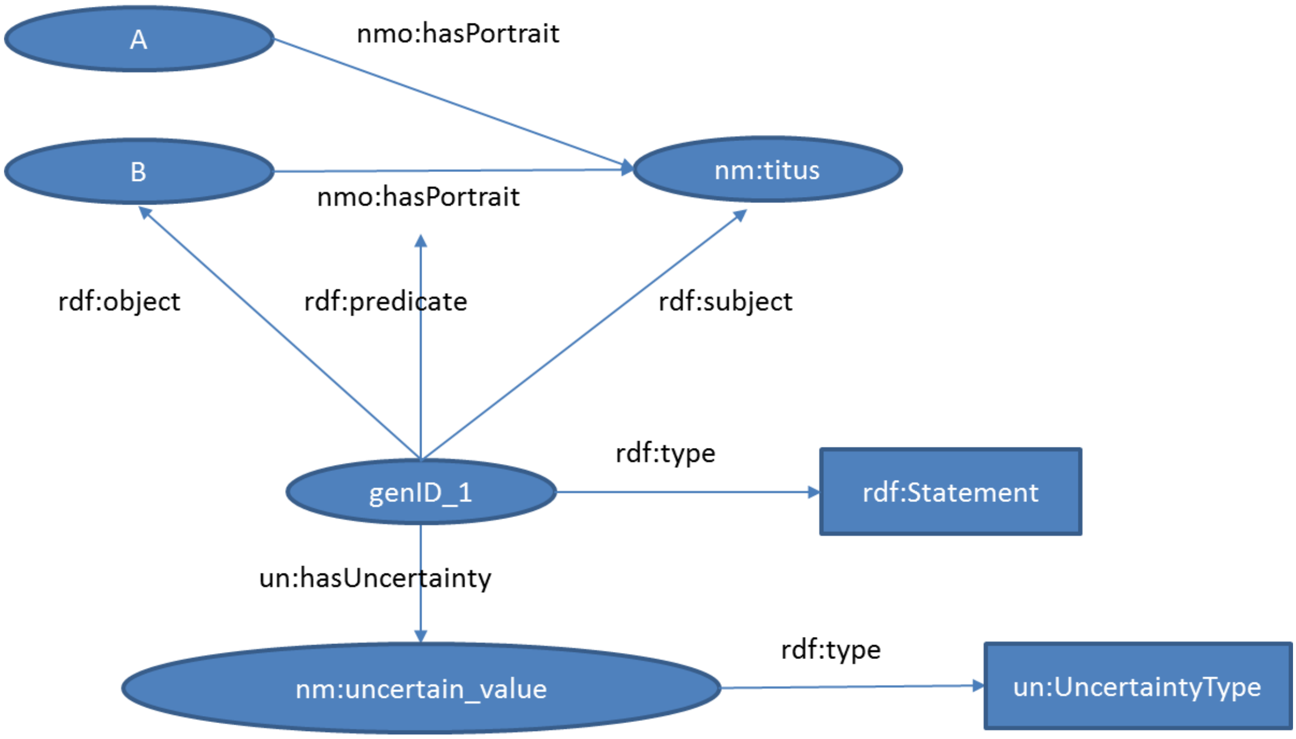


FIGURE 5: USING RDF REIFICATION IN ORDER TO EXPRESS THAT FOR COIN B IT IS UNCERTAIN IF IT HAS A PORTRAIT OF TITUS, WHILE FOR COIN A THIS IS CERTAIN.

to define the sentence '*B nmo:hasPortrait nm:titus*' as a resource, where additional statements can be attached as shown in figure 5.

Reification would be one way to model this. (Semanticweb.com 2011) discusses this reification in more detail together with some alternatives and their implication. Indeed reification is also not used very frequently due to the large number of extra triples that are needed. Whether RDF reification should be deprecated was even raised as an issue and discussed at W3C, (W3C RDF Working Group 2011).

We next studied various data models to see how uncertainty is dealt with there, for example the Europeana Data Model (EDM 2014) which does deal with provenance and source information. However, we could not find any references to uncertainty, certainty or belief.

We then looked at CIDOC-CRM and how certainty and uncertainty are handled there. But a detailed look at CIDOC-CRM produces few results. In fact under Item 9 of the CIDOC-CRM work progress (CIDOC CRM Work Progress 2012) it is stated that uncertainty and belief has been discussed as early as 2001 but the conclusion was that it was out of practical scope (which was probably true in 2001). This item also proposes another solution for reification and the usage of properties with belief classes, but would for us be no option because in our situation nearly all properties can be affected by uncertainty and therefore the number of properties would explode.

Another approach that is based on CIDOC-CRM in conjunction with the Research Space (RS) projects (Alexiev 2012; BM Association Mapping 2014; ResearchSpace 2014) shown in figure 6. It is in fact very similar to the standard reification based on RDF vocabulary.

In contrast to these two possible approaches (RDF reification and CIDOC CRM/RS), we propose to model the uncertainty by adding an additional resource directly in the property path. This could/should be a generic resource to which various comments can be added, including certainty statements. The resulting graph is shown in figure 7.

So, what are the differences? First of all, our approach needs fewer triples. This can be seen as an advantage, but the reason for proposing this approach is different. Assume you send a query to the different approaches and want to search for coins with the portrait of Titus. By using SPARQL you could write:

SELECT ?subject
WHERE { ?subject nmo:hasPortrait nm:titus }

With this query, when using reification or the CIDOC-CRM/RS approach both coins A and B would be returned, while with our approach only the certain value A would be returned. The difference is that in our approach the query needs to be more complex in order also to retrieve uncertain values, whereas in the other approaches you need to make a more complex query in order to eliminate uncertain values.

This means that our approach is safer in respect of avoiding inferring information in research that is based on uncertain data, while at the same time still allowing this to be done if wanted. But as a result of our modelling approach we have the problem that we cannot pin down the ranges of

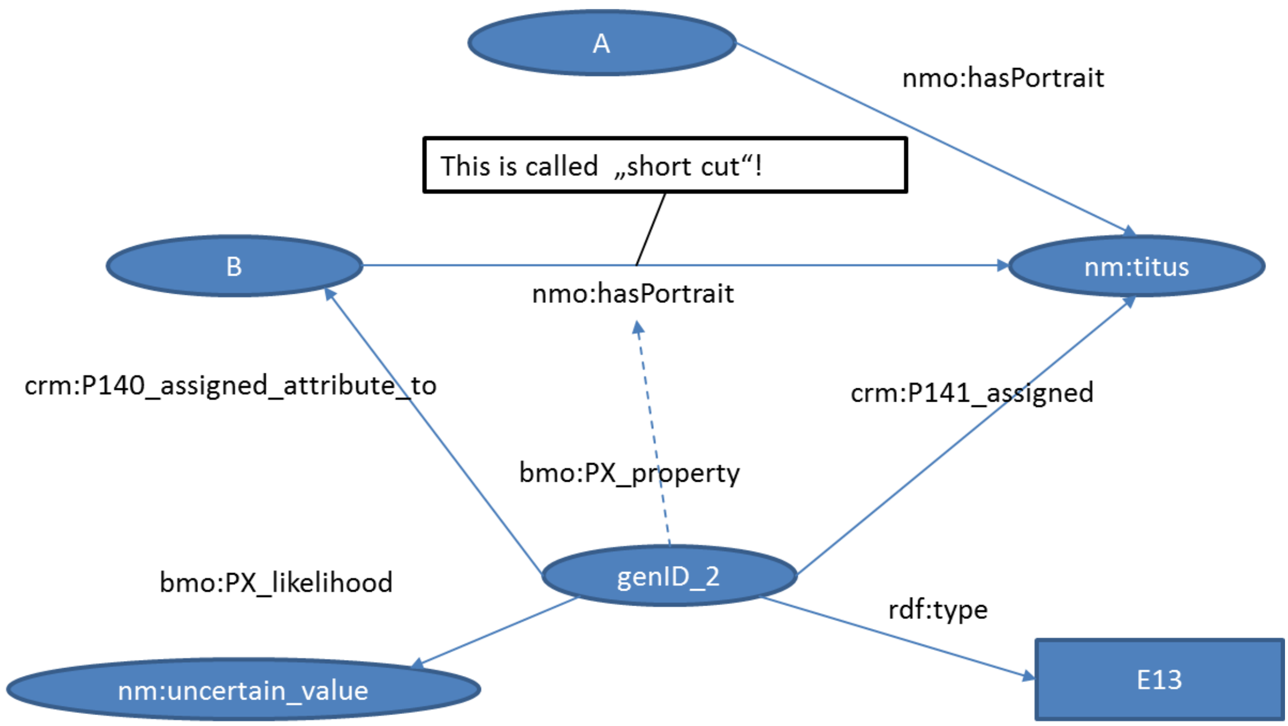


Figure 6: Using CIDOC CRM/RS in order to express that for coin B it is uncertain if it has a portrait of Titus, while for coin A this is certain.

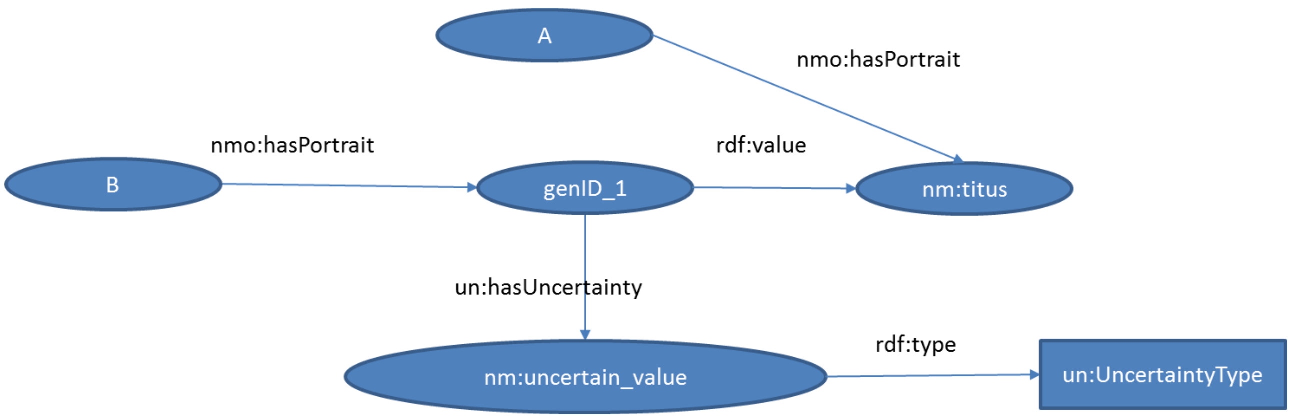


Figure 7: Our approach in order to express that for coin B it is uncertain if it has a portrait of Titus, while for coin A this is certain.

the properties we use. One possibility would be to define a person class as the range of the property *nmo:hasPortrait*, but then the *genID_1* would be inferred as being an instance of this range class, which is obviously wrong. By using reification or the CIDOC CRM/RS approach this instantiation would be no problem. However, sometimes a strict definition of domains and ranges for properties also prevents from them being reused.

5. Summary and conclusion

We have shown what kind of uncertain data we handle in our system and how we store them in a relational database. This is only one way of doing this, and how it is done internally remains the responsibility of each individual project. However, it is important to use stable and reliable modelling for uncertainty. Adding '?' at the end of entries can not be the answer for storing uncertain data, and should only be used for visualisation purposes.

For the exchange of data we showed different approaches to modelling uncertainty based on ontologies. All these approaches function; they just have different benefits and drawbacks. Our approach to modelling uncertainty provides additional security to ensure that external parties using the data do not take it as certain, and it therefore can help them avoid using flawed data when answering research questions.

However, many archaeologists are only just starting to deal with ontologies. Of course it would be best if everybody used the same ontologies and the same modelling techniques. But experience shows that this will not happen. Another conclusion from this is that in order to lower barriers for reusing ontologies, they should be built in such a way that they do not prevent different modelling approaches, e.g. by the domain and range definitions for properties being too strict.

References

ALEXIEV, V (2012). 'Types and Annotations for CIDOC CRM Properties', *Proceedings of Digital Presentation and Preservation of Cultural and Scientific Heritage*, pp. 38-51. Available from: Central and Eastern European Online Library. [20. November 2014]

BEAUBOUEF, T & PETRY, FE (1993). 'A Rough Set Model for Relational Databases' in *Rough Sets, Fuzzy Sets and Knowledge Discovery*, eds Wojciech PZ, Springer, London, pp. 100-107.

BEAUBOUEF, T, PETRY, FE (2009). 'Data mining in rough relational databases by attribute generalisation', *Int. Journal of Granular Computing, Rough Sets and Intelligent Systems* 1(1), pp.43-63.

BM Association Mapping (2014). *BM Association Mapping v2*. Available from: <https://confluence.ontotext.com/display/ResearchSpace/BM+Association+Mapping+v2#BMAssociationMappingv2-ProbablyUnlikelyProducedBy>. [15. July 2014].

CIDOC CRM Work Progress (2012). The CIDOC CRM Work Progress. Available from: <http://www.cidoc-crm.org/work_progress.php>. [15. July 2014].

ECFN (2014). *ECFN The European Coin Find Network*. Available from: <http://www.ecfn.fundmuenzen.eu/>. [17. July 2014].

EDM (2014). *Europeana Data Model (EDM) Documentation*. Available from: <http://pro.europeana.eu/edm-documentation>. [20. November 2014]

HERBORN, N. (2013). *Umgang mit und Modellierung von ungewissen Daten*, AV Akademikerverlag.

LASKEY, KJ, LASKEY, KB, COSTA, PCG, KOKAR, MM, MARTIN, T & LUKASIEWICZ, T (2008). *Uncertainty Reasoning for the World Wide Web*. Available from: <http://www.w3.org/2005/Incubator/urw3/XGR-urw3-20080331/>. [15. July 2014].

LEHMANN, I & VARUGHESE, R (2008). *Interface Meta portal for searching across different coin finds information sources*, Masterthesis, Universität Frankfurt am Main.

MILES, A & BECHHOFER, S (2009). *SKOS Simple Knowledge Organization System Reference*. Available from: <http://www.w3.org/TR/2009/REC-skos-reference-20090818/>. [17. July 2014].

OCRE 2014, *Online Coins of the Roman Empire*. Available from: <http://numismatics.org/ocre/>. [17. July 2014].

PARSONS, S. (1996). 'Current Approaches to Handling Imperfect Information in Data and Knowledge Bases', *IEEE Transactions on Knowledge and Data Engineering* 8(3), pp. 353-372.

PARSONS, S. (1998). 'Addendum to 'Current Approaches to Handling Imperfect Information in Data and Knowledge Bases''. *IEEE Transactions on Knowledge and Data Engineering* 10(5), p. 862.

ResearchSpace (2014). *ResearchSpace - a digital Wunderkammer*. Available from: <http://www.researchspace.org/>. [15. July 2014].

Semanticweb.com (2011). *Which type of reification do you use? - ANSWERS*. Available from: <http://answers.semanticweb.com/questions/2398/which-type-of-reification-do-you-use>. [15. July 2014].

W3C RDF Working Group (2011). *ISSUE-25: Should we deprecate (RDF 2004) reification of statements?* Available from: <http://www.w3.org/2011/rdf-wg/track/issues/25> [17. July 2014].

Interoperability of the ArSol (Archives du Sol) Database Based on the CIDOC-CRM Ontology

Emeline LE GOFF, Olivier MARLET, Xavier RODIER, Stéphane CURET, Philippe HUSI

UMR7324 CITERES-LAT, Université François Rabelais / CNRS, Tours, France – Consortium MASA, TGIR Huma-Num
emeline.legoff@univ-tours.fr, olivier.marlet@univ-tours.fr, xavier.rodier@univ-tours.fr,
stephane.curet@univ-tours.fr, philippe.husi@univ-tours.fr

Abstract

ARSOL (ARchives du SOL) was created in 1990 by the Laboratoire Archéologie et Territoires to exploit data from excavation sites. From the start, one of its priorities was to provide access to data. The CIDOC CRM (International Committee for Documentation, Conceptual Reference Model) was chosen as the international standard for structuring digital data for cultural heritage. This choice was imposed by the MASA (Mémoire des Archéologues et des Sites Archéologiques) consortium of the Huma-Num TGIR, whose aim is to provide unified access to digital corpora. The specific features of archaeological data present a difficulty: how can the integrity and specification of data can be respected, while generalizing the names of the fields of the data bases?

Keywords: Ontology, CIDOC-CRM, interoperability, archaeological excavations, database

1. Introduction: the ArSol database and the MASA consortium

1.1. Objectives of ArSol: data management, processing and analysis

ArSol (Archives du Sol) was created in 1990 by the *Laboratoire Archéologie et Territoires* to exploit excavation data (Galinié *et al.,*, 2005; Husi, Rodier, 2011). It has the dual aims of data management and research, and it can be used for all stratigraphic excavations. It is an open system, in other words it is modular and above all it is not constrained by predefined thesauri. It currently has two main sections:

- The *Archives de fouille* module, based on the system of stratigraphic recording used since 1970 (Galinié, 2013), concerns stratigraphic and artefact data with a view to excavation site analysis. It provides tools for management, spatialization and analysis.
- The Ceramics Database (*BaDoC*) processes ceramics data using current quantification techniques, whatever their typological facies and the scale of analysis of the archaeological contexts.

ArSol is designed as a research tool for recording and managing data during the excavation, for the manipulation and exploratory analysis of data during the post-excavation study, and finally to shift the focus from intra- to inter-site. It thus offers a complete processing chain allowing queries at different scales of analysis, essential assets for publication, with the ultimate aim of submission of evidence.

Apart from the LAT, where it is the only system in use, ArSol is also used by other research bodies and archaeology operators (INRAP, regional authorities, private operators).

1.2. Background to the interoperability of digitized resources

As in all the human sciences disciplines, the increase in digital archaeological documentation raises questions about its management, collection and use, as well as its availability to the scientific community and to a broader public. While many experiences show that the interoperability of information systems can resolve these questions without leading to a reductionist homogenization, it is clear that French archaeology is lagging behind in this area.

ArSol is being brought into line with the CIDOC CRM international standard by the Mémoire des Archéologues et des Sites Archéologiques (MASA) consortium, approved by the large research cluster (Très Grande Infrastructure de Recherche - TGRI) HumaNum (Humanités Numériques) (http://www.huma-num.fr/). The objective is to set up unified access to digitized resources via shared procedures and documentary and archiving tools. Sharing, and particularly querying, all archaeological data requires being relieved of the constraints of specific software, structure, vocabulary, accessibility and also language (Eckkrammer, Feldbacher & Eckkrammer, 2011: 154). The use of an ontology thus seems to be the best solution to make ArSol interoperable.

1.3. Implications for ArSol and committed procedures: mapping and bringing it on line

In order to achieve the desired interoperability, ArSol first had to be made available on the internet. A standardized thesaurus, PACTOLS (ISO 25964-1: 2011), was then used to free ArSol from its specific vocabulary. Finally, to allow it to become free of its software system and the structuring of its data, we have been working with the MASA consortium to study how to set up access to data using the CIDOC CRM (ISO 21127: 2006) ontology for cultural heritage. The aim is to align the ArSol data structure with that of the CIDOC CRM.

2. Understanding the CIDOC CRM ontology: between abstraction and reality

2.1. Spatio-temporal definition of archaeological entities

The interoperability of the data and the metadata used to structure the database does not necessarily mean having to ignore the needs for which they were originally produced in a given format and a specific perspective (Le Bœuf, 2013:2). This notion is particularly important in archaeology, in that all archaeological entities are recorded and become a 'research archive'. By definition, the archaeological excavation destroys the object under study, and therefore the field record is not strictly speaking an archive but nonetheless forms part of the only primary data that can subsequently be referred to.

The archaeological interest of recording excavation data concerns the material remains of past societies. From artefacts and traces of their use, the archaeologist can infer their primary or secondary use. An archaeological site is thus interpreted a *posteriori* by a function, obtained by studying the features (craftwork, private dwelling, public place, place of worship, funerary site, etc.) and by dating, usually determined by studying the artefacts.

2.2. The need to 'appropriate' the CIDOC Conceptual Reference Model

In archaeological information systems such as ArSol, particular attention is paid to the way the databases are structured: naming the tables, fields and attributed values. The precise meaning given to each field must be kept when generalizing and modelling terms and actions. The formal language of the CIDOC CRM ontology enables data to be modelled and made interoperable, but it must be 'appropriated' (Szabados *et al.*,, 2012:13).

The first step involves looking at how the data is modelled in relation to the CIDOC CRM hierarchy of classes, and then assessing the finer points of its multiple properties in order to conserve them when mapping the fields of the database. The CIDOC CRM is organized around broad categories which are not all useful here. For example, ArSol does not focus on either 'event' or 'actor', but mainly on the 'object' with the '*Faits*' (features) identified at the excavation site and from the material collected: '*Mobilier*' (artefacts). Two main classes are of interest: the '*fait archéologique*', corresponding to 'E25.Man-Made Feature', and '*mobilier archéologique*', corresponding to 'E22.Man-Made Object' (Fig. 1). They are both intrinsically linked to the activity of the archaeological excavation corresponding to 'E7. Activity'. As the excavation activity is contemporary, its metadata must be mentioned, in other words, where it takes place ('E27.Site'), the people responsible for it ('E39.Actor'), and when it was carried out ('E52.Time-Span'). With regard to the actual data, the ArSol base is

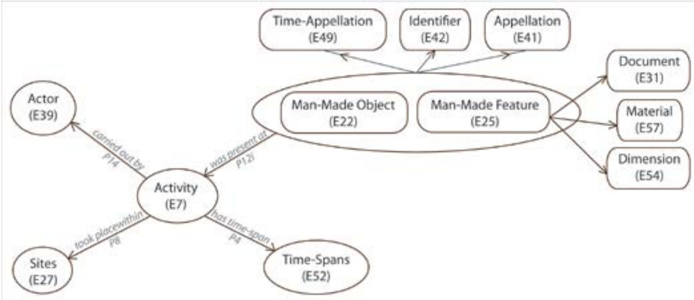


Figure 1: A synthetic view of the ArSol data model with CIDOC CRM equivalents (according to Jordal *et al.,* 2011).

composed of fields characterizing the identification of the feature or artefact ('E42.Identifier'), its description ('E41. Appellation') and its dating ('E49.Time-Appellation').

3. Steps involved in mapping the Fait and Mobilier tables in the ArSol database

3.1. Selection of basic archaeological information

ArSol has 15 main tables for excavation archives (Galinié *et al.,* 2005:5), eight tables for recording field data (*Stratigraphie, Faits, Sépultures, Murs, Structures, Séquences, Agrégations, Ensembles*), three for artefacts (*Céramique, Mobilier*) and anthropological data, and four for management of excavation archives (*Photos de fouille, Photos de mobilier, Dessins, Volumes*). As ArSol can be searched by anyone, and in view of its interoperability with other databases, we decided to initially keep only the tables containing basic information common to all archaeological excavations: features and artefacts. Mapping the Fait and Mobilier tables on the CIDOC CRM will give users access to field data from the excavated sites recorded in ArSol (including dating and documentation).

As these two tables have numerous fields, a second selection operation was performed to retain only the data characterizing the feature or artefact, the other fields having been created for data management and processing. For the *Fait* table (Fig. 2), seven fields were selected: the number of the feature, identification, description, associated structure, excavation zone, photo and drawing of the feature. For the *Mobilier* table[1] (Fig. 3), 18 fields were retained: photograph of the artefact, stratigraphic unit (US), number of the artefact, number of the associated burial place, number of the associated feature/burial place, year of excavation, category, material, description, identification, function, use, dating, length, width, height or thickness, and diameter of the artefact. The artefact is indexed by concatenating the number of the US and the number of the artefact, as several artefacts can be located in a single stratigraphic unit. Finally, two pieces of information are common to the two tables: the activity of

[1] For the mapping step, the names of the abbreviated fields have been replaced by the full name. For example, 'NumSep' = 'numéro de sépulture associée'.

the archaeological excavation and the actor/contact (LAT laboratory of Tours).

3.2. Mapping the ArSol fields on the CIDOC-CRM categories

3.2.1 Correspondence with the CRM categories: example of the Mobilier table

The following step involved mapping the selected fields on the CIDOC CRM categories (Andreussi *et al.,* 2008:2-3). The key categories are E25.Man-Made_Feature for the *Fait* table and E22.Man-Made_Object for the *Mobilier* table. Because the structure of the ontology is based on a hierarchical organization of categories and sub-categories, only the final entities are indicated (Fig. 4), opposite the corresponding fields. For purposes of readability, the fields have been combined by type of metadata, depending on whether they concern elements of contextualization, origin, inventory, identification, material description or documentation. We wished to link these combinations to the major CIDOC CRM categories: Concept (E22. Man-Made_Object // *Mobilier archéologique*), Event (E7.Activity // *Activité de fouille archéologique*), Actor (E51.Contact_Point), Place (E53.Place and E27.Site), Time (E50.Date and E49.Time_Appellation), and Thing (E31.Document, E41.Appellation, E42.Identifier, E54. Dimension, E57.Material, E55.Type).

Fields from *Fait* table	Selection
NumFait	x
Zone	x
Secteur	-
Structure	x
Etat	-
DescrFait	x
Identification	x
PhotosFait	x
DessinsFait	x

FIGURE 2: SELECTION OF BASIC ARCHAEOLOGICAL INFORMATION IN THE FAIT TABLE.

Fields from *Mobilier* table	Selection
USNum	x
NumObjet	x
NumMobilier	x
NumSep	x
FSI	x
Année	x
EtatConservation	-
Categorie	x
Matiere	x
Description	x
Identification	x
Fonction	x
Usage	x
PhotoMobilier	x
Statut	-
Datation	x
long_mm	x
larg_mm	x
HouEp_mm	x
diam_mm	x
Type	-
Poids_gr	-
Ref_biblio	-
nCatalogueObjet	-
nCatalogueMonnaie	-
RefMonnaie	-
nCaisse	-

FIGURE 3: SELECTION OF BASIC ARCHAEOLOGICAL INFORMATION IN THE MOBILIER TABLE.

3.2.2 Deployment of the ontology in RDF triples

Information from the ontology can be expressed by RDF (Resource Description Framework) triples with the fields specific to the *Mobilier* table: category, function, usage, material and dimension (Fig. 5).

In the same way as the tables of a database are inter-linked by a common identifier, so the CIDOC-CRM classes need a link between the field content (value and format of data) and their container (name). Thus, the field is first characterized (datatype) before adding the triple allowing the value of the data (value) to be introduced. For example, if we take the field '*diam_mm*' (diameter of the artefact in millimetres), its main class is 'E54.Dimension', but first the type of dimension is given, 'E55.Type = diameter', and the measurement unit, 'E58.Measurement_Unit =

Mobilier table	ArSol fields	CIDOC CRM entities	CRM categories
Contextualization	*Mobilier*	E22 Man-Made Object	Concept
	Type intervention	E7 Activity	Event
	Année	E50 Date	Time
	Contact	E51 Contact Point	Actor
Localization	*Commune*	E53 Place	Place
	Site	E27 Site	
Inventory	*IDSite*	E42 Identifier	Thing
	NumSep	E41 Appellation	
	FSI	E41 Appellation	
	USNum	E41 Appellation	
	NumObjet	E41 Appellation	
	NumMobilier	E41 Appellation	
	IDFiche	E42 Identifier	
Identification	*Catégorie*	E55 Type	Thing
	Identification	E55 Type	
	Description	E41 Appellation	
	Fonction	E55 Type	
	Usage	E55 Type	
	Datation	E49 Time Appellation	Time
Physical description	*Matière*	E57 Material	Thing
	diam_mm	E54 Dimension	
	HouEp_mm	E54 Dimension	
	larg_mm	E54 Dimension	
	long_mm	E54 Dimension	
	PhotosMobilier	E31 Document	

FIGURE 4: MAPPING THE ARSOL FIELDS OF THE MOBILIER TABLE ON THE CIDOC CRM CATEGORIES.[2]

[2] As the Mobilier table contains more fields, it allows more thorough mapping, taking into account the different CIDOC-CRM categories.

millimetre', before introducing the value, 'E60.Number = 20'. The succession of classes is clear here, and allows the conceptual classes of the CIDOC CRM to be aligned with the realities of archaeological entities: everything described in an inventory is the product of a series of events, making it possible to link an artefact to the various people who made, modified, transformed, acquired or destroyed it (Le Bœuf, 2013:6). The coherence of this series of events, once confirmed, gives the CIDOC-CRM ontology its meaning. The RDF coding provides a complete view of the fields and facilitates the production of a synthetic conceptual model.

3.2.3 Conceptual models of data and specific features of the Fait and Mobilier tables

Both conceptual models, for the *Fait* and *Mobilier* tables (Figs. 6 and 7), have common structuring elements, starting with the key class concerned ('concept'), the metadata linked to the excavation activity (the categories 'actor', 'event' and 'place'), and documentation and inventory information. Adjustments were made to each field to keep it as close as possible to its primary meaning. For example, we can see that the class E7.Activity {*activité d'inventaire*} was added to distinguish between the fait or

Metadata Categories	ArSol fields	Triples from CIDOC CRM
Physical description THING	Matière	E22 Man-Made Object {Mobilier}. P45 consists of: E57 Material {Matière} E57 Material {Matière}. P1 is identified by: E41 Appellation → value example = "cuivre"
	diam_mm	E22 Man-Made Object {Mobilier}. P43 has dimension: E54 Dimension {diam_mm} E54 Dimension {diam_mm}. P2 has type: E55 Type {diamètre} E55 Type {diamètre}. P1 is identified by: E41 Appellation → datatype = "*diamètre*" et : E54 Dimension {diam_mm}. P91 has unit: E58 Measurement Unit {mm} E58 Measurement Unit {mm}. P1 is identified by: E41 Appellation → about = "*millimètre*" et : → value example = '*20*"

Figure 5: RDF triples from the CIDOC CRM for the field 'diam_mm' in the Mobilier table.

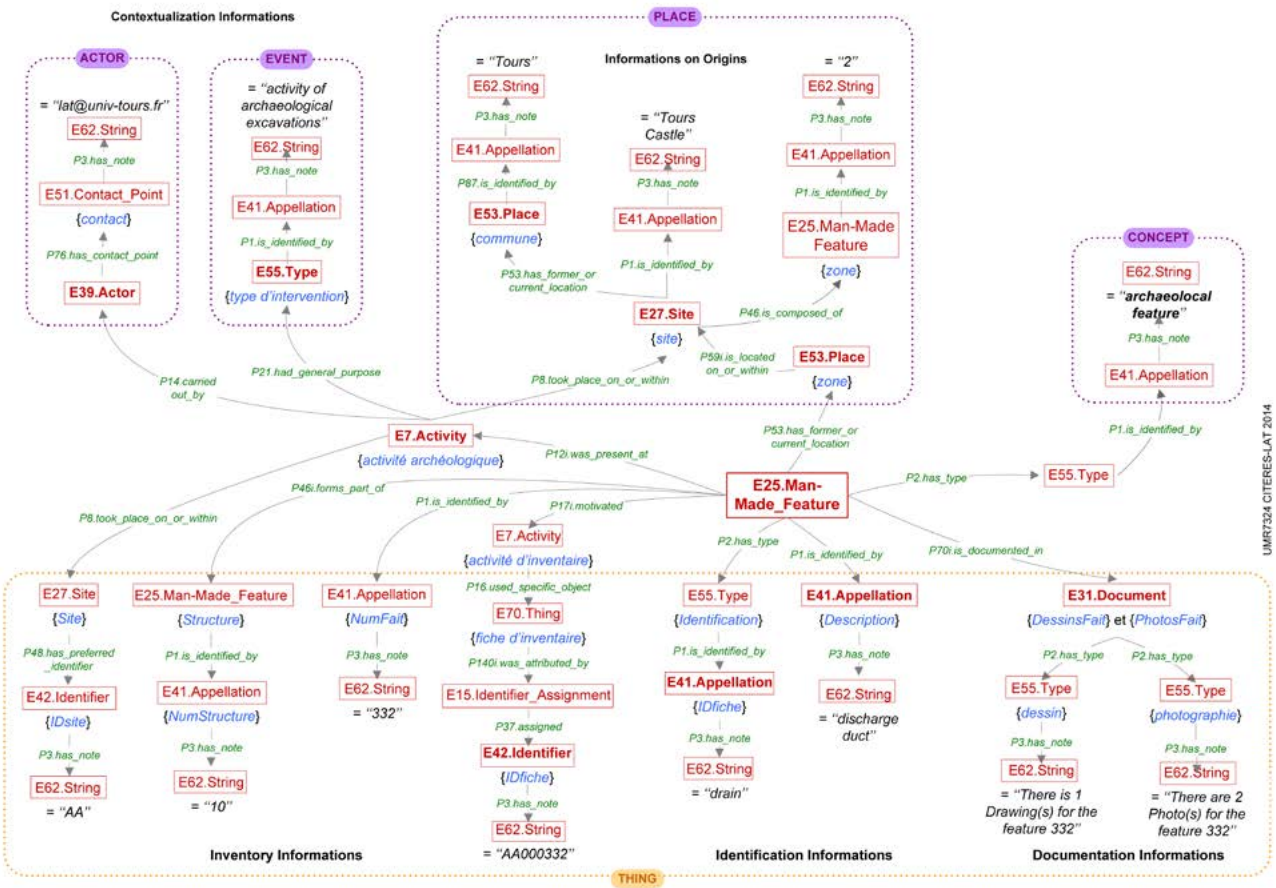


Figure 6: Conceptual model of data for the Fait table.

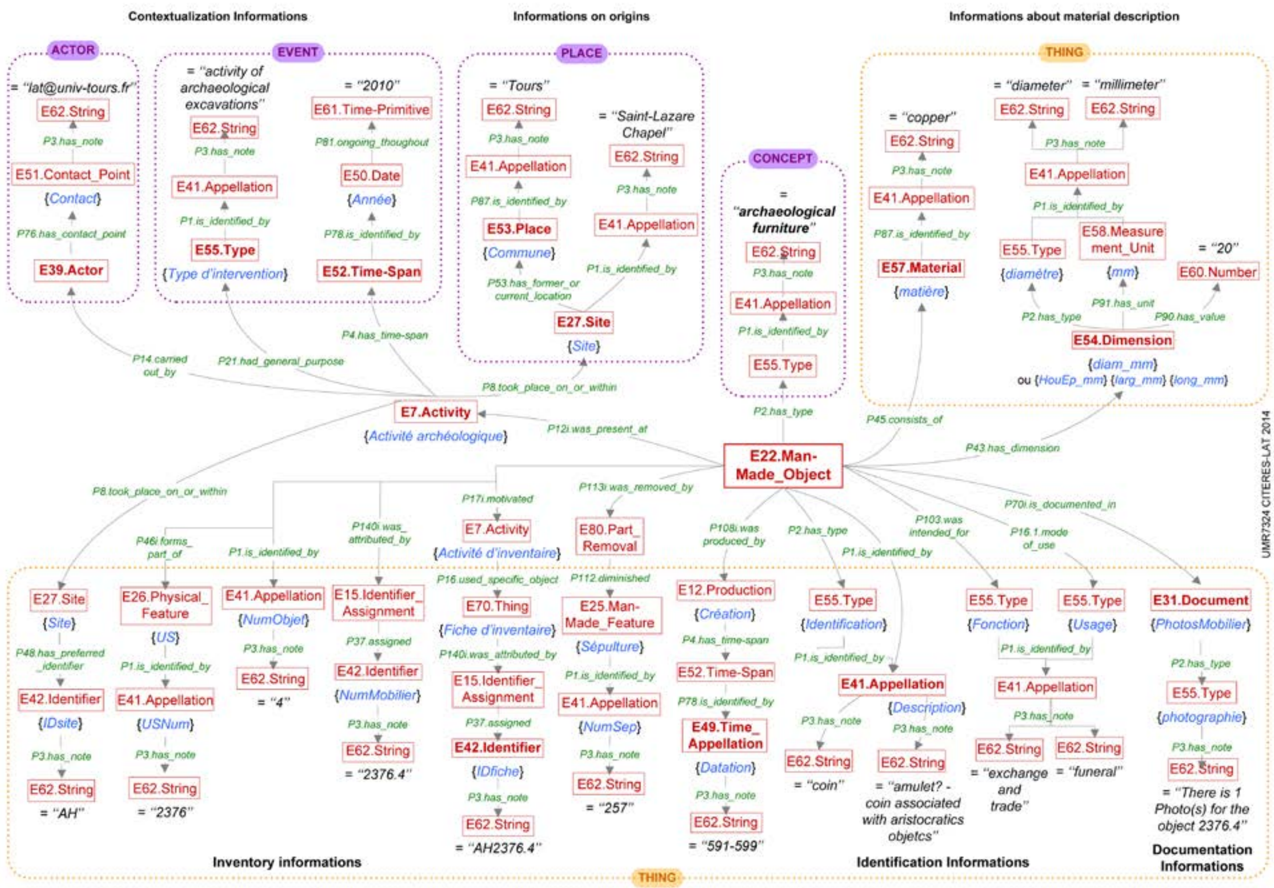


Figure 7: Conceptual model of data for the Mobilier table.

mobilier number and the inventory number in the record situated in the database.

This inventory number (IDfiche field) was created by concatenating the site inventory number and the *fait* or *mobilier* number. It was then decided to provide details of each event involved in the production or modification of the archaeological entity. The classes E25.Man-Made_Feature and E22.Man-Made_Object were first linked to the pivotal class E7.Activity, since the unearthing of a feature or artefact is clearly the result of the excavation activity. Based on the same reasoning (Fig. 4), the *Datation* field (E49.Time_Appellation) was linked to E12.Production to indicate that it was indeed a question of dating the creation of the artefact. With regard to the *Fonction* and *Usage* fields, the introduction of the class 'E55.Type' expanded from the widely used property 'P2.has_type' to the properties 'P103.was_intended_for' and 'P16.1.mode_of_use'. Finally, the *Mobilier* table also indicates the material and dimensions of the artefact, which indeed seem to be the only two classes concerning the 'Thing' category for which there is a direct match (E57.Material and E54.Dimension).

3.2. Mapping

3.3.1 Archive vs resources: divergent concepts in terms of documentation

Unlike museum or documentary objects, which are closely linked to information such as author/producer, reproduction, site of origin, repository, dating, collection, iconography, raw material (Kummer, 2010:186; Dentamaro *et al.,*, 2007:438-439; Szabados, 2012:12), archaeological entities, while including some of these elements (notably dating and raw material), have a socio-cultural component. Consequently, the structure of the ArSol database focused not only on their 'Identification', 'Description', and 'Dating', but also on 'Function' and 'Usage'. Thus, although each record can be supported by a document, drawing or photograph (E.31.Document), the documentation produced for the archaeological object is not in itself an object of study (Doerr, 2003 :76). Consequently, the characteristics of the document are not mentioned, nor for example the place where the artefacts are stored.

3.3.2 Difficulty managing the semantics specific to the archaeological domain

We give an example dealing with the two entities, Fait and *Mobilier*, to highlight the importance of the context of the discovery (Fig. 8), but a fait is not necessarily linked to the presence of *mobilier*, and vice versa. Our example concerns a coin (E22.Man-Made Object) linked to a burial site (E25.Man-Made Feature), unearthed near Tours cathedral (E.27.Site) during an excavation in 2010 (E52.Time-Span). Here, the context of the discovery is distinct from the primary purpose of the artefact, because a coin, initially produced for a trading activity (P103.was_intended_for → E55.

Type) was re-used as part of a funerary activity, and thus diverged from its primary function (P16.1.mode_of_use → E55.Type). It is the two latter fields that give this artefact its full meaning; however, they are characterized by the same class (E55.Type). It seems that the precision of the meaning of the fields is in fact borne by properties whose co-domain is ‘E55.Type’. This is one of the assets of the CIDOC CRM, but nevertheless the class ‘E55.Type’ can tend to weaken the meaning of the field and the conceptual schema of the data. To map the *Mobilier* table, we had to resort to the ‘E55.Type’ class eight times, notably for everything concerning elements identifying the *mobilier*, the distinctive features of the artefacts, and the context of their discovery.

3.3.3 How can we be sure of the conformity of the mapping?

We turn now to the issue of validating the mappings carried out by laboratories and their conformity with the CIDOC CRM. Insofar as the ontology has acquired the status of an ISO norm, the reference document becomes a tool for documenting applications based on its use (Szabados *et al.*,, 2012:7). It is true that this structure, and especially the wealth of properties, remain the strength of the ontology. However, it is unfortunate that there is no application to validate the deployment of the hierarchy of classes of the CIDOC CRM or the use of the properties according to this hierarchy (Domain or Range). However, this does not seem to be possible when there can be as many ways of mapping as there are ways of conceiving the archaeological entity (feature or artefact), or rather the events involved in its management. It is for this reason that we asked P. Le Boeuf, who is well versed in the ontology, to coordinate the mapping phase of the two ArSol tables.

4. Meeting the needs of interoperability: Current situation and future prospects

4.1. Development of tools compatible with the CIDOC CRM

First of all, a script has been developed to allow access to the content of each *Mobilier* and *Fait* record, based on the architecture used for the mapping phase, in line with the CIDOC CRM, in XML-RDF format (Fig. 9). By authorizing the conversion of each record to the CIDOC CRM norms, this tool also enabled us to check the efficiency of the mapping.

Secondly, the aim is to exploit this CIDOC CRM compliant mapping for the whole of the ArSol base – a pre-requisite phase to bring ArSol into the semantic web (Fig. 9). To achieve this, an OBDA (Ontology Based Data Access) system will make external interrogation of the ArSol database possible from the CIDOC CRM. Thanks to this OBDA layer, the ontology becomes the access point for searching data, while leaving the information in the databases used by the researchers. An OBDA system is composed of: an ontology (semantic level), various data sources, a set of mappings showing the relationships between source data and the ontology, and an application layer to manipulate and query the system.

This stage is currently under development based on the use of *-ontop-*, created by the University of Bozen-Bolzano (Italy), and which is an extension for the ontology editor Protégé created by Stanford University (California).

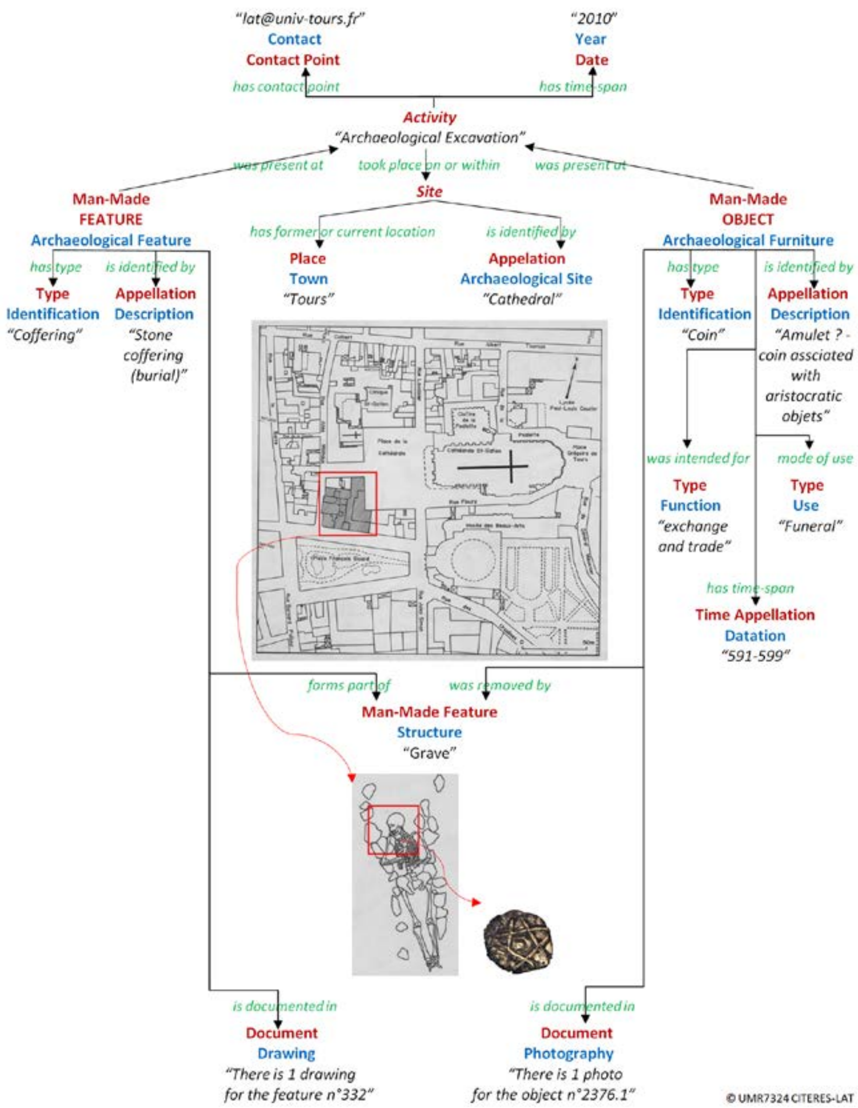


Figure 8: Example for a coin found in funerary context.

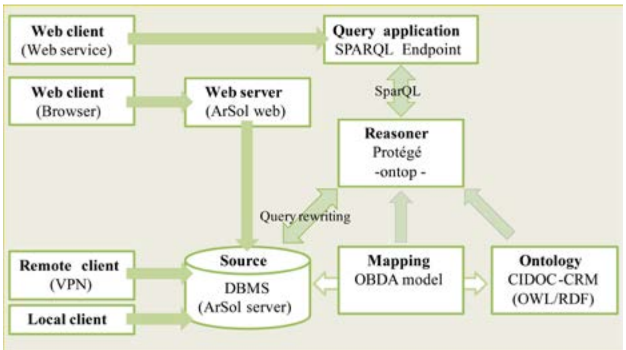


Figure 9: Project to bring ArSol to the semantic web with an ODBA system.

4.2. Accessibility of the database and export tools

ArSol is now on line (http://arsol.univ-tours.fr), enabling scientific data to be accessed via the Web by connecting to the database server used by LAT researchers. The information is filtered so that the visitor does not see information specific to database management, notably concerning storage or management of man-made objects and field archives.

To avoid exporting the base in a standard form but to keep it in its current form, which is perfectly mastered by the researchers, and without having to worry about up-dates, the interoperability of ArSol will stem from its compatibility with an application enabling it to be queried externally. It will thus be possible to formulate queries in a unified manner based on the CIDOC CRM ontology, alongside other archaeological bases, without having to change the original structure.

5. Conclusion

Mapping ArSol data on the CIDOC CRM ontology is specific to the information structure used by the *Laboratoire Archéologie et Territoires* but is widely shared by others. While it cannot be reproduced as such for other archaeological databases, the experience can be shared.

In spite of the difficulty of aligning some ArSol fields with CIDOC CRM categories, bringing the databases into line is indispensible for knowledge sharing. The interoperability of applications requires semantic homogenization to allow data to be queried and shared.

For the next stage of our work, we plan to add the SKOS format (Simple Knowledge Organization System) to the RDF version using the dedicated tool -*ontop*-, in order to improve the interoperability of data in the web-semantic framework (export and interrogation), meeting the recommendations of the W3C (World Wide Web Consortium) as closely as possible.

Acknowledgements

The This research and development work is supported by the MASA consortium of the Huma-Num TGIR (http://www.huma-num.fr). The web development of ArSol database is supported by SCREVO project funded by the Région Centre (http://www.regioncentre.fr).

We would like to thank Patrick Le Bœuf (Bibliothèque Nationale de France, Paris) for his invaluable and tireless help in the mapping phase of the ArSol database tables, and also Anne-Violaine Szabados (CNRS UMR7041 - ArScan, Université de Paris 10, MAE, Nanterre), Rosemonde Letricot (CNRS, Institut de Recherches et d'Histoire des Textes, Paris) and Andrea d'Andrea (Centro Interdipartimentale di Servizi di Archeologia, Università degli Studi di Napoli L'Orientale, Vicoletto I, Italy) for sharing their experience of using the CIDOC CRM. We also thank Béatrice Bouchou Markhoff (Université François-Rabelais, Laboratoire d'Informatique, Blois) for recommending the use of ODBA and -*ontop*-, and for her commitment to the development of this project.

References

ANDREUSSI, M. and FELICETTI, A. (2008). The CIDOC CRM Encoding of the 'Fontes ad Topographiam Veteris Urbis Romae Pertinentes' by Giuseppe Lugli. In: Posluschny, A., Lambers, K. and Herzog, I. (eds.) *Layers of Perception. Proceedings of the 35th CAA Conference, Berlin, Germany, 2-6 April 2007*. Bonn: Dr. Rudolf Habelt GmbH, pp. 1-6.

BACHIMONT, B. (2000). Engagement sémantique et engagement ontologique : conception et réalisation d'ontologies en Ingénierie des connaissances. In: Charlet, J., Zacklad, M., Kassel, G. and Bourigault, D. (eds.) *Ingénierie des connaissances, évolutions récentes et nouveaux défis*, Paris: Eyrolles, 16 p.

CRUZ, I.F. and XIAO, H. (2005). The Role of ontologies in data integration. *Engineering intelligent systems* 13 (4), pp. 245-252.

CUENCA GRAU, B. and GIESE, M. (2013). *Towards Query Formulation and Query-Driven Ontology Extensions in OBDA Systems*. 12p. Available from: http://www.cs.ox.ac.uk/files/5511/owled2013_wp3.pdf [Accessed: 20th July 2014].

D'ANDREA, A., NICCOLUCCI, F. and FERNIE, K. (2012). *3D-ICONS: European project providing 3D models and related digital content to Europeana*. Florence: EVA, 6p. Available from: http://fr.slideshare.net/3dicons/3dicons-european-project-providing-3d-models-and-related-digital-content-to-europeana [Accessed: 20th July 2014].

DENTAMARO, F., DE LUCA, P.G., GENCO, L., PERRINO, G., CANNITO, C., STUFANO, M.A. and SIBILANO, M.G. (2007). A CIDOC CRM-Based Ontology System. In: Jeffrey, C., Jeffrey, T. and Hagemeister, E.M. (eds.) *Digital Discovery. Exploring New Frontiers in Human Heritage, Proceedings of the 34th CAA Conference, Fargo, United States, April 2006*, p. 437-444.

DOERR, M. (2003). The CIDOC Conceptual Reference Model: An Ontological Approach to Semantic Interoperability of Metadata. *AI Magazine* 24 (3), p. 75-92.

ECKKRAMMER, F., FELDBACHER, R. and ECKKRAMMER, T. (2011). Cidoc CRM in Data Management and Data Sharing. Data Sharing between Different Databases. In: Erzsébet, J., Redő, F. and Szeverényi, V. (eds.). *On the Road to Reconstructing the Past. Proceedings of the 36th CAA Conference, Budapest, 2-6 April 2008*, Budapest: Archaeolingua, pp. 80-85.

GALINIE, H., HUSI, P., RODIER, X., THEUREAU, C. and ZADORA-RIO, E. (2005). ARSOL. La chaine de gestion des données de fouilles du laboratoire Archéologie et Territoires, *Les petits cahiers d'Anatole* 17, 27/05/2005, 25 p. Available from: http://citeres.univ-tours.fr/doc/lat/pecada/F2_17.pdf [Accessed: 20th July 2014].

HOLMEN, J., ORE, C.E. and EIDE, Ø. (2004). Documentation two histories at once: Digging into archaeology. In: Ausserer, K.F., Börner, W., Goriany, M. and Karnlhuber-Vöckl, L. (eds.). *Enter the Past, the E-way into Four Dimensions of Cultural Heritage, Proceedings of the CAA Conference 2003*, Oxford: Archaeopress, pp. 221-224.

HOLMEN, J. and ORE, C.E. (2010). Deducing Event Chronology in a Cultural Heritage Documentation System. In: Bernard, F., Crawford, J.W. and Koller, D. (eds.). *Making History Interactive, PComputer Applications and Quantitative Methods in Archaeology (CAA). Proceedings of the 37th International Conference, Williamsburg, Virginia, United States of America, March 22-26 (BAR International Series S2079),* Oxford: Archaeopress, pp. 122-129.

HUSI, P. and RODIER, X. (2011). ArSol: An Archaeological Data Processing System. In: Erzsébet, J., Redő, F. and Szeverényi, V. (eds.). *On the Road to Reconstructing the Past, Proceedings of the 36th CAA Conference, Budapest, 2-6 April 2008*, Budapest: Archaeolingua, pp. 86-92.

JORDAL, E., ULEBERG, E. and HAUGE, B. (2012). Was It Worth It? Experiences with a CIDOC CRM-based Database. In: Zhou, Mingquan, Iza Romanowska, Zhongke Wu, Pengfei Xu and Philip Verhagen (eds) 2012. *Revive the Past. Computer Applications and Quantitative Methods in Archaeology (CAA). Proceedings of the 39th International Conference, Beijing, April 12-16.* Pallas Publications, Amsterdam, pp. 255-260.

KUMMER, R. (2010). Archaeology and the Semantic – Prospects and Challenges. In: Bernard, F., Crawford, J.W. and Koller, D. (eds). *Making History Interactive, PComputer Applications and Quantitative Methods in Archaeology (CAA). Proceedings of the 37th International Conference, Williamsburg, Virginia, United States of America, March 22-26 (BAR International Series S2079),* Oxford: Archaeopress, pp. 178-190.

LE BŒUF, P. (2009). Modélisation conceptuelle de l'information bibliographique et muséologique : CIDOC CRM et FRBROO. *Documentation et bibliothèques* 55 (4), p. 209-214. Available from: http://hal-bnf.archives-ouvertes.fr/hal-00807980 [Accessed: 20th July 2014].

LE BŒUF, P. (2013). *De la sémantique des inventaires aux musées en dialogue : la modélisation CIDOC CRM*. Avril 2013. Available from: http://halshs.archives-ouvertes.fr/hal-00807664/ [Accessed: 20th July 2014].

LE BOEUF, P., DOERR, M., ORE, C.E. and STEAD, S. (2013) *Definition of the CIDOC Conceptual Reference Model*. Version 5.1.2. October 2013, 186p. Available from: http://www.cidoc-crm.org/docs/cidoc_crm_version_5.1.2.pdf [Accessed: 20th July 2014].

LOURDI, I., PAPATHEODOROU, C. and DOERR, M. (2009). Semantic integration of collection description: combining CIDOC/CRM and Dublin Core Collections Application Profile. *D-Lib Magazine* 15 (7/8). [Online] Available from: http://www.dlib.org/dlib/july09/papatheodorou/07papatheodorou.html [Accessed: 20th July 2014].

MCGUINNESS, D.L. and VAN HARMELEN, F. (2004). *OWL Web Ontology Language Overview*. [Online] Available from: http://www.w3.org/TR/owl-features/ [Accessed: 20th July 2014].

RODRIGUEZ-MURO, M. and CALVANESE, D. (2008). *Towards an open framework for Ontology Based Data Access with Protégé and DIG 1.1*. 4p. Available from: http://ceur-ws.org/Vol-432/owled2008eu_submission_37.pdf [Accessed: 20th July 2014].

RODRIGUEZ-MURO, M., KONTCHAKOV, R. and ZAKHARYASCHEV, M. (2013). *OBDA with Ontop*. 6p. Available from: http://www.dcs.bbk.ac.uk/~roman/papers/ORE1.pdf [Accessed: 20th July 2014].

RODRIGUEZ-MURO, M., KONTCHAKOV, R. and ZAKHARYASCHEV, M. (2013). *Ontop at work*. 12p. Available from: http://webont.org/owled/2013/papers/owled2013_18.pdf [Accessed: 20th July 2014].

SZABADOS, A.-V. (2012). Du système documentaire du LIMC au portail CLAROS. Interopérabilité et optimisation de l'information archéologique grâce à l'usage des normes. In: Giligny, F., Costa, L., Djindjian, R., Ciezar, P. and Desachy, B. (eds.). *2e Journées d'Informatique et Archéologie de Paris - JIAP 2010, Paris, France, 11-12 juin 2010, Archaeologia e Calcolatori*, Suppl. 3, p. 11-25.

SZABADOS, A.-V. and LETRICOT, R. (2012). L'ontologie CIDOC CRM appliquée aux objets du patrimoine antique. In: Giligny, F., Costa, L., Djindjian, R., Ciezar, P. and Desachy, B. (eds.). *3e Journées d'Informatique et Archéologie de Paris - JIAP 2012, Paris, France, 1-2 Juin 2012*, 14p. Available from: http://halshs.archives-ouvertes.fr/halshs-00752996 [Accessed: 20th July 2014].

Find the Balance - Modelling Aspects in Archaeological Information Systems

Franz SCHWARZBACH,[1] Felix F. SCHÄFER,[2] Alexander SCHULZE[3]

[1] Dresden University of Applied Sciences, Germany
[2] German Archaeological Institute, Germany
[3] Brandenburg University of Technology, Germany
schwarzbach@htw-dresden.de, felix.schaefer@dainst.de, schulal2@tu-cottbus.de

Abstract

The article describes specific aspects of data modeling in archaeological information systems. The key message is that the implementation of highly specialized models is extremely difficult or even impossible under certain conditions. This is illustrated by, among others, using an example from a non-archaeological domain (INSPIRE).
Accordingly, a solution is needed that is sufficiently general, generic and thereby flexible, but which also supports the domain-specific standards with the aim of an interoperable exchange of data.
Finally, one example for such a data model is presented. Based on concrete project experience advantages and disadvantages of this solution are discussed. In this context the balance between a universal and therefore broadly usable data model on the one hand and the support of interoperability on the other hand is an important aspect.

Keywords: archaeological information systems, generic data models, interoperability

1. Problem

At the very beginning of developing an information system, one has to define the underlying data model.

Generally, the most straightforward approach to this is to implement an existing standardized application schema (e.g. CIDOC-CRM) including the associated implementation specifications. This approach provides a number of benefits:

- savings of analysis and modelling effort,
- subsequent use of available software and
- a high level of semantic and syntactic interoperability.

The alternative is the definition of a schema designed for a specific application: Traditionally, in a protracted and complex process, a project-specific data model is developed and described in a formalized language. Thereafter, the resulting Platform Independent Application Schema can be transformed into several Platform Specific Application Schemata like database schemata or XML schemata. This methodology has been developed by the Object Management Group (OMG) as 'Model Driven Architecture' (Figure 1).

Typically the Platform Independent Application Schema is very sophisticated, fine-grained and highly specialized. It is very likely that it cannot be re-used for other projects or purposes. Additionally, it is more or less inflexible and consequently subject of frequent updates. As a result, derived Platform Specific Application Schemata must also be changed and probably one has to migrate data, too. Therefore these updates are expensive.

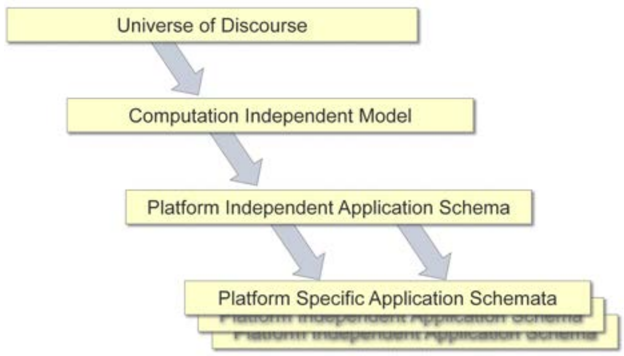


Figure 1: Model Driven Architecture.

- One has to know the stakeholders, especially the data producers and potential users and their requirements.
- One needs an agreement regarding a shared professional point of view on the universe of discourse and the corresponding scope.

Unfortunately these conditions are not always met as the two following examples show.

The first example is INSPIRE (http://inspire.ec.europa.eu). The acronym stands for Infrastructure for Spatial Information in the European Community and defines the aim of this project of the European Commission. INSPIRE is based on a European law ('DIRECTIVE 2007/2/EC of the European parliament and of the council') and covers a huge scope. Both the legislation and the implementation are taking many years to complete. In detail INSPIRE comprises 34 so-called 'Themes'. In recent years, data specifications were adopted for each of these themes which are often very broad. With the development of the data specifications, both different professional requirements and a variety of specific situations in the countries of the European Union had to be taken into account. These conditions led to (at least partly) comprehensive data models.

In other cases, a harmonization of the different points of view could not be reached and thus the data models are kept more general. One example for this is the sub-theme 'Administrative and Social Governmental services' (European Commission, 2013) which, at first glance, seems to be easy to model. Apart from the question: 'What is a governmental service?', existing data of all member states of the European Union had to be considered, where the majority of the data is under the responsibility of municipalities. Finally it was impossible to provide an application schema for such fuzzy scope and content. As result, a model with only one class (Figure 2) plus belonging code lists and data types was defined.

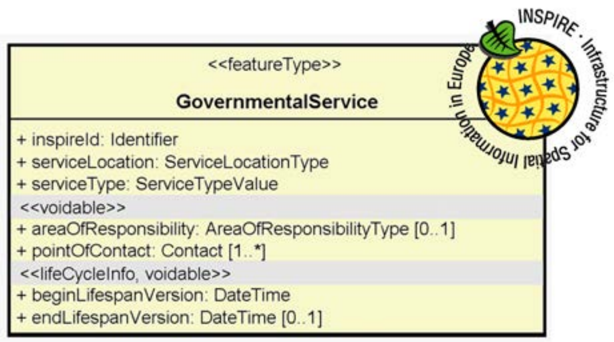


FIGURE 2: INSPIRE DATA MODEL ' ADMINISTRATIVE AND SOCIAL GOVERNMENTAL SERVICES ' (CUTOUT).

2. Approach in a special project

The second example is OpenInfRA (http://www.tu-cottbus.de/projekte/de/openinfra/). OpenInfRA is a cooperative project of two German universities and the German Archaeological Institute. It is funded by the 'Deutsche Forschungsgemeinschaft' as part of a program 'Information infrastructures for research data'. The project idea was already presented on the CAA conference in Southampton, 2012 (Schulze *et al.*, 2012). OpenInfRA shall be completely implemented in 2015.

The overarching aim of OpenInfRA is to provide a research data infrastructure component, which focuses on data capturing, storage and presentation of primary research data and which can be used for documentation purposes in different projects. Several methods relevant for archaeological and historical fieldwork as well as, for the time being, unknown users have to be taken into account. Therefore the starting position is comparable to the case above: fuzzy scope and content. Another goal is to keep the obstacle for potential users as low as possible. In particular, the users shall get the possibility to keep working according to their own professional point of view. As a result of these considerations, flexibility and adaptability are key aspects of OpenInfRA.

The conclusion at this stage was, that the preconditions mentioned above (knowledge about the stakeholders and a common professional perspective) will not be met, when the information system needs to be flexible enough to be used by numerous projects, in different areas of application and for user groups with non-harmonized professional perspectives. This situation typically occurs, when the information system is offered as a service in a digital infrastructure, e.g. for geographic data or archaeological research data. Therefore a solution was needed which is general, generic and flexible but which also supports domain-specific standards with the aim of an interoperable exchange of data.

As a result of these considerations, a highly generic data model was defined. Figure 3 depicts the main part of the application schema (UML class diagram).

The application schema comprises only a few classes: There are topics (themes, classes) (e.g., 'room') with properties (e.g. 'function of the room') and corresponding instances (objects) (in that case: a specific room). 'Properties' include both attributes and relationships between topic instances. The project-specific configuration of the topics regarding attributes, relationships, domains, etc. is, in the first instance, completely free. Based on the experience from legacy systems, a set of predefined topics, including the associated properties is provided. Users can freely decide to adopt these existing themes and properties, but in principle this approach allows users also to implement their own concept of, for example, 'room'. From the user's perspective the system is therefore 'open'. Since there are just a few limitations in the data model, the obstacle to use OpenInfRA is very low.

In order to create a basis both for cross-project analysis and for an interoperable data exchange, the range of the majority of model elements is limited to pre-defined values. These values are implemented as items in a single structured value list (left class in figure 4), where each item has a reference to its type (right class in figure 4). Within individual projects, both lists can be limited or expanded. The corresponding UML class diagram is shown in Figure 4.

The last part of the model represents the desired multilingualism. To fulfil this requirement, the conceptual point of view of ISO 19139 (2007) has been implemented. ISO 19139 is a standard in the field of Geographic Information and focuses on the topic metadata. One advantage of this approach is that within ISO 19139 an XML mapping is already provided (Figure 5).

It is obvious that this data model isn't directly applicable in special projects. In fact, it is a meta model, which has to be adapted first. This includes the compilation of topics, assignment of properties or customization of value lists as part of the project initialization (Figure 6).

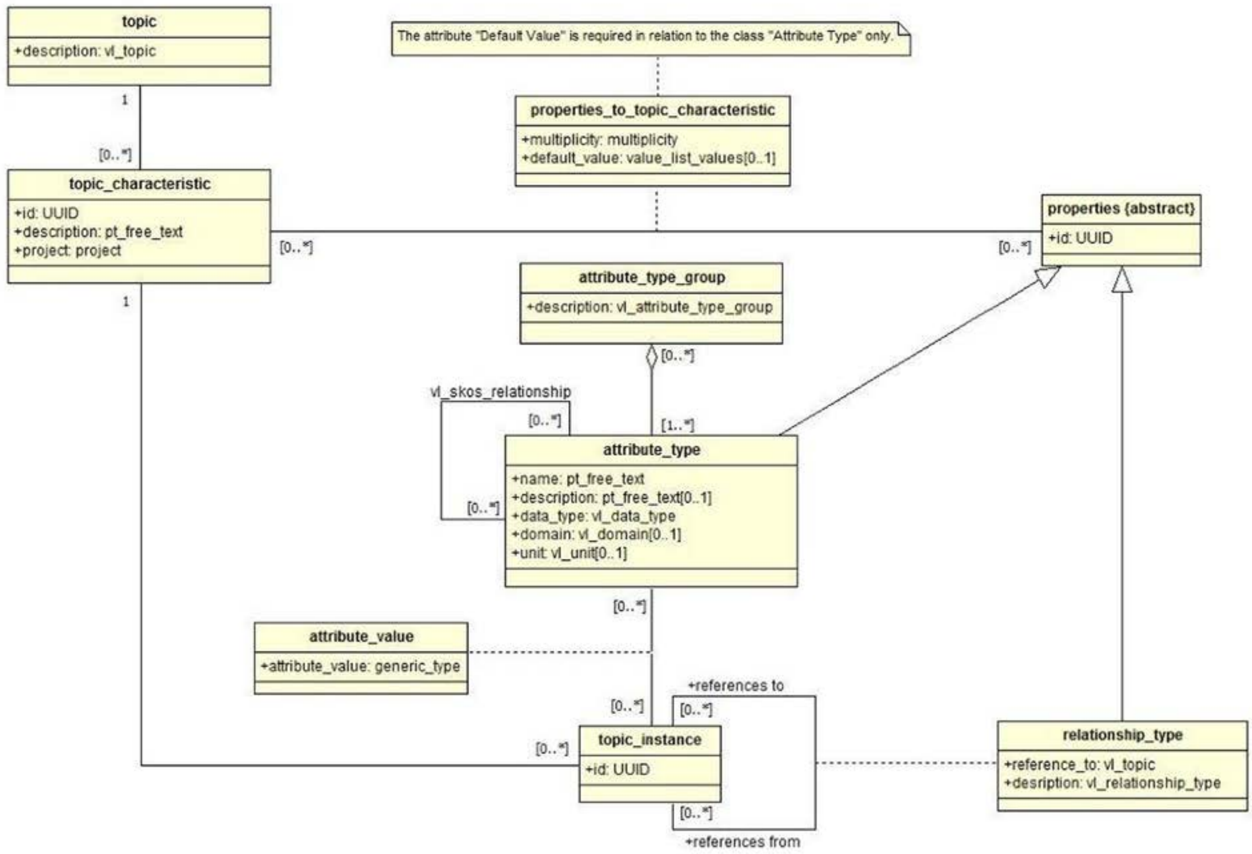


FIGURE 3: OPENINFRA CLASS DIAGRAM.

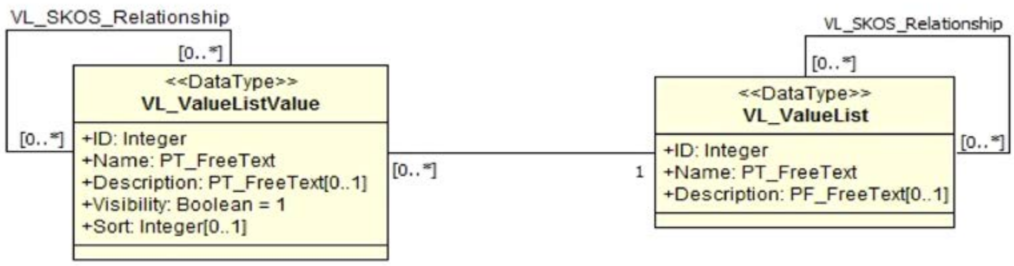


FIGURE 4: OPENINFRA VALUE LISTS.

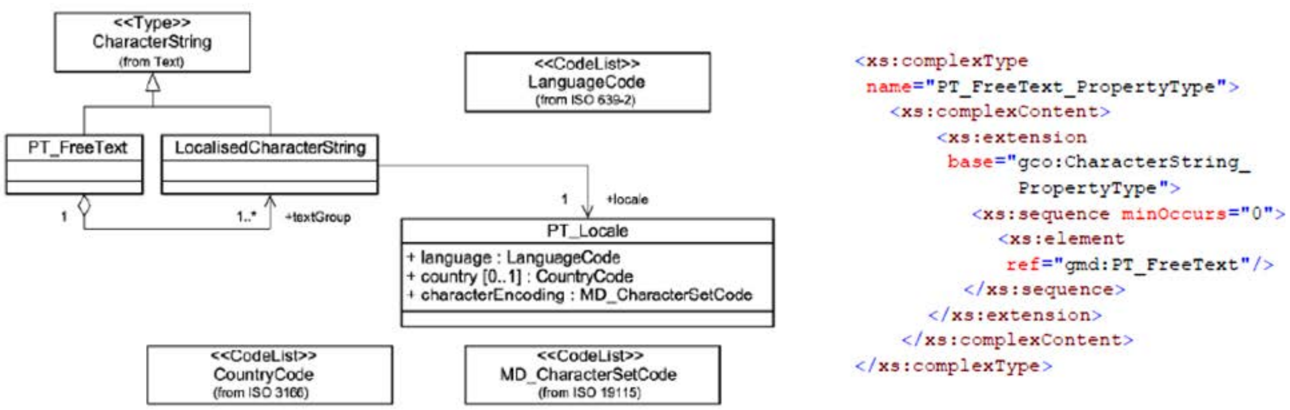


FIGURE 5: MULTILINGUALISM ACCORDING TO ISO 19139.

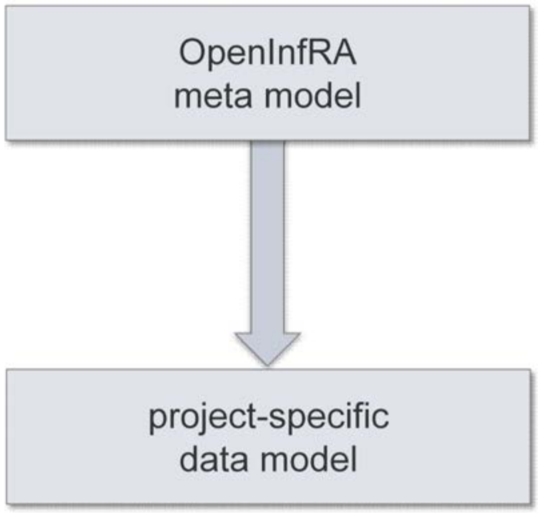


FIGURE 6: GENERATING OF PROJECT-SPECIFIC DATA MODELS.

3. Discussion of impacts – lessons learned

From implementation viewpoint (database, application, interfaces, graphical user interfaces) the most important advantage of the approach taken by OpenInfRA is that its data model is highly flexible, robust and sustainable. One has to consider, that changes in the data model usually are expensive, since they affect the whole system. This effect is completely eliminated by the generic approach described here: Both the professional view on the universe of discourse as well as its scope is mapped to the data level, not to the schema level. Therefore newly added topics, attributes and relationships require no changes to the data model and the corresponding software components.

Nevertheless, there are also a number of problems: The first one relates to the aforementioned instantiation of the meta model as part of the project initialization. When a project starts, the project-specific data model must be derived from the meta model first. So, the project administrator is burdened with modelling effort.

Another aspect is even more important:

The essential function of a formalized application schema within an IT project is to provide a common understanding of analysts, developers and users regarding the universe of discourse. Even with knowledge of UML the meta model is not understood spontaneously by the stakeholders. E.g., the 'meta-model-approach' with a high level of abstraction led to problems during the migration of data from legacy systems into the OpenInfRA database. It is obviously not easy to achieve an understanding of this approach in order to be able to map existing data into the OpenInfRA model.

The next problem concerns the database. Since the application schema is comparatively small, the associated database was implemented fairly quickly. Meanwhile the database (PostgreSQL/PostGIS) has been operable for about two years. During this time some issues have arisen, e.g.:

- More constraints than usual are needed. The reason is that the model contains both types and instances. These two levels have to be kept consistent. E.g., a topic instance can only enter into a relationship with another topic instance, when both the topic of the relationship partner and the relationship type is pre-defined for the associated topic.
- Since the attribute types are generic, this also applies to their data types. E.g., one has no priori knowledge of whether a particular attribute has the data type 'text' or 'geometry'. The database design had to take this into account.
- The database contains only a few tables, but with a large number of records. Therefore, performance tests are executed in parallel.

And, last but not least: In general the use of a standardized application schema, as it is not the case here, facilitates the data exchange between different systems considerably. That aspect is discussed in the next paragraph.

4. Interoperability aspects

An important challenge for modern information systems is to combine systems not only technically, but also semantically. In this context the first objective in OpenInfRA is to link various projects together in order to enable a cross-project analysis. This aim is supported by the set of predefined topics and values lists mentioned above.

Furthermore, due to the lack of a common conceptual basis, an infrastructure has to be provided in order to map data from OpenInfRA projects to other data models and vice versa. This issue is solved on the basis of standard XML technology. The conceptual schema has been mapped (parallel to the database schema) to an XSchema. The required interfaces are implemented as XSL transformations:

The XSL transformations shown in the figure are operable. Currently, the interfaces are used to transfer data from legacy systems into the OpenInfRA database.

The main advantages of this solution are that

- no new software needs to be developed for the support of other data models, only a stylesheet has to be written and
- the correctness of the transformed data can be checked (within the limits of an XML validation).

In the future, external data models will also be supported in this way. So far, there is a project-specific interface to CIDOC-CRM, but the corresponding style sheet covers at the moment only a part of the project data.

5. Summary

With the conception and design of (archaeological) information systems, the question arises whether existing standardized application schemata can be utilized,

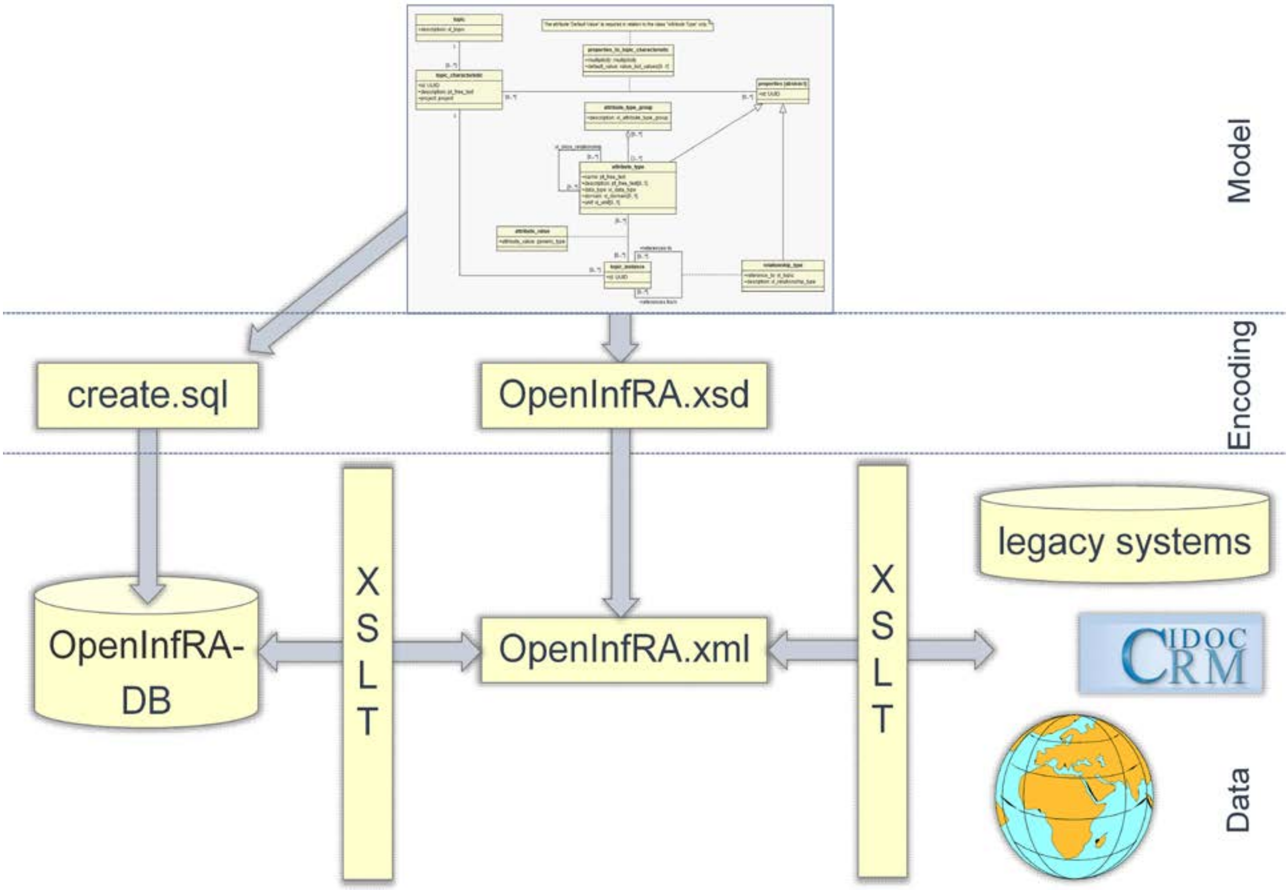


Figure 7: OpenInfRA XML Interfaces.

including the associated implementation specifications. In other cases, an existing conceptual schema cannot be used as a fall back option. Traditionally a project-specified data model will then be developed and described. A highly specialized but more inflexible data model also arises as a result of this analysis.

Both approaches require the fulfilment of certain conditions. The article has shown, that these conditions are not always met, especially in the context of digital infrastructures. Under certain conditions a generic data model should be preferred compared to the use of a standardized or project-specific developed application schema. It has been illustrated, that both the data model and the corresponding database are small, flexible to use and robust regarding changes in user requirements.

The advantages and disadvantages outlined in the article must be evaluated in the context of specific data and project objectives. It should be emphasized, that the generic model used in the project OpenInfRA should not be understood as a template for arbitrary archaeological information systems. The triggers for its usage were the specific conditions of a digital infrastructure project.

References

European Commission. (2013). INSPIRE *Data Specification on Utility and Government Services – Technical Guidelines*. European Commission Joint Research Centre [Online]. Available at: http://inspire.ec.europa.eu/.

SCHULZE, A., SCHÄFER, F.F., GEHRT, P., HENZE, F. and SCHWARZBACH, F. (2012). OpeninfRA – *Storing and Retrieving Information in a Heterogeneous Documentation System*. In: E., Graeme, T. Sly, A. Chrysanthi, P. Murrieta-Flores, C. Papadopoulos, I. Romanowska and D.Wheatley (eds). *Archaeology in the Digital Era. Papers from the 40th Annual Conference of Computer Applications and Quantitative Methods in Archaeology (CAA), Southampton, 26-29 March 2012*. Amsterdam University Press, Amsterdam [Online] Avalaible from: http://dare.uva.nl/cgi/arno/show.cgi?fid=516092

International Organization for Standardization. (2007). ISO/TS 19139:2007. *Geographic information – Metadata – XML schema implementation*. Geneva: ISO.

Integration of Archaeological Datasets Through the Gradual Refinement of Models

Cesar GONZALEZ-PEREZ, Patricia MARTÍN-RODILLA

Institute of Heritage Sciences (Incipit), Spanish National Research Council (CSIC), Spain
cesar.gonzalez-perez@incipit.csic.es, patricia.martin-rodilla@incipit.csic.es

Abstract
Archaeological data is usually stored as software datasets, which differ not only by their supporting technology but, most importantly, by the conceptualisations that underpin them. Any intellectual effort that tries to relate information from multiple datasets needs to be aware of such conceptual differences and achieve optimal integration for the results to be meaningful. Data comparison, unification or transfer are examples of this. Current approaches to dataset integration in archaeology suffer from several problems, including that of severe inattention to abstraction issues; this means that conceptualisations at very high levels of abstraction such as CIDOC CRM are routinely 'mapped' to others at very low levels of abstraction such as physical database schemata. In this paper, we propose an alternative approach for the integration of archaeological datasets through the gradual refinement of models; this approach caters for abstraction and allows model users to choose how much information they want to see, and how much detail they need.

Keywords: Data integration.Archaeological datasets.CHARM.Model refinement.Conceptual modelling.

1. Introduction

Archaeological data often resides in software-based datasets, i.e. computer-stored files that can be accessed through digital media. Datasets are extremely heterogeneous, ranging from large and complex infrastructures such as Europeana Cloud (Europeana, 2014) or the forthcoming ARIADNE infrastructure (ARIADNE, 2014) to small, project-specific local databases implemented as e.g. Microsoft Excel spreadsheets. The most obvious differences between datasets are those of the supporting technologies, and a number of efforts have been made to bridge the technological gap between heterogeneous datasets (May *et al.*, 2012; Meyer *et al.*, 2007; Eckkrammer *et al.*, 2008). However, non-technological differences and, in particular, conceptual differences, arguably have a larger impact on dataset interoperability, but have received a lower degree of interest from the community, perhaps due to the fact that they are not as evident, perhaps because they are harder to deal with, or perhaps due to the belief that the systematic conceptualisation of reality is difficult, impossible or not desirable within the humanities.

In any case, many interpretive efforts in archaeology work from the premise that multiple datasets are necessary in order to achieve a particular goal, thus giving rise to the problem of dataset integration. This is a well-known problem that has been studied in the literature, albeit with the above described strong focus on technological integration. The few works that emphasise conceptual integration are usually related to the adoption of standards such as ISO 21127:2006, i.e. CIDOC CRM (ISO, 2006) or the use of 'semantic web' approaches such as ontologies or thesauri, e.g. (Isaksen *et al.*, 2010).

In this paper we argue that these approaches, while having contributed very useful insights to the problem of dataset integration, suffer from a number of issues, most importantly that of being oblivious to abstraction problems. In Section 2 below we describe the details of these problems and explain why current approaches are insufficient. Section 3 describes the proposed solution, based on the gradual refinement of models. Section 4 illustrates this solution through an example that integrates three archaeological datasets at different levels of abstraction and shows how users would benefit from this integration. Section 5 discusses some benefits and limitations of this approach, and Section 6 presents some conclusions.

2. Motivation and Context

Literature on archaeological data integration is not scarce. However, the proposed approaches usually suffer from one or many of the following issues. Most are heavily oriented towards data rather than information, not to mention knowledge. It is true that computers often store data, and information (or knowledge) is inferred by humans from these data through processes that are only mildly mediated by software or digital technology (but see (Martín-Rodilla, 2013) for a counterexample). Data is the most visible and easy to manipulate, being what is directly stored, and this arguably explains why most efforts towards integration focus on data rather than higher levels of abstraction. It is interestingly paradoxical that the so-called 'semantic technologies', often embodied by linked data approaches, focus so much on data and lexical connections between symbols, and so little on the actual connections between these symbols and whatever they refer to (i.e. their *semantics*). Works such as (May *et al.*, 2012; Meyer *et al.*, 2007; Isaksen *et al.*, 2010; Eckkrammer *et al.*, 2008) clearly focus on the very specifics of the data being stored and how to map from one data container to another, but rarely mention the connections between data and the entities they represent.

Along similar lines, current approaches devote large efforts to reconciling the technological differences between datasets, without realising that technological solutions come and go very quickly over the years. Whether an archaeological dataset is backed as a relational database, a SOAP web service, a SPARQL endpoint or a public Microsoft Azure blob container, all these technologies will have most likely changed in a few years' time, and any efforts done today to integrate datasets at the technological level will become wasted time and money. As archaeologists, we should leave technological integration issues to systems engineers, and use our expertise to address problems that only we can address: those of diverging archaeological conceptualisations.

Unfortunately, not many existing proposals to dataset integration directly consider *conceptual integration*. As a remarkable exception, (Kintigh, 2006) states that:

> '*Concept-oriented archaeological data integration will enable the use of existing data to answer compelling new questions and permit syntheses of archaeological data that rely not on other investigators' conclusions (that even their authors may now consider outdated) but on analyses of meaningfully integrated new and legacy data sets.*
>
> *In particular, because of archaeology's unique ability to provide centennial-and millennial- scale comparative data and comparative data from geographically dispersed areas, such a knowledge-based data-integration system would allow archaeology to contribute substantially to scientific understandings of long-terms social dynamics*.'

We cannot agree more. Making knowledge, rather than data, the focus of integration is thus crucial for attaining results that are meaningful to archaeologists.

It is still true, however, that most archaeological information is stored at a very low level of detail, i.e. as inter-linked atomic data items; this is the case, for example, of rows and fields in a relational database or triples in a triple store. We need, therefore, a reliable mechanism to manage abstraction in a disciplined manner, i.e. to work with representations of the relevant area of reality (the archaeological record, presumably) that leave out more or less detail depending on the specific representational purpose. In the absence of a mechanism like this, arbitrary jumps in abstraction level will most likely result in semantic mismatches that will be extremely difficult to amend or, even worse, go unnoticed.

Figure 1 illustrates this with an example. On the left-hand side, a screenshot of SIA+, an archaeological information system that has been used at Incipit CSIC for over 15 years, is shown. The data entry form for the 'Site' entity can be seen, including some of the fields that users need filling in. From a conceptual point of view, it is evident to the archaeologist using this system that SIA+ manages entities of the 'Site' type, and that sites, in SIA+, have a Code, a Type, a Chronology, some Dimensions and a Description; this is what the concept of a site looks like for an archaeologist at Incipit using SIA+.

However, if the SIA+ dataset where data about over 5,000 sites is stored were to be integrated with other dataset by using a data-oriented approach, an extremely different situation would be found. The database that underpins SIA+ does not contain any structure that recognisably maps to the conceptual 'Site' entity as described above. On the right-hand side of Figure 1, a fragment of the database schema is depicted, showing the tables and relationships that are used to store sites. Little resemblance to the clean and simple concept of 'Site' can be found, due to a number of technical details related to relational technology,

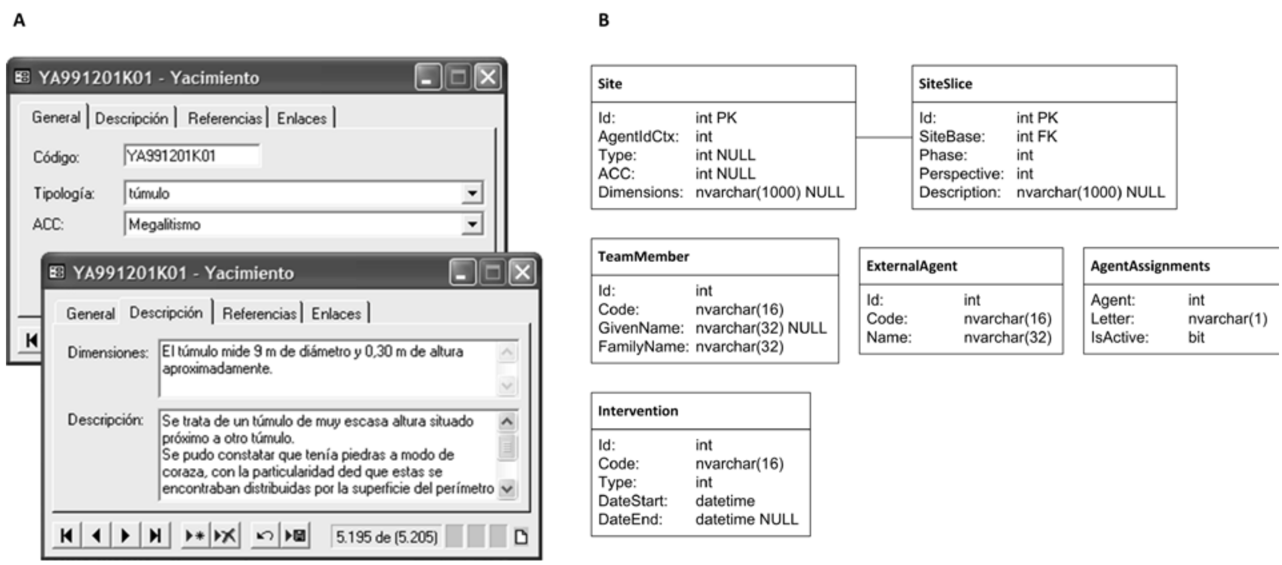


Figure 1: Contrast between abstraction levels. On (A), the data entry form for the 'site' entity of SIA+ can be seen, including fields for each site's Code, Type, Chronology, Dimensions and Description. On (B), a fragment of the underlying database schema is depicted, showing the tables, columns and relationships that are necessary to store the data that is input or retrieved by archaeologists through the previously described user interface.

variable links between tables, subjectivity and temporality management implementation, automatic code generation, and other 'implementation noise' that the user is rarely aware of.

The most important point that we would like to make with this example is not that going from the clear conceptualisation of a 'Site' that a user has to the apparently messy database implementation is difficult and arbitrary; while this is true, the most vital point here is that an approach that is strongly data-focussed would only look at the database, finding the complex, apparently unfathomable schema in Figure 1B, and would almost certainly fail to realise that this schema is supporting the 'Site' abstraction, among others. Unless the conceptual level is taken into account, the database would be hardly useful.

For these reasons, the commonly held belief that database schemata nicely represent the concepts in the users' minds is, at best, naïve, and very often plainly wrong. A better approach to archaeological dataset integration is necessary, one that is aware of abstraction issues and puts the emphasis on the conceptual level. The next section describes our proposal.

3. Proposed Solution

The proposed approach to archaeological dataset integration is based on the gradual refinement of models, and works along the following structure, which is depicted in Figure 2:

1. Agree on a very abstract reference model.
2. Obtain a particular model for each dataset to integrate.
3. Reconcile the obtained particular models into a common, shared model.
4. Look at each dataset through the common model.

First of all, an abstract reference model within which the necessary ideas can be expressed needs to be determined. If we assume that the relevant domain of discourse is that of the archaeological record, then this abstract reference model can be any solid ontology of the archaeological record plus any additional related elements that need to be dealt with. CHARM (Incipit, 2013c) or CIDOC CRM (CIDOC, 2011) are examples of adequate abstract reference models for archaeology. This reference model must be very abstract, since it should be applicable to a wide range of situations, projects, periods, geographical locations and cultures; the more abstract it is, the wider the range of datasets we will be able to integrate under it.

Secondly, a particular model for each dataset must be obtained. A particular model is a conceptual model that is derived from the agreed-upon abstract reference model to describe the information in the dataset at the most specific level of detail that is attainable. Particular models may be pre-existing if the dataset was constructed with such foresight in mind, or they may need to be developed for the occasion if they don't exist.

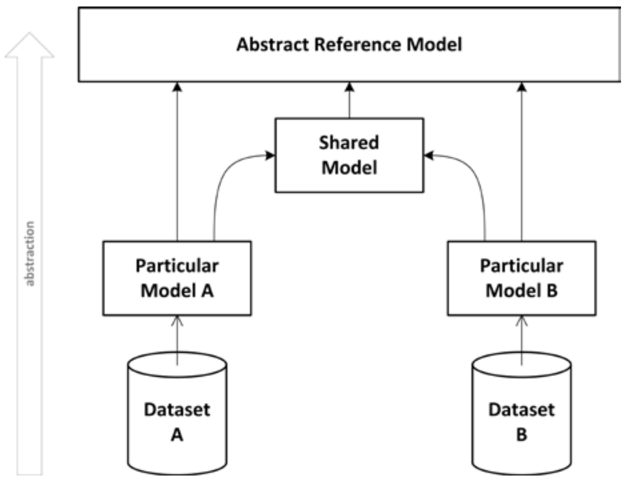


Figure 2: Overall architecture of the proposed solution. Datasets are depicted as cylinders and models as rectangles. Stick arrowheads depict data conformance to a model, and filled in arrowheads depict conformance between models. Abstraction increases from bottom to top.

Then, the different particular models obtained (one for each dataset) must be reconciled into a common, *shared model*. This is achieved through an abstraction process that removes the specific details of each particular model and leaves only those details that are common to all of them. Thus, the resulting shared model will be necessarily more abstract than any of the individual particular models, while still being conformant to the abstract reference model from step 1. Sometimes it is the case that the particular models are so different that no shared details can be found other than those given by the abstract reference model; in this case, the shared model is equivalent to the abstract reference model.

Finally, the information in any of the considered datasets can be accessed through the reconciled shared model. Obviously, this would result in an abstracted view of the data where some detail is lost for the sake of interoperability. In particular, any information accessed through the shared model is guaranteed to be structurally and semantically interoperable regardless of what dataset it resides in.

A few comments are necessary to complement this description. First of all, the notion of *conformance* must be defined. A dataset is said to be conformant to a model if all the entities stored in it are (direct or indirect) instances of the types declared in the model. In turn, a model is said to be conformant to another, more abstract, model if all the types declared in it are subtypes of the types declared in the more abstract model. Model conformance involves Liskov substitutability (Liskov and Wing, 1994), by which we can state that if a dataset conforms to a model M, and model M conforms to model N, then the dataset also conforms to model N. In Figure 2, for example, datasets A and B are conformant to the Shared Model and to the Abstract Reference Model as well by virtue of this rule.

A second important comment concerns the notion of *model derivation*. We said that particular models are derived from the abstract reference model; derivation entails the construction of a model from a base one through well-defined mechanisms that allow for the creation of a more specific, less abstract model that is conformant to its base. Model derivation is also called model extension in the literature (Incipit, 2013b), and requires that the base model is expressed in a clear, non-ambiguous language and that the derivation mechanisms are also well documented and semantically unambiguous. For this reason, informal models or models expressed in a language without a formal basis are poor candidates for dataset integration or other kinds of semantic operation.

As third and last comment, it is important to bear in mind that shared models can be created at multiple levels of abstraction. For example, in a scenario where multiple datasets are being integrated, and once a particular model for each one has been obtained, particular models that are especially similar to each other can be identified and a 'local' shared model developed. Then, a 'global' shared model that covers all the 'local' shared models can be created. This way, abstraction is managed in a two-step manner, making differences between datasets more manageable and providing for conceptual 'anchor points' that are stable and can work as future references.

The following section illustrates the proposed approach through an example where three datasets were integrated in order to perform queries across the combined information in a transparent manner

4. Case Study

The chosen example involves the integration of three datasets. The first dataset, named SIA, comprises a subset of the SIA+, the 20-year old, half-million object database that Incipit CSIC uses as their regular archaeological database (Gonzalez-Perez, 1999). The second dataset, named Caminho, is a subset of the 'Caminho Primitivo' ('Primitive Way') database, a purpose-built data store also created by staff at Incipit CSIC for the cataloguing of sites and associated ethnographic elements along the Way of St. James in Spain (Martins da Fonte, 2009). The third dataset, named Essex, is a subset of the 'Town and Country in Roman Essex' database as offered by the Archaeological Data Service (ADS) of the University of York, a Roman settlements database from Essex in the UK (Perring, 2011).

These datasets were selected under the hypothesis that SIA and Caminho, having been developed within the same organisation, would likely share many of the involved conceptualisations, while Essex, being completely alien to Incipit CSIC, would not share a significant amount of specific semantics with them. Under this hypothesis, a shared model that would capture Incipit CSIC's overall conceptualisations and integrates SIA and Caminho under a common view was planned, as was a second, more abstract shared model that would cover the three datasets. This hypothetical architecture is shown in Figure 3.

Also, the technological choice to use CHARM (Cultural Heritage Abstract Reference Model) (Incipit, 2013c) as an abstract reference model was made. CHARM is an 'ontology' of cultural heritage, i.e. it represents 'anything that may be the recipient of cultural value ascribed by any individual, plus the associated valorizations ascribed to said things, plus the representations of these things that may exist. In this way, CHARM does not only represent the specific entities that might receive cultural value, but also other entities which, without doing so, are necessary in order to describe and understand the former' (Incipit, 2013d). CHARM was chosen over e.g. CIDOC CRM because of two major reasons. First of all, CHARM is expressed in a well-known, clearly documented modelling language, namely ConML (Incipit, 2013a), so that semantic ambiguity is minimized. Secondly, guidelines for extending the model have been documented and made public (Incipit, 2013b), providing specific advice for users to create particular models that are guaranteed to be Liskov-compatible with CHARM and therefore interoperable with one another. In other words, CHARM was conceived from the viewpoint of gradual refinement and conceptual modelling, rather than under the premise of data-oriented direct mapping.

The fact that CHARM is expressed in ConML means that the figures that illustrate most of this paper take the form of ConML type diagrams. Type diagrams represent concepts through classes, hierarchical subsumption of concepts through generalisation relationships, properties of concepts through attributes, and relationships between concepts through associations. Classes are usually depicted as rectangles containing the class' attributes (see e.g. class 'Photograh' on the bottom left of Figure 4); generalisation relationships are depicted as triangular arrows pointing from the more specific to the more abstract class (see e.g. the arrow going from 'PotteryFind' to 'Find' in Figure 4); and associations are shown as lines linking two class rectangles together (see e.g. 'OccursTo' between

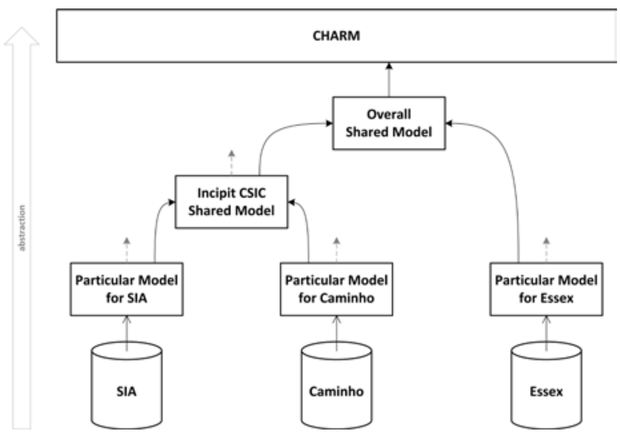


Figure 3: Hypothetical architecture for the case study. Like in Figure 2, datasets are depicted as cylinders, models are depicted as rectangles, stick arrowheads depict data conformance to a model, and filled in arrowheads depict conformance between models. Indirect conformance between models (i.e. that which is a result of Liskov substitutability) is shown as grey dashed arrows for clarity.

'CreationEvent' and 'PrimaryEntity' at the top of Figure 4), often labelled with a name. For a complete description of ConML and its graphical notation, please see (Incipit, 2013a).

The following sections describe how the three datasets were integrated under this approach, and how the outcomes of this integration were used.

4.1. Integrating Datasets

Following the steps outlined at the beginning of Section 3, and once CHARM had been selected as abstract reference model, a particular model for each dataset to integrate was developed. The creation of a conceptual model from a dataset is far from a trivial process, and depending on the case may require a substantial effort in terms of study and abstraction of the source dataset. In the case of SIA, one of the authors of this paper was the chief developer of the source dataset, so the task was relatively easy. For Caminho, the dataset creators were interviewed and information gathered this way. In the case of Essex, the documentation that accompanies the database was studied and a likely conceptual model was derived from both the documentation and the database structure. Due to these varying approaches, the degree of reliability of the resulting particular models decreases in this order: SIA, Caminho, Essex. The fact that the documentation of the Essex dataset as offered to the public through ADS does not describe its abstract semantics but only its database-level implementation details made the task more difficult than expected, and was found particularly noteworthy.

4.1.1 Obtaining Particular Models

Figure 4 shows the particular model for the SIA dataset. Classes have been introduced to represent features, pottery finds and lithic finds. Also, photographs are considered. Areas and point of interest were also modelled and a relationship added to capture the fact that an area of interest is defined through a collection of points. Archaeological impact and conservation status interpretive classes were also added according to the methodology employed by SIA.

Figure 5 shows the particular model for the Caminho dataset. Classes have been introduced to represent architectural elements such as real estate elements and villages. A separate class is used to describe element locations. Intangible elements are also considered since they played a prominent part in the Caminho project. Heritage impact and conservation status interpretive classes were also added in a very similar fashion to those of SIA.

Finally, Figure 6 shows the particular model for the Essex dataset. Classes have been introduced to represent two kinds of finds: pottery finds and coins. A separate class is used to describe locations, and another class has been introduced to represent events associated to finds.

4.1.2 Developing a Shared Model

Once the particular models were obtained, it was easy to observe that, in fact, the SIA and Caminho models shared a number of common features. For example, both include a 'ConservationStatus' class with nearly identical

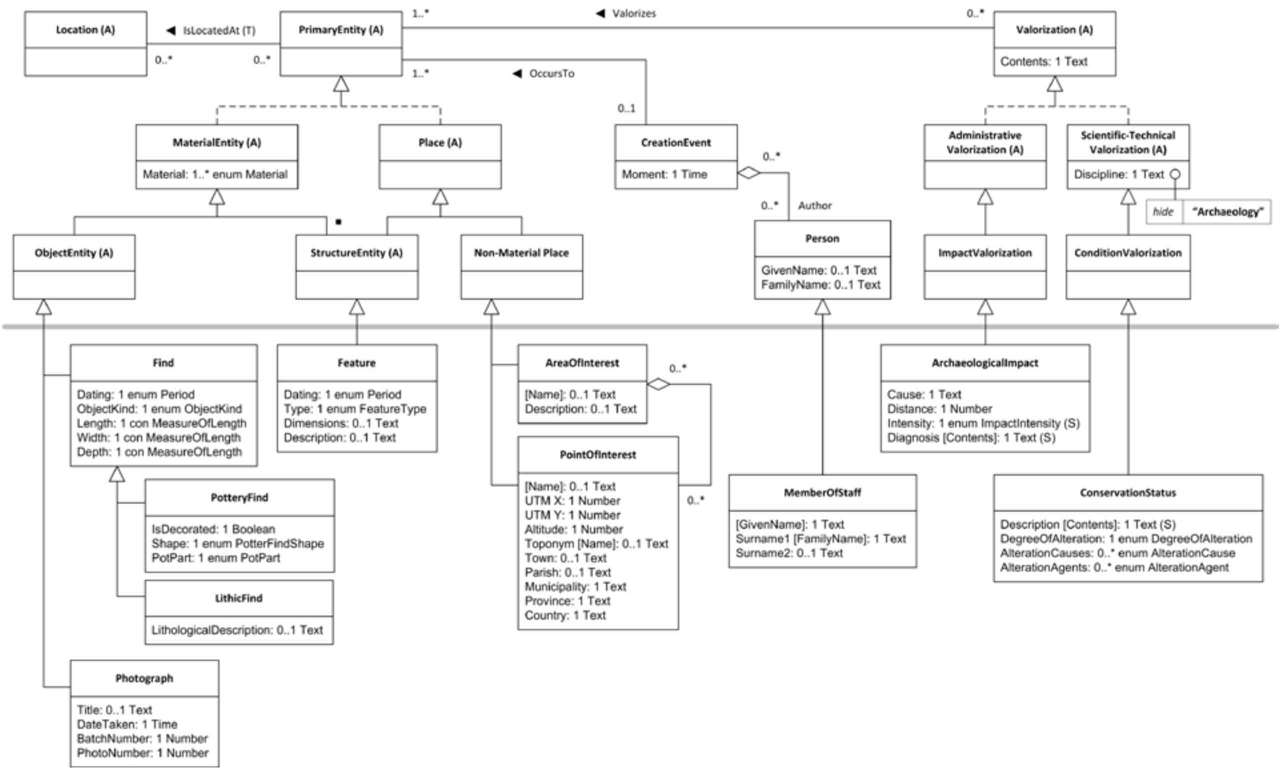


Figure 4: Particular model for the SIA dataset. Classes above the grey line are part of CHARM; classes below the grey line have been added to capture concepts specific to SIA.

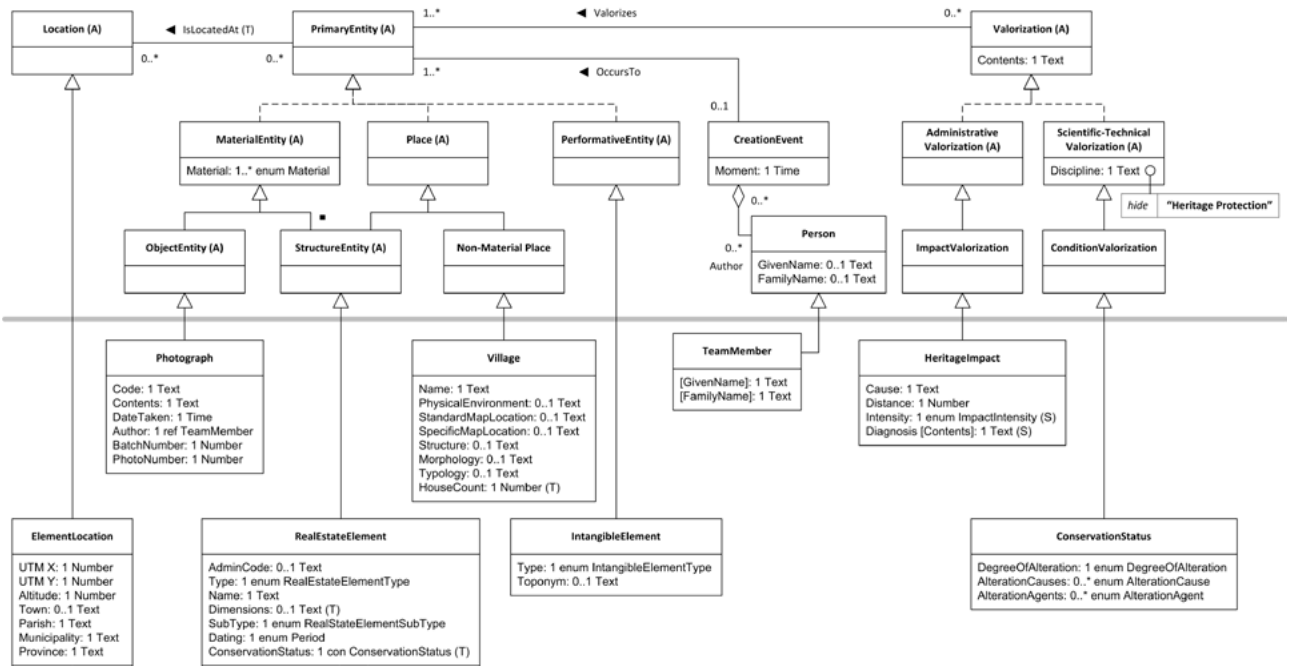


Figure 5: Particular model for the Caminho dataset. Classes above the grey line are part of CHARM; classes below the grey line have been added to capture concepts specific to Caminho.

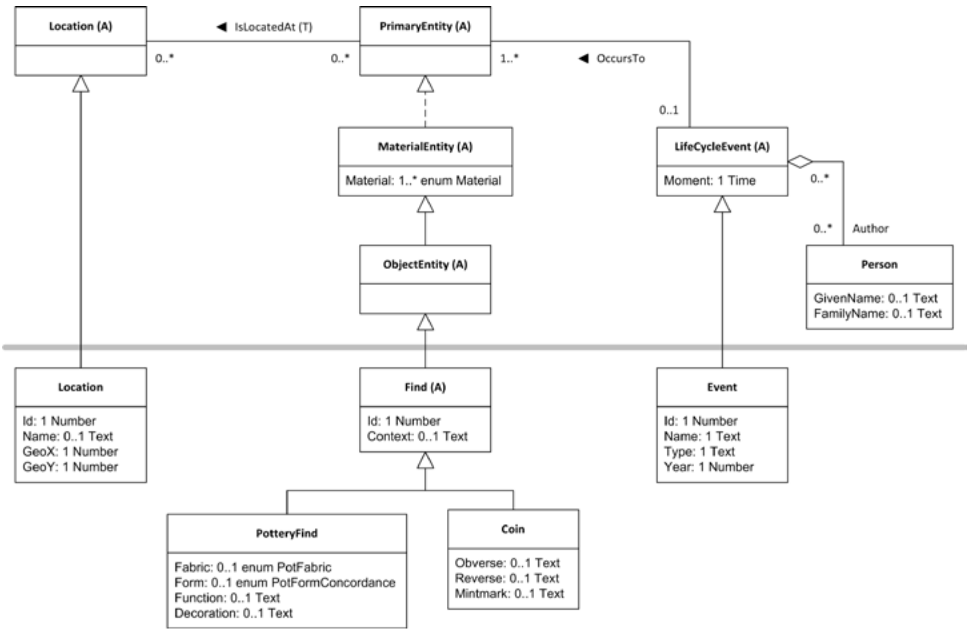


Figure 6: Particular model for the Essex dataset. Classes above the grey line are part of CHARM; classes below the grey line have been added to capture concepts specific to Essex.

attributes; the classes 'HeritageImpact' of Caminho and 'ArchaeologicalImpact' of SIA, despite having different names, have identical attribute structure; both models contain a 'Photograph' class, albeit with slightly different attributes; etc. According to plan, these commonalities were factored out into an Incipit CSIC shared model (see Figure 3 for its place in the hypothetical integration architecture). This model is shown in Figure 7.

Creating the shared model is not enough; reconciling the particular models with it is necessary in order to exploit the benefits of the shared model. This 'reconciliation' involves re-expressing the relevant particular models in terms of the shared one. In our case, the SIA and Caminho particular models (shown in Figure 4 and Figure 5) were modified so that they would be based on the Incipit CSIC shared model (Figure 7) rather than on CHARM directly. Reconciliation is usually an easy task, and often involves the refactoring of some classes and associations while fully maintaining semantics.

Figure 8 shows the particular model for the SIA dataset expressed in terms of the Incipit CSIC shared model. Note that a common class such as 'IntangibleElement' is removed from the model since it is not needed for the SIA dataset. The 'Impact' common class, in turn, is simply

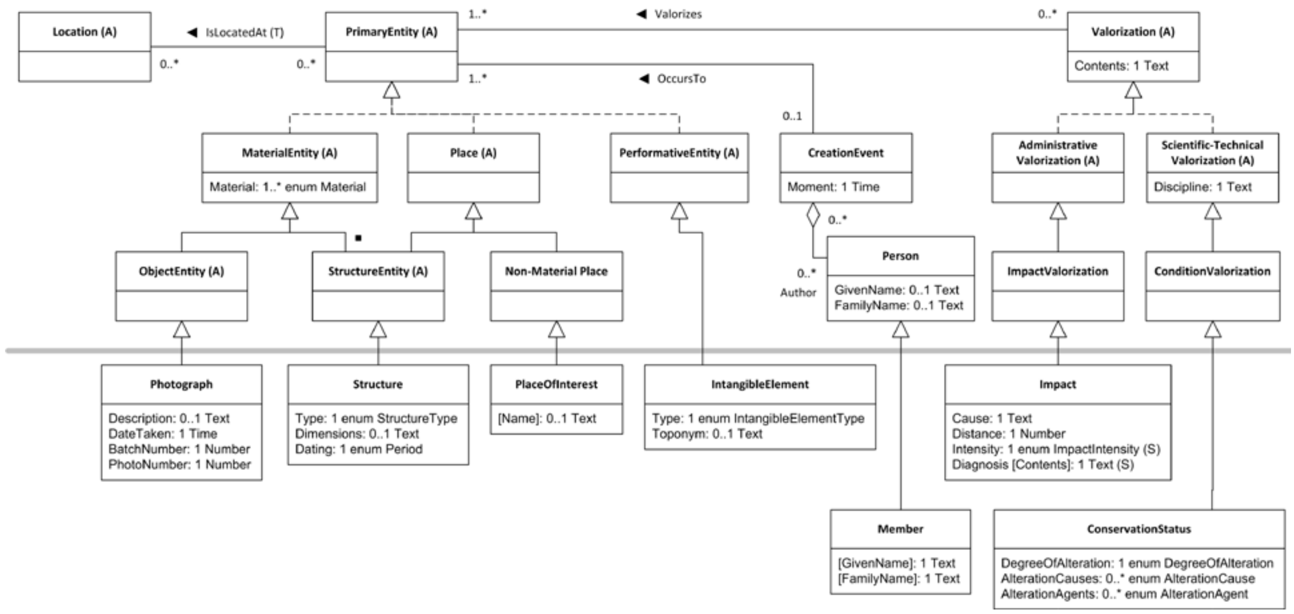


Figure 7: Incipit CSIC shared model. Classes above the grey line are part of CHARM; classes below the grey line have been added to capture concepts specific to Incipit CSIC as determined by commonalities to both SIA and Caminho.

renamed to 'ArchaeologicalImpact'. Other classes such as 'ConservationStatus' are refined in order to add extra attributes that are specific to SIA, and even some others such as 'Find' are newly added.

Similarly, Figure 9 shows the particular model of Caminho expressed in terms of the Incipit CSIC shared model. In this case, 'IntangibleElement' is not removed but left intact, since it is necessary for the Caminho dataset. The 'Impact' common class, in turn, is renamed to 'HeritageImpact' but otherwise left untouched, and some classes such as 'Village' are added.

This way, a 3-tier model structure was created, following the architecture sketched in Figure 3:

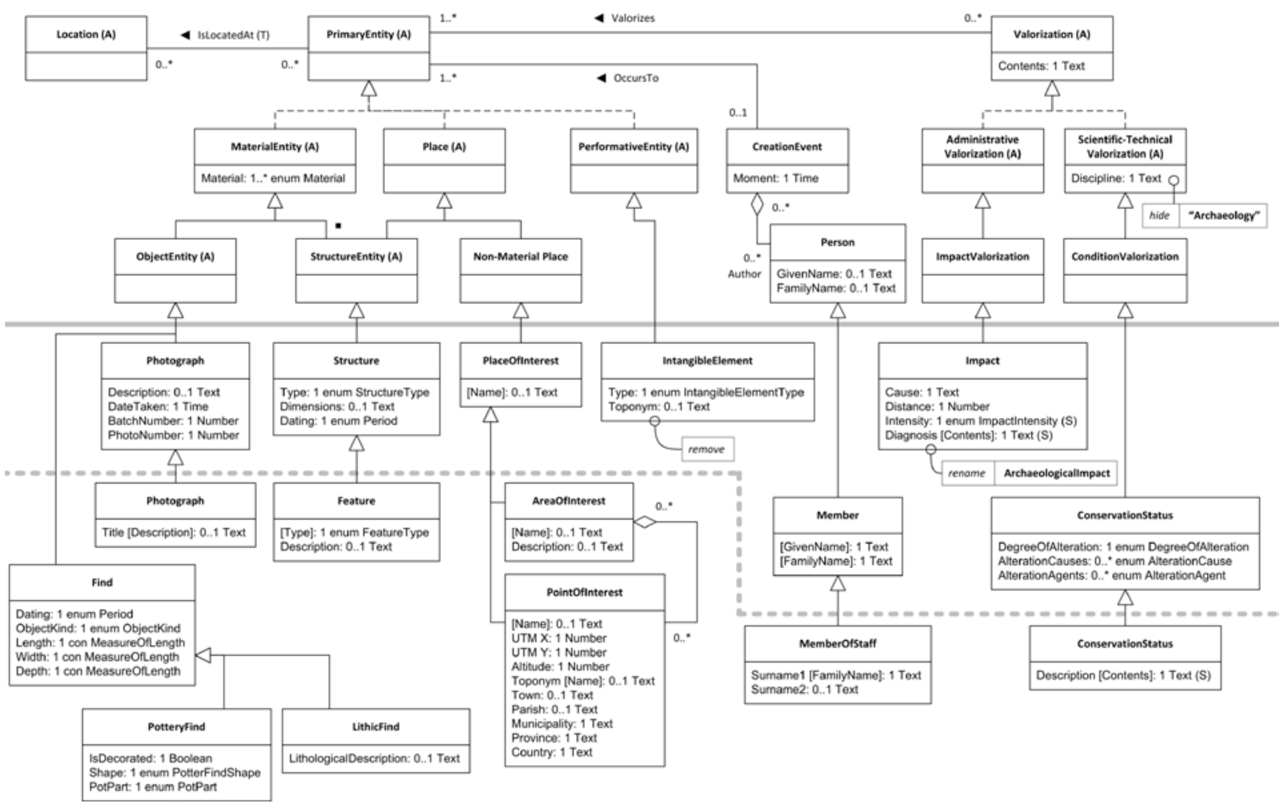


Figure 8: Incipit Particular model for SIA expressed in terms of the Incipit CSIC shared model. Classes above the solid grey line are part of CHARM. Classes below the solid grey line and above the dotted grey line have been added to capture concepts specific to Incipit CSIC as determined by commonalities to both SIA and Caminho (see Figure 7). Classes below the dotted grey line have been added to capture concepts specific to SIA.

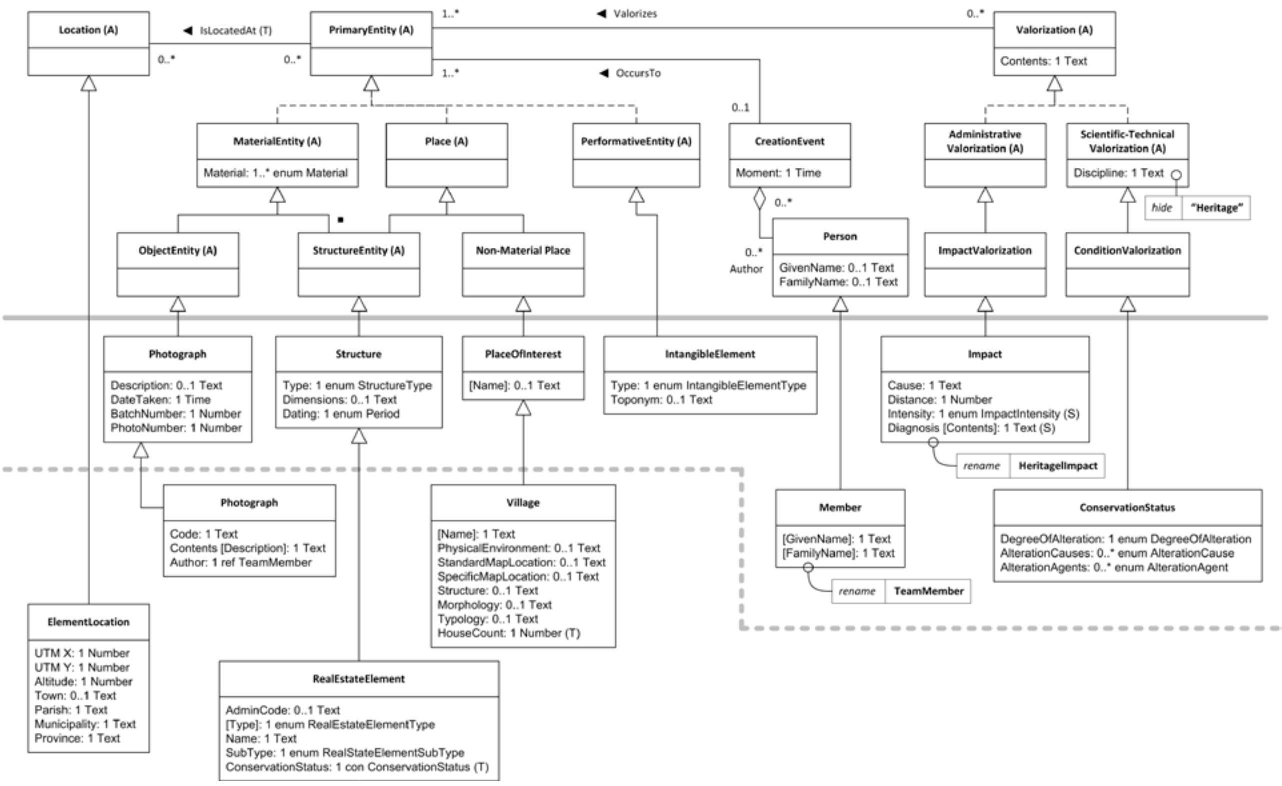


Figure 9: Particular model for Caminho expressed in terms of the Incipit CSIC shared model. Classes above the solid grey line are part of CHARM. Classes below the solid grey line and above the dotted grey line have been added to capture concepts specific to Incipit CSIC as determined by commonalities to both SIA and Caminho (see Figure 7). Classes below the dotted grey line have been added to capture concepts specific to Caminho.

- The particular models, even in their reconciled versions, describe the information in each of the datasets (SIA, Caminho) at the most precise level of detail. Abstraction is minimal: every piece of data is represented in these models. However, each of these models is tightly bounded to its dataset, and cannot describe any other dataset that has a different structure. This means that interoperation cannot occur.
- The Incipit CSIC shared model, having been obtained as an abstraction of SIA and Caminho, describes the information in both of these datasets, albeit at the price of having dropped some details. For example, its 'Photograph' class takes into account details such as the date the picture was taken and the photo number, but ignores particularities such as SIA's photo title or Caminho's photo code; still, most relevant details are preserved. At this level, abstraction is medium, and this shared model can successfully describe all the datasets that 'fall under it', but only if some details are omitted. Interoperation can occur within this range of datasets.
- The CHARM abstract reference model describes the information in potentially any dataset, but a significant amount of detail needs to be dropped in order to bridge the gap between the very abstract CHARM and the very specific dataset. CHARM's abstraction is maximal but, in exchange, its coverage is extremely wide. Interoperation can occur across any dataset, but at the high price of dropping much detail.

Thus, we argue that the more abstract a model is, the larger the number of datasets that can be made interoperable through it, and the more different these datasets can be. The value in having a multi-tier model structure (3 tiers in our example, but more are perfectly possible) resides in the fact that users can choose at which point in this structure they wish to work at any point in time, and thus select an acceptable compromise between information detail and interoperation range.

4.1.3 Integrating the Third Dataset

As for the Essex dataset, and according to the hypothetical architecture of Figure 3, its particular model and the Incipit CSIC shared model should be abstracted out into an even more abstract overall shared model. However, and as we explained in Section 3, sometimes it is the case that two models are so different that no significant shared details are found other than those given by the abstract reference model; this was the case of Essex and the Incipit CSIC shared model. After exploring some abstraction options, it was decided that a common model would be small that its existence was superfluous. In cases like this, CHARM is better used as a shared model. It is worth mentioning that the fact that classes named 'Find' exist both in the SIA and Essex particular models does not necessarily entail a strong semantic correlation between these models. A close examination of the models reveals that attributes and associations of the 'Find' classes are clearly different, and for this reason it was decided that finds in SIA and finds in Essex would rather be integrated through the CHARM abstract concept of 'ObjectEntity'. Of course, a different strategy could have been used, and an alternative solution is described in Section 5.

4.2. Using the Integrated Datasets

As we describe at the end of the previous section, the power of a multi-tier model structure lies in the fact that users can select an appropriate compromise between information detail and interoperation range when they work with the datasets. We will illustrate this through the use of data queries against the three datasets considered in this paper. Let's imagine that we are interested in the information about rock carvings stored in the different datasets. First, queries are issued individually against the SIA and Caminho datasets using the matching particular model for maximum detail (but minimum integration coverage). Then, a similar query is issued again both datasets using the common Incipit CSIC shared model; this allows for a consolidated result set but limits the details obtained. Finally, a query is issued against the three datasets using CHARM; this allows for a wide overview of the information but severely constrains the data that is returned.

4.2.1 Querying SIA

According to the SIA particular model (Figure 8), and assuming that rock carvings are conceptualised under the class 'Feature' with Type = RockCarving, this query would return something like the following (table 1).

4.2.2 Querying Caminho

According to the Caminho particular model (Figure 9), and assuming that rock carvings are conceptualised under the class 'RealEstateElement' with Type = RockCarving, this query would return something like the following (table 2). Please note that the values of 'ConservationStatus' are omitted for the sake of brevity, since they involve references to nested objects rather than simple scalar values.

4.2.3 Querying the Integrated SIA and Caminho

According to the SIA particular model (Figure 8) and Caminho particular model (Figure 9), it is the class 'Structure' which abstracts out the SIA's 'Feature' and Caminho's 'RealEstateElement'.This means that rock carvings are abstractly conceptualised under this class 'Structure' as far as the Incipit CSIC shared model is concerned. Also, given the fact that the Type attribute is of an enumerated data type, it can be safely assumed (from the ConML semantics) that the 'FeatureType' enumerated type in SIA and the 'RealEstateElementType' enumerated type in Caminho both specialise from the 'StructureType' enumerated type in the Incipit CSIC shared model. The query can be thus issued against this model and having Type = RockCarving, and would return something like the following (table 3).

Material	Type	Dimensions	Dating	Description
Rock/Granite	RockCarving	3.0 x 1.5 m.	Neolithic	Two maze shapes plus three concentric circles. Very eroded.
Rock	RockCarving	3.5 x 4.8 m.	(unknown)	Four oval shapes and five crosses.
Rock/Granite	RockCarving	1.1 x 0.6 m.	Neolithic	Two cups.
Rock/Granite	RockCarving	6.0 x 2.1 m.	BronzeAge	One deer shape, a large spiral and three crosses.

TABLE 1: QUERY RETURN CLASS 'FEATURE' WITH TYPE = ROCKCARVING.

Material	Type	Dimensions	Dating	AdminCode	Name	SubType	ConservationStatus
Rock/ Granite	RockCarving	5.6 x 3.1 m.	Neolithic	YA961029Z02	Pena Cabaleira	Archaeo	...
Rock/ Granite	RockCarving	5.5 x 2.5 m.	Neolithic	YA961029Z01	Pena do Zapato	Archaeo	...

TABLE 2: QUERY RETURN CLASS 'REALESTATEELEMENT'.

Material	Type	Dimensions	Dating
Rock/Granite	RockCarving	3.0 x 1.5 m.	Neolithic
Rock	RockCarving	3.5 x 4.8 m.	(unknown)
Rock/Granite	RockCarving	1.1 x 0.6 m.	Neolithic
Rock/Granite	RockCarving	6.0 x 2.1 m.	BronzeAge
Rock/Granite	RockCarving	5.6 x 3.1 m.	Neolithic
Rock/Granite	RockCarving	5.5 x 2.5 m.	Neolithic

TABLE 3: QUERY RETURN.

This result set clearly shows that some of the details that were present for SIA and Caminho have disappeared; specifically, the 'Description' column of SIA is not available, and the 'AdminCode', 'Name', 'SubType' and 'ConservationStatus' columns used in Caminho are not available either. However, the result set successfully consolidates entities from two datasets, since the common model is capable of describing both at this level of abstraction. Operations such as counts or clustering over present attributes such as 'Material' or 'Dating' are possible in a consolidated fashion

4.2.4 Querying at a High Abstraction Level

Finally, a high abstraction level query can be issued through CHARM for maximum integration coverage, which would include the Essex dataset as well. As shown in Figure 7, the CHARM class that is the closest to 'Structure' in the Incipit CSIC shared model is 'StructureEntity'. This class, however, does not exhibit any manner to filter by type (since CHARM is highly abstract), so there is no way that rock carvings can be separated from other structures (such as pits or walls) at such a high level of abstraction. This, however, should not be surprising.

In addition, the Essex particular model (Figure 6) does not make use of the CHARM class 'StructureEntity', so we can safely assume that no structure entities are stored in the Essex dataset. However, both the Essex particular model and the Incipit CSIC shared model converge at the 'MaterialEntity' class, which is a generalized class of 'StructureEntity' used by the Incipit CSIC shared model and, at the same time, of the 'ObjectEntity' class used by Essex to implement finds. 'MaterialEntity' is an extremely abstract concept, encompassing every entity that is made of matter and is fundamentally perceived in a direct fashion and through its materiality (Incipit, 2013c); it is arguable that such an abstract concept has practical use for information integration, but the approach here outlined works equally on concepts like this.

Material
Rock/Granite
Rock
Rock/Quartzite
Rock/Granite
Rock/Quartzite; Wood
Rock/Granite
Rock/Granite
Rock/Granite
Metal/Copper
Metal/Copper
Metal/Silver

Table 4: Query return 'MaterialEntity'.

A query on the three datasets and through 'MaterialEntity' would return something like the following (table 4).

At this level of abstraction, and as according to 'MaterialEntity', only the 'Material' attribute is available. Still, this may be useful to obtain counts and other simple statistics about objects such as grouping by 'Material'.

5. Discussion

The approach presented here has some important advantages over conventional approaches, as introduced in Section 2, as well as some limitations that are worth mentioning. With regard to advantages, the fact that knowledge, rather than data, becomes the focus of integration is a crucial benefit of the proposed approach, which arguably improves meaningfulness to archaeologists. As illustrated by the example in Section 4, the fact that shared models exist at intermediary levels of abstraction also provides an excellent anchor point for further reasoning about a collection of datasets in contrast with others; for example, the Incipit CSIC shared model (Figure 7) reveals how structures are conceptualized and described across Incipit CSIC, and makes further integration with additional datasets from the same source a much easier task.

In addition and as previously pointed out, shared models do not need to be limited to one level between particular models and CHARM; in fact, any number of tiers is possible. For example, a large organisation such as a heritage government agency or an archaeological unit may develop an organisation-wide model to describe the overall adopted approach to dealing with archaeology, and then each department or division may refine this model into a more specific one that addresses their needs, such as one for urban excavation or one for underwater archaeology. Furthermore, each non-trivial project may refine one of these into a project-specific model that takes into account the particular needs, methodological choices, stakeholder requirements and even project team preferences regarding the description of the archaeological record. This way, a fine-grained hierarchy of models is constructed where the best possible balance between abstraction and interoperability can be selected by data users.

Regarding limitations, we can mention the following. The proposed approach is strongly based on conceptual modelling, and therefore assumes that valid conceptual models of the datasets can be constructed; it also assumes that these models can be successfully abstracted out into shared models. Conceptual modelling is not a trivial endeavour, and often requires skilled training (Eide *et al.*, 2008). However, our experience with CHARM and ConML shows that this training can be obtained quite quickly and with a great rate of success (Gonzalez-Perez, 2012). Another potential limitation is the need to create conceptual models for the datasets before any tangible result can be obtained, with the associated need for time and effort. In fact, when teaching this approach, we sometimes receive the complaint that it looks too complex, and that a 'direct mapping' method that avoids models should be faster. Evidently, it is true that the suggested approach needs more work than any 'direct mapping' one; however, this extra cost needs to be compared against the benefits that the approach entails, which have been already described.

A very similar approach, also based on the gradual refinement of models, has been adopted by the International Organization for Standardization (ISO) for the subcommittee 7 (SC7) of the Joint Technical Committee 1 (JTC1) special working group on ontologies. This special working group aims to develop an overarching ontology that is capable of describing the whole domain of activity of SC7 (namely, software and systems engineering), in a similar fashion as CHARM is capable of abstractly describing the whole domain of cultural heritage. An architecture has been developed (Henderson-Sellers *et al.*, 2014) and a proof of concept is being created. Although the application areas of the SC7 special working group (software and systems engineering) and the proposal described here (archaeological data integration) are extremely different, the modelling theory behind the approach is the same, and we argue that both experiences will be able to feed back into each other.

Finally, it is worth mentioning that conceptual modelling is not a single-solution technology: multiple valid models can be obtained that reliably represent a given portion of reality or a dataset. This means that the suggested approach contains a significant amount of inherent subjectivity. This is not a drawback, since it is through this subjectivity that model creators can inject their expertise, background and tacit knowledge into the models that they develop. For instance, the example described in Section 4 treats the 'Find' class in the Essex particular model as a different concept to the 'Find' class in the SIA dataset. As we explained in Section 4.1.3, this is one valid solution, backed by the very different attributes and associations that these classes exhibit in both models. However, alternative solutions can also be considered. For example, a finds specialist may have chosen to integrate first the Essex and SIA datasets based on the fact that both deal with finds in a very central manner, whereas Caminho does not, by creating an Essex+SIA shared model; and only then, as a second step, integrate the Essex+SIA shared model with the Caminho particular model. The resulting features of the integration would naturally be different to those of the described example, responding to the purpose and preferences of the alternative situation.

6. Conclusions

In this paper we have proposed a knowledge-level integration approach for archaeological datasets, which departs from the usual data-oriented ones by employing gradual refinement of conceptual models. The CHARM abstract reference model is suggested as a conceptual framework, and ConML proposed as an infrastructural modelling language.

The proposed solution, despite involving more work that 'direct mapping' ones, yields richer archaeological semantics, and allows information users to select the optimal balance between interoperation and information detail at any point in time.

Acknowledgement

This work has been partially funded by the Spanish National Research Council (CSIC) through the program 'Junta de Ampliación de Estudios' (B.O.E. 146, 20/06/2011) and the FP7 INFRASTRUCTURES project 'ARIADNE' (grant number 313193).

References

ARIADNE (2014). ARIADNE (web site). Accessed on 16 July 2014. Available from: http://www.ariadne-infrastructure.eu/

CIDOC (2011). The CIDOC Conceptual Reference Model (web site). Accessed on 26 November 2012. Available from: http://www.cidoc-crm.org/

ECKKRAMMER, F., FELDBACHER, R. & ECKKRAMMER, T. (2008). CIDOC CRM in Data Management and Data Sharing. Data Sharing between Different Databases. In Jerem, E., Redő, F. & Szeverényi, V. (Eds.) *On the Road to Reconstructing the Past: Computer Applications and Quantitative Methods in Archaeology (CAA): Proceedings of the 36th International Conference: Budapest, April 2–6, 2008*. Budapest, Archaeolingua Foundation. pp. 153-158.

EIDE, O., FELICETTI, A., ORE, C. E., D'ANDREA, A. & HOLMEN, J. (2008). Encoding cultural heritage information for the semantic web. Procedures for data integration through CIDOC-CRM mapping. In Arnold, D., Niccolucci, F., Pletinckx, D. & Gool, L. V. (Eds.) *Proceedings of the EPOCH Conference on Open Digital Cultural Heritage Systems*.1-7.

Europeana, (2014). Europeana Cloud (web site). Accessed on 16 July 2014. Available from: http://pro.europeana.eu/web/europeana-cloud

GONZALEZ-PEREZ, C. (1999) *Tecnologías de la Información y Patrimonio Cultural: Una Metodología Integral Orientada a Objetos para Desarrollo de Software*, Santiago de Compostela, Spain, Laboratorio de Arqueología y Formas Culturales, University of Santiago de Compostela.

GONZALEZ-PEREZ, C. (2012). A Conceptual Modelling Language for the Humanities and Social Sciences. IN Rolland, C., Castro, J. & Pastor, O. (Eds.) *Sixth International Conference on Research Challenges in Information Science (RCIS 2012)*. Valencia (Spain), IEEE Computer Society.

HENDERSON-SELLERS, B., GONZALEZ-PEREZ, C., MCBRIDE, T. & LOW, G. (2014). An ontology for ISO software engineering standards: 1) Creating the infrastructure. *Computer Standards & Interfaces*, 36, 563-576.

Incipit, (2013a). *ConML Technical Specification*. ConML 1.4. Available from: http://www.conml.org/Resources_TechSpec.aspx

Incipit, (2013b). *CHARM Extension Guidelines*. Incipit, CSIC. Available from: http://www.charminfo.org/Resources/Technical.aspx

Incipit, (2013c). *CHARM Web Site* (web site). Accessed on 30 October 2013. Available from: http://www.charminfo.org

Incipit, 2013d. *CHARM White Paper*. Incipit, CSIC. Available from: http://www.charminfo.org/Resources/Technical.aspx

ISAKSEN, L., MARTINEZ, K., GIBBINS, N., EARL, G. & KEAY, S. (2010) Linking Archaeological Data. In Frischer, B., Webb Crawford, J. & Koller, D. (Eds.) *Proceedings of Computer Applicationsand Quantitative Methods in Archaeology* (CAA 2009), Williamsburg. Oxford (UK), Archaeopress

ISO, 2006. *Information and documentation -- A reference ontology for the interchange of cultural heritage information*. ISO 21127:2006.

KINTIGH, K. (2006) The Promise and Challenge of Archaeological Data Integration. *American Antiquity*, 71, pp. 567-578.

LISKOV, B. & WING, J. M. (1994) A Behavioral Notion of Subtyping. *ACM Transactions on Programming Languages and Systems*, 16, pp. 1811-1841.

MARTÍN-RODILLA, P. (2013). Software-Assisted Knowledge Generation in the Archaeological Domain: A Conceptual Framework. In Indulska, M. & Weber, B. (Eds.) *Proceedings of the Doctoral Consortium of the 25th International Conference on Advanced Information Systems Engineering (CAiSE 2013)*, Valencia, Spain, June 21, 2013.CEUR-WS.org

MARTINS DA FONTE, J. M., (2009). Aplicações dos Sistemas de Informação Geográfica (SIG) à Gestão dos Recursos Patrimoniais: O Caminho Primitivo de Santiago, in Faculdade de Letras. Master's Thesis. Universidade de Porto. Porto.

MAY, K., BINDING, C., TUDHOPE, D. & JEFFREY, S. (2012) Semantic Technologies Enhancing Links and Linked Data for Archaeological Resources. In Verhagen, P. (Ed.) Revive the past: *Proceedings of the 39th Annual Conference of Computer Applications and Quantitative Methods in Archaeology (CAA),* Beijing, China, 12-16 April 2011. Amsterdam, Amsterdam University Press, pp. 261-272.

MEYER, E., GRUSSENMEYER, P., PERRIN, J.-P., DURAND, A. & DRAP, P. (2007). Integration of Heterogeneous Cultural Heritage Data in a Web-based Information System: A Case Study from Vianden Castle, Luxembourg. In Clark, D. & Hagemeister, E. M. (Eds.) *Digital Discovery: Exploring New Frontiers in Human Heritage: CAA 2006: Proceedings of the 34th Conference, Fargo*. David Brown Book Company. pp. 329-340.

PERRING, D. (2011). Town and Country in Roman Essex: Settlement Hierarchies in Roman Essex. BY Archaeology South-East, through Archaeology Data Service. Available from: http://archaeologydataservice.ac.uk/archives/view/romanessex_eh_2011/

Linked Open Greek Pottery

Ethan GRUBER,[1] Tyler Jo SMITH [2]

[1] American Numismatic Society
[2] McIntire Department of Art, University of Virginia, USA
gruber@numismatics.org, tjs6e@virginia.edu

Abstract
Linked open data methodologies have tremendous potential in facilitating data interoperability, aggregation, and analysis. With lessons learned in the development of Nomisma.org, a collaborative project dedicated to the definition of numismatic concepts in the semantic web, the authors have endeavored to apply these methodologies to the field of Greek pottery, through Kerameikos.org. While several thesauri that contain pottery concepts have been published online, not all employ linked data, and none of them are ―five star‖ linked data, as defined by Berners-Lee. This paper discusses the development of a discipline-specific thesaurus which serves as a bridge between existing vocabulary systems, the open source XForms/REST/SPARQL framework for its publication, and the development of web-based tools to analyse and visualise pottery data aggregated from the Getty Museum and British Museum as a proof of concept of the utility of Kerameikos.org as a research tool for Greek pottery.

Keywords: linked data, pottery, xforms, information architecture

1. Accessing Greek Vases: A Potted History

The study of figure-decorated pottery (or 'vases') from archaic and classical Greece has undergone tremendous change over the past century. Open and closed vessels, varying greatly in their form, scale, function, subject, and quality of embellishment, not only line the shelves of many Classical art collections in the major museums of Europe and the United States, but also continue to be discovered in abundant quantities from excavated contexts. A cursory glance at publications devoted to the locally produced vases excavated from the Athenian Agora, such as those in Volume 23 for black-figure (Moore and Philippides, 1986), offers a sense of the range of shapes and styles, as well as their standard method of presentation as archaeological material - i.e. the data is presented in catalogue form according to shape categories rather than according to specific context, findspot, or assemblage. On closer inspection one would notice that, in addition to the expected information about condition and size, the individual pots, cups, and fragments have been associated with or attributed to specific painters or workshops. The depth of detail and the number of artistic personalities to choose from would seem staggering to archaeologists unfamiliar with Classical (and esp. Greek) archaeology and its unique system of pottery classification.

Greek vases were produced in great quantities in the potter's quarters of ancient Corinth and Athens, as well as in other regions of Greece, among them Laconia, Boeotia, the Cyclades, Crete, and various cities in the East (Boardman, 2008; Cook, 1997). Although a large number of vases have been discovered in Etruscan tombs, and their Etruscan reception and context has begun to interest many scholars, it is important to recall that vast amounts of figure-decorated pottery have been uncovered in both Greece itself and from Greek colonial sites across the Mediterranean and around the Black Sea. The mechanisms of production for these vases, regardless of place of manufacture or final spot of deposition, are not fully understood. Ancient authors, who have a great deal to tell us about sculptors, architects, and painters, are virtually silent on the topic pottery, which may indicate something significant about its standing as a craft as opposed to _high' art. That being said, many would argue that Greek vases, especially those produced in the city of Athens from the middle years of the 6th century onward, have much to offer aesthetically, and that figural vases of all dates, styles, and locations are the best surviving visual source for myths about the gods and heroes, religious life and practices, and many other everyday activities, such as dancing, drinking, and dining. Iconography, namely the imagery (or 'vase-painting') and its interpretation, has also been a dominant trend in the modern analysis of Greek vases.

While the early history of Greek vase scholarship cannot be recounted in detail here, it is useful to remember that their scholarly presentation as archaeological material derives from the combination of their position in the early history of antiquities collecting and their high esteem in that regard, their large survival numbers, and the classification established for Athenian (or 'Attic') examples during the last century by Sir John Beazley of Oxford University (Nørskov, 2002; Rouet, 2001; Kurtz, 1985). At present there are several major resources used by archaeologists to access, identify, and classify Greek vases, and each is applicable regardless of the nature (i.e. whole pots vs sherds) or origin (i.e. Athenian vs non-Athenian) of the evidence they showcase. Volumes, such as one listed above for the Agora in Athens, are among the first places vase specialist consult in order to identify forms and their decoration (both subsidiary ornamentation and human/

animal figures), particularly if one is making a study of unpublished excavation pottery. The large quantities of well-studied figure-decorated pottery from Athens, Corinth, and Samos, among other places – complete with drawings, photographs, and detailed descriptions - can provide either immediate comparanda or at the very least serve as a vital starting point for further exploration. A second go-to place, and one where many would turn for Athenian vases, is the Beazley Archive Pottery Database (BAPD) maintained by the Beazley Archive at Oxford University (www.beazley.ox.ac.uk). The BAPD was based initially on the famous 'lists' of Attic black- and red-figure vase-painters compiled by Sir John Beazley and published in several volumes. Since Beazley's death in 1970, the paper archive comprised of photographs, notes, and drawings, has been digitized and made available as an open access searchable database (Smith, 2005). A large number of entries and accompanying photographs have been added since the death of Beazley, and the Beazley Archive was able to provide the following statistics as of 21 July 2014:

- Greek vases all fabrics: 107,610
- Greek vases with images: 60,289
- All Athenian vases: 80,183
- Athenian black-figure: 42,664
- Athenian red-figure: 37,519
- All vases in Beazley's lists: 34,000

Although on average the BAPD adds 5000 entries per year based on newly published material, it is most readily accessible to specialists and assumes a certain level of technical expertise. For non-Athenian vases, other scholars, such as A.D. Trendall for South Italian wares, have applied similar methods of attribution to establish their own classification lists complete with artists' names and a legacy archive (http://www.latrobe.edu.au/trendall/). A third important resource for vase scholars is the Corpus Vasorum Antiquorum [CVA] an ongoing international project founded in 1922 with the aim of publishing in a more or less standardised form all vases from all collections around the world (http://www.cvaonline.org/cva/). Digitisation of the existing CVA volumes, many of which are now out of print, was undertaken by the Beazley Archive beginning in 2000. The contents of the CVA volumes were scanned as PDFs and are now searchable in their own right or as part of a BAPD search for an individual object. As with the excavation reports and the BAPD itself, the CVA volumes are written by specialists for experts in the field of Greek vase scholarship; and despite their production cost, more are being put into print each year. Each of these resources, furthermore, should be understood as a starting point for research and, one that leads on to the next step, rather than as an end in itself. A fourth resource, and one which has become a bit less fashionable (and even slightly controversial in the wake of Beazley), is the scholarly monograph devoted to a single vase-painter (e.g. Oakley, 1990). Unlike other areas of archaeology, Greek vase scholars have shied away from statistical analysis and are well aware of the shortcomings of using any of the resources listed above (including the BAPD) for such a purpose. Although attempts in this direction can be met with harsh criticism, the tide may slowly be turning towards a slightly more scientific approach to Greek pottery presentation (cf. Walsh, 2014; Smith, 2007).

The place of Greek figural pottery within the larger context of both online and print resources for the study of Classical art and archaeology more generally, should be emphasized. Because of their rich and informative iconography, Greek vases from all known centres, have found their place in major print reference works such as the Lexicon Iconographicum Mythologiae Classicae (LIMC) and its companion publication Thesaurus Cultus et Rituum Antiquorum (ThesCRA). Digital resources for Classical studies, most notably the Perseus Digital Library (http://www.perseus.tufts.edu/hopper/), includes Greek vases in its Art and Archaeology Artifact Browser, where one is able to search according to: collection, context, painter, period, potter, region, shape, ware, and keyword. The catalogue of vases in Perseus is by no means comprehensive, but the keyword function makes it more user-friendly than the more focused options mentioned above. It is also of critical importance to stress that ―most online citation databases subscribed to by academic and research libraries do not cover the scope of ancient art‖ (Stylianopoulos 2012: 714). And, although an increasing number of museums are putting their collections on the Web, their searchable databases are of most use to Greek vase scholars in search of publication or teaching illustrations (Stylianopoulos, 2012: 720-721), and are only of limited value either as research tools or as reliable introductions for non-specialists.

Finally, due to the wide variation in the native languages of these databases (and lack of consistent, controlled vocabulary even among databases of the same languages), it is difficult to conduct research across multiple systems. Inspired by the recent successes of Nomisma.org, a collaborative project dedicated to the definition of numismatic concepts with Linked Open Data (LOD) methodologies, we are endeavoring to create a similar project, Kerameikos.org, which would define the intellectual concepts of Greek pottery according to similar standards. Rather than relying on textual strings as controlled vocabulary (the painter, Exekias, in most Western scholarship is transliterated in Chinese to:

埃克塞基亚斯

concepts are represented by HTTP URIs, following the definition of Linked Open Data by Tim Berners-Lee (Berners-Lee, 2006). These URIs, when dereferenced, may offer both human-readable HTML or machine-readable RDF or JSON-LD, following standard ontologies. Lastly, by definition these concepts, encoded in RDF, link to concepts in other information systems. Kerameikos.org may provide a pathway to normalisation for many pottery databases, enabling the large scale data aggregation and subsequent analyses that are not currently possible within the discipline.

2. Kerameikos.org: Linked Data Applied to Ceramics

Kerameikos.org is primarily a thesaurus of ceramic concepts, but the user interface includes a handful of visualisation and analysis features that will be detailed further in this paper. In its initial development phase, more than 60 URIs were minted to define concepts in a small handful of categories: production place, person or organization responsible for some aspect of production (i. e. a painter or potter), shape, material, style, period, ware, and technique. These URIs contain a list of preferred labels in a variety of languages, which would enable multilingual user interfaces, and a list of URIs of identical or related concepts defined in other linked data systems. The latter component is an especially important one in the development of large-scale interoperable systems, as we will demonstrate.

Many pottery concepts defined on Kerameikos.org have also been defined in the LOD thesauri published by the British Museum (http://collection.britishmuseum.org/) and the Getty (http://vocab.getty.edu). Unfortunately, neither thesaurus is ―five star linked data (yet), as defined by Berners-Lee― that is, they do not link to related external resources. Kerameikos.org is intended to serve as the intermediary bridge between disparate systems, enabling the aggregation of content from collections that have adopted Kerameikos.org URIs natively, as well as collections that provide data that conform to British Museum or Getty identifiers. Therefore, the concept of the Black Figure technique, defined by http://kerameikos.org/id/black_figure, is the same as http://collection.britishmuseum.org/id/thesauri/x14736 in the British Museum thesaurus and http://vocab.getty.edu/aat/300387209 in the Getty Art and Architecture (AAT) thesaurus. Matches are also made with dbPedia, when possible. Painters and potters defined in Kerameikos.org are linked to URIs in the Virtual International Authority File (http://viaf.org) and the Lexicon of Greek Personal Names (LGPN: http://lgpn.ox.ac.uk/). Ancient production places are linked to URIs defined by the Pleiades Gazetteer of Ancient Places (http://pleiades.stoa.org). Kerameikos.org concepts will incorporate links to other systems as they become available.

Pottery concepts defined in this project have a varied audience, predominately vase scholars and their students. However, the benefit of Kerameikos.org to the museum community cannot be understated, particularly those institutions that own ceramics, but whose databases do not already incorporate controlled vocabulary that conform to existing thesaurus systems (e.g., the Getty). Even for those institutions that use Getty or British Museum identifiers to categorize their objects, Kerameikos.org could provide a pathway to wider publication, access, and scholarly query through its SPARQL endpoint, as semantic reasoning makes it possible to associate similar objects together regardless of whether they are categorised by Kerameikos.org, Getty, or British Museum URIs.

Archaeological databases have tremendous potential in shaping scholarly research in the discipline. Kerameikos.org could represent a stable set of identifiers, curated and maintained by pottery specialists, that excavations might incorporate into their own databases. This reduces the need for individual projects to maintain internal typologies, while making it easier for them to publish content into the larger LOD cloud. Even among those projects that implement semantic web methodologies, like CLAROS (http://www.clarosnet.org/XDB/ASP/clarosHome/), URIs for concepts (whether ceramic ones from Kerameikos.org or numismatic ones published by Nomisma.org) can aid in normalisation and enhancement of the user experience.

With the goals of Kerameikos.org in mind, the authors developed specifications for the first phase of the project:

- The ability for pottery specialists to create, edit, and publish identifiers (with training for the semantic meaning of properties, but without need to understand RDF).
- Human readable pages for each identifier.
- Geographic and quantitative visualisation features that specialists are accustomed to create in their research, delivered nearly instantaneously.
- Machine readable serialisations in relevant models (e.g., RDF/XML, Turtle, JSON-LD, KML, geoJSON, etc.).
- REST APIs that deliver data efficiently.
- A published ontology, with accompanying documentation.
- Receive, validate, and post RDF of vase data into an RDF triplestore to facilitate queries and visualisations mentioned above.

The prototype interface (both the front end and administrative back end) was developed upon a small subset of several dozen identifiers that represent the essential categories by which Greek vases are classified. The architecture of this system will be discussed in greater detail below, but we must first detail the thought processes that informed the development of an ontology and model that maps the organization of specialised ceramics knowledge into machine terms.

2.1. Ontologies and Models

2.1.1 Ontology and Model of Kerameikos.org Concepts

Before delving into the technical specifics relating to RDF models and ontologies, it was imperative to approach the challenge of knowledge representation by evaluating the traditional classification methods of Greek vases. Firstly, how does a pottery specialist categorize objects? As indicated above, Greek vases are classified and presented by the following fairly limited set of typologies: artist/creator (painter and/or potter: could be an individual or organization), decoration/iconography, material, period, place of production, shape, style, technique, and ware. Although these typologies (or classes, in ontological terms) apply specifically to Greek pottery, they are generally applicable to ancient ceramics of other periods and cultures.

In order to promote reusability, we opted to incorporate classes and properties from other RDF ontologies when possible. Several classes from CIDOC-CRM were adopted: E4_Period, E53_Place, and E57_Material. While the CRM provides an Actor class, we adopted Person and Organization from Friend of a Friend (FOAF) to designate painters, potters, and workshops. Most identifiers are simultaneously designated as concepts, as defined by Simple Knowledge Organization System (SKOS). Accordingly, since a production place, e.g. Athens, is reckoned to be a concept, but also has a geographic component, we have chosen to implement SpatialThing from the World Wide Web Consortium (W3C) geographic vocabulary (World Wide Web Consortium 2004) to encapsulate coordinates in the form of simple points (with geo:lat and geo:long) or more complex polygons that represent regions, encoded in geoJSON (Figure 1).

Some classes—such as shape, ware, or technique—are particular to the discipline and have no equivalent in other ontologies. The British Museum and Getty thesauri are intended to be applied to a diverse array of object types found in museums, and so their applications are more general. For example, in the British Museum thesauri, vessel shapes such lekythos and amphora are considered object types (rather than forms or shapes). As a result, we endeavored to create an ontology, identified by http://kerameikos.org/ontology# (prefix: 'kon'), for pottery-specific classes and properties that would fill in these gaps. Therefore, http://kerameikos.org/id/lekythos and http://kerameikos.org/id/amphora are defined by the kon:Shape class; http://kerameikos.org/id/black_figure is a kon:Technique; http://kerameikos.org/id/boeotian is a kon:Ware.

```
<http://kerameikos.org/id/athens>
   a ecrm:E53_Place, skos:Concept ;
   geo:location <http://kerameikos.org/id/athens#this> ;
   skos:broader <http://kerameikos.org/id/attica> ;
   skos:definition "Athens dominates the Attica region and is one of the
world's oldest cities, with its recorded history spanning around 3,400
years."@en ;
   skos:exactMatch
<http://collection.britishmuseum.org/id/place/x22744>,
<http://dbpedia.org/resource/Athens>,
<http://pleiades.stoa.org/places/579885>,
<http://vocab.getty.edu/tgn/7001393> ;
   skos:prefLabel "Atenas"@es, "Athens"@en, "Athènes"@fr, "アテネ"@ja .

<http://kerameikos.org/id/athens#this>
   dcterms:isPartOf <http://kerameikos.org/id/attica> ;
   a geo:SpatialThing ;
   geo:lat "37.972"^^<xsd:float> ;
   geo:long "23.726443"^^<xsd:float> .
```

Figure 1: Athens ID as RDF/TTL.

As mentioned previously, the most vital aspects of any concept defined in Kerameikos.org are the multilingual labels and links to related concepts in other LOD systems. SKOS provides a clear avenue for encoding labels with skos:prefLabel (preferred label) and skos:altLabel (alternative label) and linking external concepts with skos:exactMatch and skos:relatedMatch. The Exact Match property is used when linking to the Getty AAT, British Museum, VIAF, LGPN, and most other systems. The Pleiades Gazetteer of Ancient Places is the exception; skos:relatedMatch is used because the distinction between places as concepts and spatial features is not clearly defined. All IDs require an English definition (denoted by the skos:definition property), but definitions in other languages may be used. SKOS scope notes may be included to define rules governing the semantic application of a concept.

The other essential linking property incorporated from the SKOS ontology is skos:broader, used to connect concepts hierarchically. Logically, this would apply to production places: Athens is in Attica, which is in Greece. Techniques, shapes, periods, and styles may also be linked together with skos:broader, following the traditional organizational structure of Greek pottery. Bell kraters and volute kraters are specific renditions of the krater shape. The Black Figure technique, by definition, comprises both silhouette and incision techniques. Early Helladic III is the third phase of the Helladic period, which is one of a number of eras of the Aegean Bronze Age. Applying the SKOS ontology to these concepts makes it possible for Kerameikos.org identifiers to reflect the current methods of pottery classification.

2.1.2 Using CIDOC-CRM with Kerameikos.org Properties to Encode Vases

In order to demonstrate the utility of Kerameikos.org as both a LOD thesaurus and research portal, we ingested vase data into the project's underlying triplestore to facilitate analysis and visualisation. Following examples of applying CIDOC-CRM to Greek vases established by the British Museum, we developed two small datasets of vases: one derived from a CSV file provided by the Getty Museum (and processed with PHP into RDF/XML conforming to the British Museum CIDOC-CRM model) and the other extracted programmatically from the British Museum SPARQL endpoint. In all, data for more than 20 vases in the Getty and several hundred from the British Museum were processed into CIDOC-CRM and published into Kerameikos.org's RDF triplestore. An example of one vase is shown in Figure 2.

In this example, although many CRM properties were used, several properties from other ontologies were incorporated to simplify the model and improve processing efficiency: title from Dublin Core Terms and FOAF properties for encoding image URLs. In addition to a handful of pottery-specific classes, the Kerameikos.org ontology also includes a number of properties that more closely resemble the real-world organization of ceramic knowledge, e.g., kon:hasShape. With the exception of crm:P50_has_current_keeper, which links to a Kerameikos.org URI defining a legal body that curates the object, British Museum or Getty vases are defined by URIs in their respective vocabulary systems. Semantic reasoning inherent to Kerameikos.org's linked data architecture makes it possible to query objects defined in different thesauri.

```
<http://www.getty.edu/art/gettyguide/artObjectDetails?artobj=7982>
  ecrm:P108i_was_produced_by [
    ecrm:P14_carried_out_by <http://vocab.getty.edu/ulan/500018372> ;
    ecrm:P4_has_time-span [
      ecrm:P82a_begin_of_the_begin "-0449"^^xsd:gYear ;
      ecrm:P82b_end_of_the_end "-0449"^^xsd:gYear ;
      a ecrm:E52_Time-Span
    ] ;
    ecrm:P7_took_place_at <http://vocab.getty.edu/tgn/7001393> ;
    a ecrm:E12_Production
  ] ;
  ecrm:P32_used_general_technique <http://vocab.getty.edu/aat/300387208> ;
  ecrm:P45_consists_of <http://vocab.getty.edu/aat/300010669> ;
  ecrm:P50_has_current_keeper <http://kerameikos.org/id/getty_museum> ;
  kon:hasShape <http://vocab.getty.edu/aat/300264937> ;
  dcterms:identifier "71.AE.442" ;
  dcterms:title "Oil Jar with a Departing Warrior"@en ;
  a ecrm:E22_Man-Made_Object ;
  foaf:depiction <http://www.getty.edu/art/collections/images/l/00798201.jpg> ;
  foaf:thumbnail <http://www.getty.edu/art/collections/images/
m/00798201.jpg> .
```

FIGURE 2: RDF OF A VASE FROM THE GETTY MUSEUM.

2.2. Architecture

Kerameikos.org's software architecture is nearly identical to that of Nomisma.org and similar to other American Numismatic Society projects, due to Gruber's software development role at the American Numismatic Society. Kerameikos.org is an open source collection of scripts that handle the front-end user interface and back-end administrative functions. The framework is available on Github at https://github.com/kerameikos/framework. This collection of scripts--predominantly XSLT, XForms (and supporting standards), and Javascript--binds together open source server applications that run in Java, specifically in Apache Tomcat.

The web application architecture is based on XRX (XForms, REST, and XQuery) (McCreary 2008), though XQuery has been substituted for the combination of SPARQL and Apache Lucene/Solr. Kerameikos.org identifiers are stored in the filesystem as RDF/XML files, maintained in a Github repository (https://github.com/kerameikos/ceramic-ids), which enhances access in addition to providing version control. Apache Solr is the de facto standard search index in the cultural heritage sector, prevalent in repository applications, like Fedora, or content aggregators, like Europeana. XForms, a W3C standard schema for advanced web form functionality, is the backbone of Kerameikos.org, handling the editing and publication process of ceramic concepts in addition to processing RDF data dumps from contributors (World Wide Web Consortium 2009). XForms enables the publication of data into the project's internal Solr index and RDF triplestore (Apache Fuseki is currently implemented in Kerameikos.org), but also supports REST interactions with a variety of external Linked Open Data services, which will be discussed in greater detail below. The XForms processor used by Kerameikos.org is Orbeon (http://orbeon.com). Orbeon provides an array of processors that serialise the canonical RDF/XML concepts, Solr and SPARQL query results, and various APIs into all of the content types required to power the user interface—from human-readable HTML to machine-readable geoJSON-LD, Turtle/RDF, and KML for mapping. Kerameikos.org's architecture is modularised. With minor modifications, the Github/filesystem-based storage could be replaced with any number of off-the-shelf NoSQL database solutions. Fuseki could be substituted with any SPARQL 1.1-compliant RDF triplestore, Solr could be replaced with any search index software that supports REST, and Orbeon could be swapped with other XForms processors, such as XSLTForms or betterFORM. While the Kerameikos.org framework does include code fragments which are specific to this project, we aim to make this application more generalisable and adaptable for publishing other types of LOD thesauri.

The system is flexible and scalable. This framework differs from the LOD thesauri published by the British Museum, Library of Congress, and the Getty Museum. Kerameikos.org concepts are made available through linked data methodologies, but the project's user interface includes various query APIs and visualisation tools; it is a research portal for Greek pottery in addition to a vocabulary system. Furthermore, the back-end system is enterprise-ready: scalable to millions of ceramic concepts and hundreds of millions of triples, managed through stable information systems workflows. These applications scale well beyond the ceiling we ever anticipate encountering, even if we expand beyond Greek pottery into the typologies of ceramics from other periods and cultures.

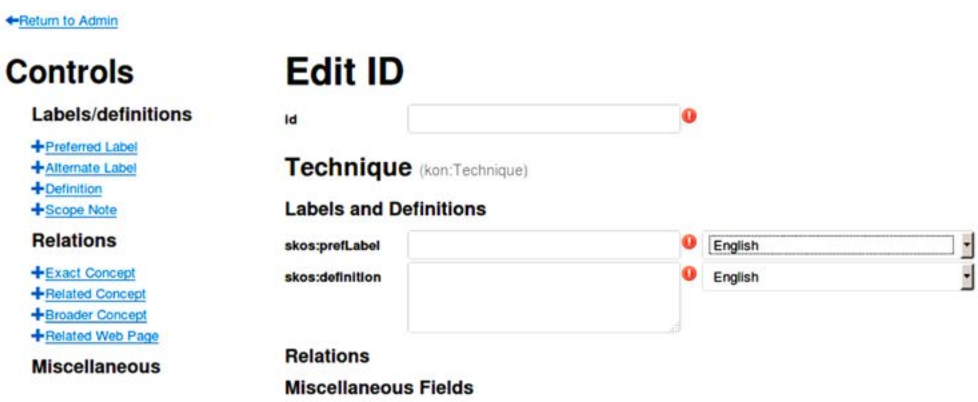


FIGURE 3: EDITING IDS WITHIN THE WEB FORM.

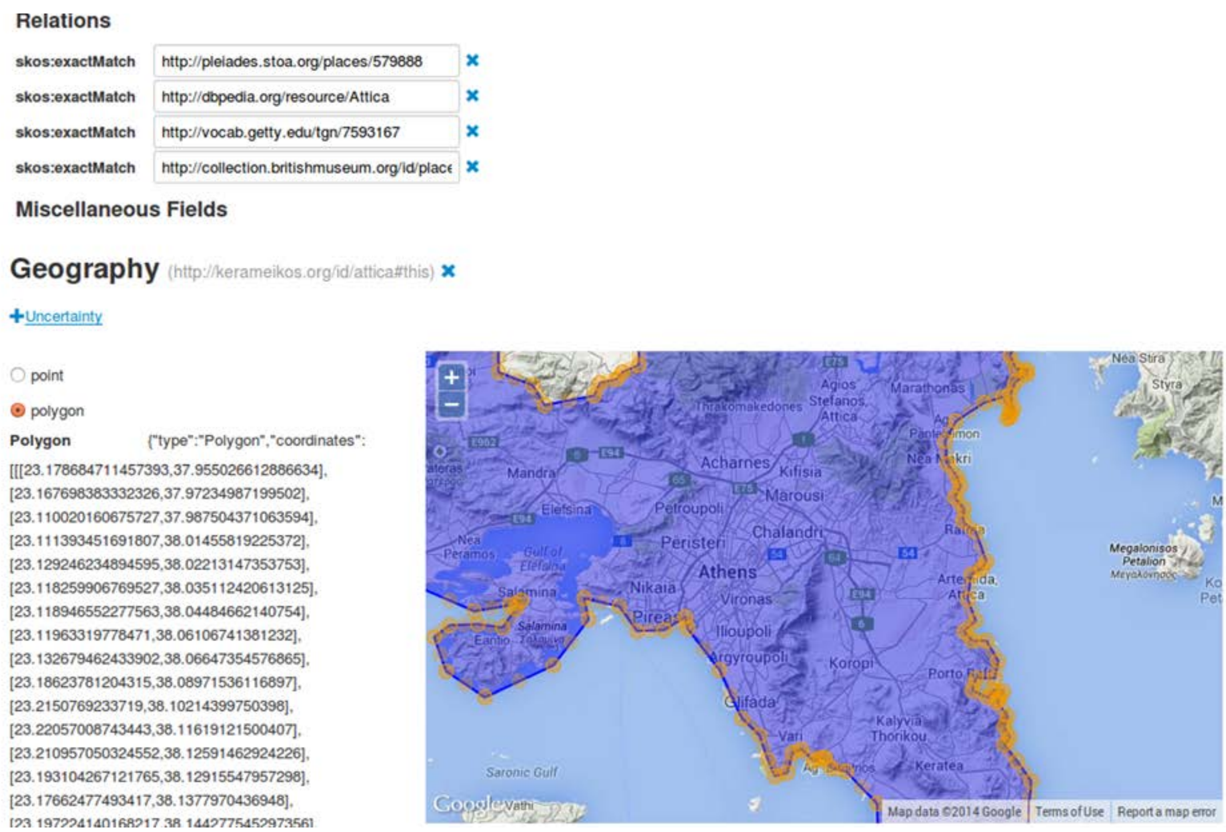


Figure 4: Drawing a map in OpenLayers and geoJSON.

2.2.1 Administrative Editing and Publishing Interface

The most important aspect of the administrative interface for Kerameikos.org is the editing interface that pottery scholars may use to create well-defined concepts without the need to edit RDF directly or have any specific technical knowledge about the RDF publication process (although knowledge of what a property means semantically, e.g., when to use skos:exactMatch rather than skos:relatedMatch, is required). Kerameikos.org presents editors with an intuitive web-based interface, with controls to expedite the lookup and linking mechanisms with other LOD thesaurus systems (Figure 3).

To reiterate, XForms drives the editing interface. While a comprehensive discussion of the standard is beyond the scope of this paper, a brief synopsis will be provided.[1]

In recent years, support for the standard has moved from native browser support to standalone processors, whether client or server-side. XForms' XML controls are embedded within XHTML files, and the processor is responsible for transforming this document into a functional HTML-based web form, including all of the necessary CSS for styling and Javascript to handle sophisticated validation, REST interactions, and other controls on the web page (e.g., to add a new skos:prefLabel field into the document model when a user clicks a button in the user interface). XForms adheres to the Model-View-Controller (MVC) architectural framework. As an XForms developer, one needs only to focus on the development of the schematic functionality of the web form. The XForms code can be ported from processor to processor with the expectation of continued performance. XForms applications support the editing and validation of complex XML models based on XPath and XML Schema.

Below are some of the more common validation scenarios:

- A concept must have an English preferred label and definition, and there may be no duplicate languages for these properties.
- Latitude and longitude must be decimal values between -180 and 180.
- URIs inserted into linking properties, like skos:exactMatch, must be valid.
- A geo:SpatialThing must contain either geo:lat and geo:long or a geoJSON-encoded polygon.

If all validation scenarios are met while editing an ID, the Save button will become active, and upon saving, the XForms application will execute events to serialise the RDF/XML model into an XML file and store it on the disk (for pushing to Github later). The RDF triples will be refreshed in the triplestore through SPARQL/Update, and the RDF/XML will be transformed into another XML model conforming to the Solr schema and indexed for keyword and faceted search in the public user interface.

This RDF editor allows a user to insert all of the aforementioned properties used in Kerameikos.org's data model, but the application includes a number of features that expedite the editing process. When creating a Place concept,

[1] For further information about XForms applied to cultural heritage informatics, see Gruber et al. 2010.

FIGURE 5: LINKING TO VIAF.

the user does not manually enter geographic coordinates, but rather uses an OpenLayers-based map to either create a point or draw a polygon. With a Javascript extension of the application, Orbeon interacts with OpenLayers to update the RDF with these features (Figure 4).

The XForms application also interacts with REST APIs provided by external thesaurus systems. The editor may wish to link a Greek vase painter defined in Kerameikos.org with identities defined on VIAF. VIAF provides query results in an RSS feed. A search for ―Exekias‖ may yield numerous results, as VIAF still has much work to do in the disambiguation of entities (Figure 5). The editor may select applicable boxes, defining them as Exact Matches, and the URIs will be imported in the RDF model. Furthermore, the XForms application will request the RDF for each of these entities from VIAF, extracting owl:sameAs URIs for matches in other systems, such as dbPedia or the Gemeinsame Normdatei (GND), maintained by the Deutsche Nationalbibliothek. If the Kerameikos.org identifier contains a skos:exactMatch for a dbPedia entry, preferred labels may be extracted directly from dbPedia's RDF. Similar lookups are available for other types of concepts--like shapes and techniques—in the Getty AAT, which are made available through their SPARQL endpoint. This mechanism will be extended

FIGURE 6: MAP AND TIMELINE OF USER INTERFACE.

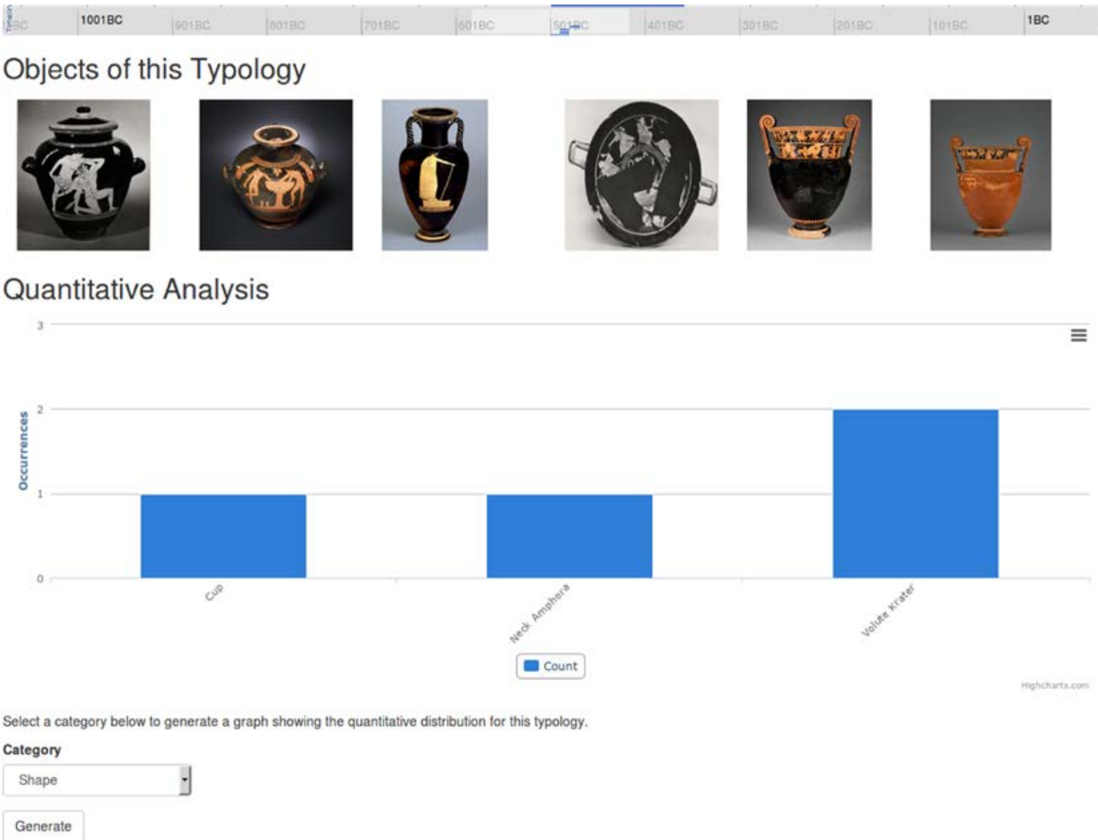


Figure 7: Thumbnails and analyses.

in the Kerameikos.org editing interface when the other Getty vocabularies become available as LOD. Essentially, specialists can hit the ground running, rapidly making links to other vocabulary systems in the creation of Kerameikos.org concepts, with labels harvested programmatically from dbPedia.

2.2.2 Public User Interface

Presently, the public user interface of Kerameikos.org is generated primarily through Orbeon and the XML Pipeline Language (World Wide Web Consortium 2002). It includes a browse page generated from Solr query results (http://kerameikos.org/id/), a page which serves as a placeholder for documentation of the ontology (http://kerameikos.org/ontology), a SPARQL endpoint (http://kerameikos.org/sparql), and pages for each individual concept. We intend to introduce a variety of REST APIs which either interface with SPARQL to deliver common queries more efficiently or deliver machine readable serializations more expediently. The ontology page is generated dynamically from the OWL/RDF file and available in RDF/XML and Turtle. There is much work to be done throughout the public user interface, but the pages for individual concepts provide the clearest picture of the potential for Kerameikos.org as both a dynamic thesaurus of ceramic concepts and a research portal for students and scholars of the discipline.

The pages for each entity typically contain the following in the body: the RDF predicates and objects rendered in HTML5 (with underlying RDFa attributes in the HTML, enabling extraction of triples with an RDFa parser), a map and timeline (with the Timemap Javascript library [https://code.google.com/p/timemap/]) showing the distribution of a particular typology over time and space (Figure 6), a list of thumbnails representing the typology, and a very simple form which allows a user to perform basic quantitative analysis of the typology, e.g., to show the numeric distribution of shapes created by a Greek potter (Figure 7). The maps and quantitative analyses are generated by using XSLT to serialise underlying SPARQL query results into the JSON models required by the map and chart Javascript libraries. The sidebar includes links to alternate serialisations of the concept in RDF/XML, Turtle, JSON-LD, and KML, as well as contextual information extracted from other LOD sources (e.g., birth and death dates extracted from Text Encoding Initiative (TEI) XML provided by LGPN). Furthermore, Orbeon supports semantic HTTP 303 redirects and content negotiation, enabling technical users to extract JSON-LD by requesting the content type of ―application/json from http://kerameikos.org/id/black_figure.

We support both content negotiation and REST to maximise access and reuse of data.

We will eventually enhance these quantitative analysis features in order to deliver the types of visualisations that scholars are accustomed to generating in their own research. With more typologies defined in Kerameikos and more vase data in the system, we might alter research possibilities in the discipline. Scholars could, in theory, spend more time on analysis and interpretation rather than data gathering, with museum and archaeological materials made available for query through the same system. Although the BAPD achieves this to the same extent at present, its primary function is to serve as a repository of images and bibliography (built on the format of Beazley's 'lists'), its nucleus is Athenian pottery specifically, and its search terms are not linked externally. With Kerameikos.org's analysis and visualisation tools, it might be possible to conduct sophisticated queries based on geography and statistics, paving the way for studies of the ancient economy (akin to research methodologies in numismatics). These tools may even extend the boundaries of the discipline, making it possible to formulate new types of questions that may never have been considered without large-scale aggregation and visualisation of ceramics data.

3. Future

In this paper, we have effectively demonstrated the potential of Kerameikos.org and provided a glimpse of the sophisticated query and visualisation mechanisms made possible by applying linked open data methodologies to Greek vases. On a technical level, this is not experimental digital work. We are applying stable, tried-and-true technologies to a field in need of a re-envisioning of the role that digital tools play enhancing the research and synthesis process. Kerameikos.org will only succeed with community buy-in, and this interest can take years of hard work to generate. With this in mind, we are forming a steering committee that consists both of Greek pottery and cultural heritage informatics specialists. Our goal is to develop a system that effectively represents the intellectual organisation of ceramics knowledge, based upon current Linked Open Data standards.

Our initial phase has focused on the creation of a thesaurus of the simplest typologies to model, and we will progress toward more complex challenges in later iterations of the Kerameikos.org project: a system for maintaining data dumps of CIDOC-CRM describing vases, iconographic topic modeling, enhanced models for representing and linking painters, potters, and workshops (which, in turn, would enable social network analysis of the ancient pottery industry), and further interaction with other LOD

projects, such as the Pelagios Project and Standards for Networking Ancient Prosopographies. We will continue to build this project and adapt it to the needs of the Classical archaeology and museum communities.

Acknowledgements

The authors would like to acknowledge Natasha Dakouri-Hild at the University of Virginia, Thomas Mannack of the Beazley Archive at Oxford University, Sebastian Rahtz, Director (Research Support) of Academic IT Services at Oxford, Vladimir Alexiev at Ontotext, for insights into ontology development, and the linked data team at the Getty Museum (specifically, Joan Cobb, Patricia Harpring, and Brenda Podemski), for providing vase data.

References

BERNERS-LEE. (2006). *Linked Data - Design Issues*. [Online]. Available from: http://www.w3.org/DesignIssues/LinkedData.html. [Accessed: 28 July 2014].

BOARDMAN, J. (2008). *The History of Greek Vases*. London: Thames and Hudson.

COOK, R.M. (1997) *Greek Painted Pottery* (3rd edn). London: Routledge.

GRUBER, E., FITZPATRICK, C., PAROD, B. and PRATER, S. (2010). XForms for Libraries: An Introduction, *code4lib* 11, Available from: http://journal.code4lib.org/articles/3916.

KURTZ, D.C. (1985). Beazley and the Connoisseurship of Greek Vases, *Greek Vases in the J. Paul Getty Museum 3*, pp. 237-250.

MCCREARY, D. (2008). *XRX: Simple, Elegant, Disruptive*. [Online]. Available from: http://archive.oreilly.com/pub/post/xrx_a_simple_elegant_disruptiv_1.html. [Accessed: 28 July 2014].

MOORE, M B., PHILIPPIDES, M.Z. (1986). A*gora 23: Attic Black-Figured Pottery*, Princeton: American School of Classical Studies at Athens.

NØRSKOV, V. (2002). *Greek Vases in New Contexts*, Aarhus: Aarhus University Press.

OAKLEY, J.H. (1990). *The Phiale Painter*, Mainz am Rhein: Phillip von Zabern.

ROUET, Ph. (2001). A*pproaches to the Study of Attic Vases: Beazley and Pottier*, Oxford: Oxford University Press.

SMITH, T.J. (2005). The Beazley Archive: Inside and Out. *Art Documentation* 24, pp. 22-25.

SMITH, T.J. (2007). Athenian Black-figure Pottery.‖ In S. Solovyov (ed.), Borysthenes-Berezan. *The Hermitage Archaeological Collection 2*, St. Petersburg: State Hermitage Museum.

STYLIANOPOULOS, L.W. (2012). Surveying the Scholarship. In: Smith, T.J. and Plantzos, D. (eds.). *The Companion to Greek Art*, Malden, MA and Oxford: Wiley-Blackwell, pp. 711-722.

WALSH, J. (2014). *Consumerism in the Ancient World: Imports and Identity Construction. Routledge Monographs in Classical Studies* 17, London; New York: Routledge.

World Wide Web Consortium. (2002). *XML Pipeline Definition Language Version 1.0.* [Online]. Available from: http://www.w3.org/TR/xml-pipeline/. [Accessed: 28 July 2014].

World Wide Web Consortium. (2004) *W3C Semantic Web Interest Group: Basic Geo (WGS84 lat/long) Vocabulary*. [Online]. Available from: http://www.w3.org/2003/01/geo/. [Accessed: 28 July 2014].

World Wide Web Consortium. (2009) *XForms 1.1*. [Online]. Available from: http://www.w3.org/TR/xforms/. [Accessed: 28 July 2014].

The Digital Archaeological Workflow: A Case Study from Sweden

Marcus J. SMITH

Swedish National Heritage Board (RAÄ)
marcus.smith@raa.se

Abstract

The Digital Archaeological Workflow (DAP) is a programme of work being carried out at the Information Development Unit at the Swedish National Heritage Board, in partnership with other Swedish archaeological stakeholders. The programme aims to streamline the flow of archaeological data between different actors in the Swedish archaeological process, and to ensure that this data is preserved in a sustainable and accessible manner. It aims to address a number of problems which have hampered the practice of archaeology in Sweden for some time, but which have now started to become more acute as digital technology saturates the processes involved.
This paper describes the work accomplished so far and the progress that has been made, as well as our plans for the rest of the programme. In particular, it focuses on our mapping of the existing archaeological processes in Sweden and our use of the semantic web as a fundamental method.

Keywords: archaeological process, semantic web, linked open data, sweden, digitisation

1. Introduction

The Digital Archaeological Workflow (DAP) (Riksantikvarieämbetet, 2014b) is a programme of work being carried out at the Information Development Unit of the Swedish National Heritage Board (Riksantikvarieämbetet, 2014b), in partnership with the major Swedish archaeological stakeholders. The programme aims to streamline the flow of archaeological data (and its associated metadata) between different actors in the Swedish archaeological process, and to ensure that this data is preserved in a sustainable and accessible manner. It aims to address a number of problems which have hampered the practice of archaeology in Sweden for some time but which have now started to become more acute as digital technology saturates the processes involved. It also aims to raise the quality of archaeological and heritage data curated by the Swedish National Heritage Board.

2. Background

With the transition from analogue to digital archaeological fieldwork recording in the mid-to-late 1990s, and the parallel rise of digital systems supporting local and national government, increasing amounts of digital data about archaeological fieldwork and ancient monuments is being produced. However, the dissemination and long-term preservation of this digital material are problems which have yet to be satisfactorily solved in Sweden, and which grow more acute with each passing year. Furthermore, the benefits offered by these digital technologies are not yet being fully exploited: many processes within archaeological planning and fieldwork appear superficially to be digital, but in reality remain mired in increasingly incongruous analogue paradigms. The DAP programme aims to address a number of these problems, which are as follows:

- There is no centralised register of archaeological fieldwork in Sweden, or indeed of any events which affect the status of ancient monuments. This makes it difficult to keep track of what fieldwork – both research and commercial – is going on where. It also impedes the collation of past fieldwork (in relation to a particular site or monument, for example).
- Sweden has no central digital archive or repository for the storage of either archaeological fieldwork data or written reports; nothing corresponding to the ADS in the UK (ADS, n.d.) or EDNA in the Netherlands (DANS, n.d.). Site records are now produced almost exclusively digitally both during and post-excavation, and so the lack of a central digital repository to receive, curate, publish, and preserve such records means that valuable archaeological data is thus increasingly at risk of being lost.
- Despite the fact that almost all of the data and administrative metadata surrounding archaeological work are digital-born, they are nonetheless still handled according to analogue paradigms. We still speak in terms of – and create! – digital *documents* (rather than data) multiple copies of which are distributed to various parties. This is particularly apparent when information must be shared between different organisations, which often results in digital data being printed out, sent through the post, re-digitised, edited, printed out again, and posted on. Such practices are time-consuming, inefficient, and cause data loss – particularly in the case of geographical information. Where documents *are* sent digitally, it is typically as an attached documents via email rather than as structured data in open formats.

- As a consequence of the lack of a digital repository for archaeological data, there is a good deal of raw data – primarily fieldwork records – which remains unpublished and inaccessible. Sources of archaeological data which *are* currently made available digitally by various national and local bodies are often unstructured and are not typically linked together: each database is a silo, unconnected with other data sources to which it might be related. Even in datasets which include references to related resources, these links are not typically published in a machine-readable format as semantic links. This leads to inefficiencies in information transfer, duplication of data and effort, and to information describing the same 'objects' (in part or in full) being stored in different systems within different organisations.

3. Aims and methodology

The DAP programme is intended to address these problems, initially over the course of a five-year period in concert with an effort to raise the quality of archaeological data maintained by the National Heritage Board. As much as possible the programme will try to make use of existing open, standardised, platform-agnostic data formats and protocols to streamline information transfer between organisations. It will release a series of open taxonomies and ontologies for common Swedish archaeological terms and concepts on the semantic web in order to facilitate data interoperability. And it will designate a secure digital repository both for the raw data and reports arising from fieldwork and research. The Swedish National Heritage Board aims to make this information freely available in the Web as linked open data. The programme seeks to address the problems described above in the following ways:

- Implement fully digitised, seamless transfer of information between different stakeholders in the archaeological process;
- Designate (or create) a central digital archive for archaeological data;
- Ensure that this data remains accessible to all; not just reports and syntheses, but also the 'raw' fieldwork data, GIS, geophysics, laboratory analyses, etc;
- Ensure that this data is released in a structured, open, machine-readable format which includes semantic links to related digital objects using established standards for linked open data;
- Strive, wherever possible, to ensure that the data is made available under an open licence (e.g. Creative Commons (n.d.) or ODC (Open Knowledge Foundation, n.d.) licences)
- Ensure that the quality of the data is described in its metadata, so that judgements may be more easily made about its suitability for a particular task, and its comparability with other datasets (this will hopefully also involve raising the quality of the data, in addition to describing it);
- Include, as part of this data, information about archaeological 'events': things which create information about or change the status of sites and monuments: anything from an excavation or field survey to the granting of planning permission, or even a natural disaster that has caused damage to a monument.

A central part of the vision for the DAP programme is Linked Open Data (Berners-Lee, 2009). Not only does linked data facilitate connecting heterogeneous datasets, it also places that data into a broader context, aids discoverability, and enables a hitherto unachievable degree of openness in the archaeological process, both in terms of the publication of results and raw data, but also in terms of the transparency of the process itself. We feel that it is important to address these issues if DAP is to succeed and archaeology is to move forward as a discipline.

4. Work on DAP to date

In the preparatory groundwork for the DAP programme over the preceding two years (2012–2014), representatives from the archaeological and heritage sector have been involved in a variety of modelling workshops aiming to map the state of the Swedish archaeological process as it stands today, and to identify problems and causes of friction. As a result, awareness and engagement with the programme within the sector is already high, and we have established contacts with local administrative bodies, fieldwork units, museums, and researchers. This cooperation at an early stage is key, as the execution of the DAP programme is not something that the National Heritage Board intends to (or is capable of) carrying out alone; rather, it requires the cooperation of, investment from, and changes from, all participants in the archaeological process.

The outcomes from these early modelling workshops also mean that much of the work of describing and mapping the processes as they exist today is already done (although some details remain): we have a good working model of how things work in Swedish archaeology today, and – more importantly – a good idea of where things don't work, and why (fig. 1). The results of this modelling work will form the basis upon which we will define a new, fully digitised archaeological process.

In parallel with the modelling of the processes involved in Swedish archaeology, an equal amount of attention has been devoted to the conceptual modelling of the information objects involved in these processes. This conceptual modelling (fig. 2) describes the information objects within the archaeological process both according to current established practice but also as laid down in the appropriate laws governing and protecting Swedish cultural heritage (primarily Kulturmiljölagen [Regeringskansliet / Lagrummet, 2014]). The resulting conceptual models will inform the creation of a new information model for archaeological data, as well as a new systems architecture for the National Heritage Board. Thus far, the conceptual models we have arrived at are very similar to English

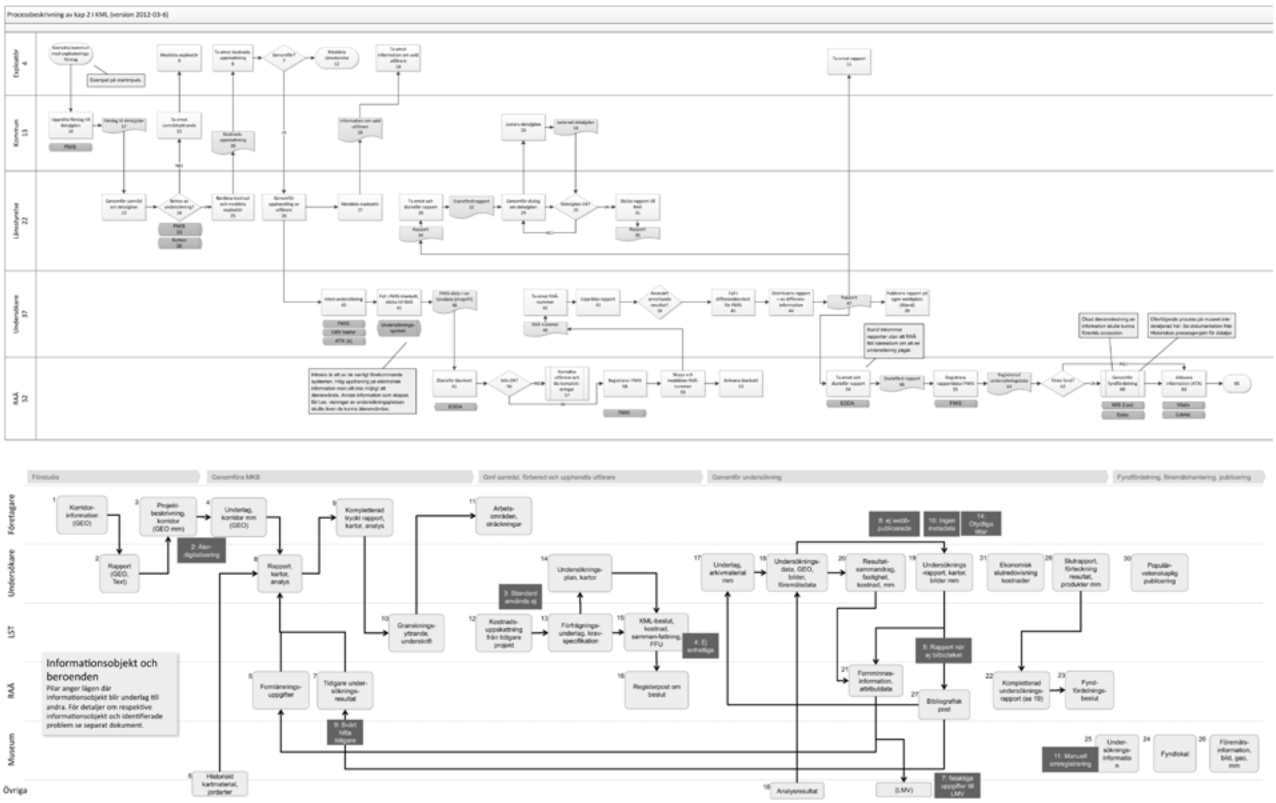


Figure 1: The results of the work modelling the current Swedish archaeological process, showing the course of a typical case in terms of processes (top) and information flow (bottom). The darker boxes in the lower diagram highlight 'pain points' – where there is friction in the current process, or where information is lost. (CC-BY 2012 Riksantikvarieämbetet).

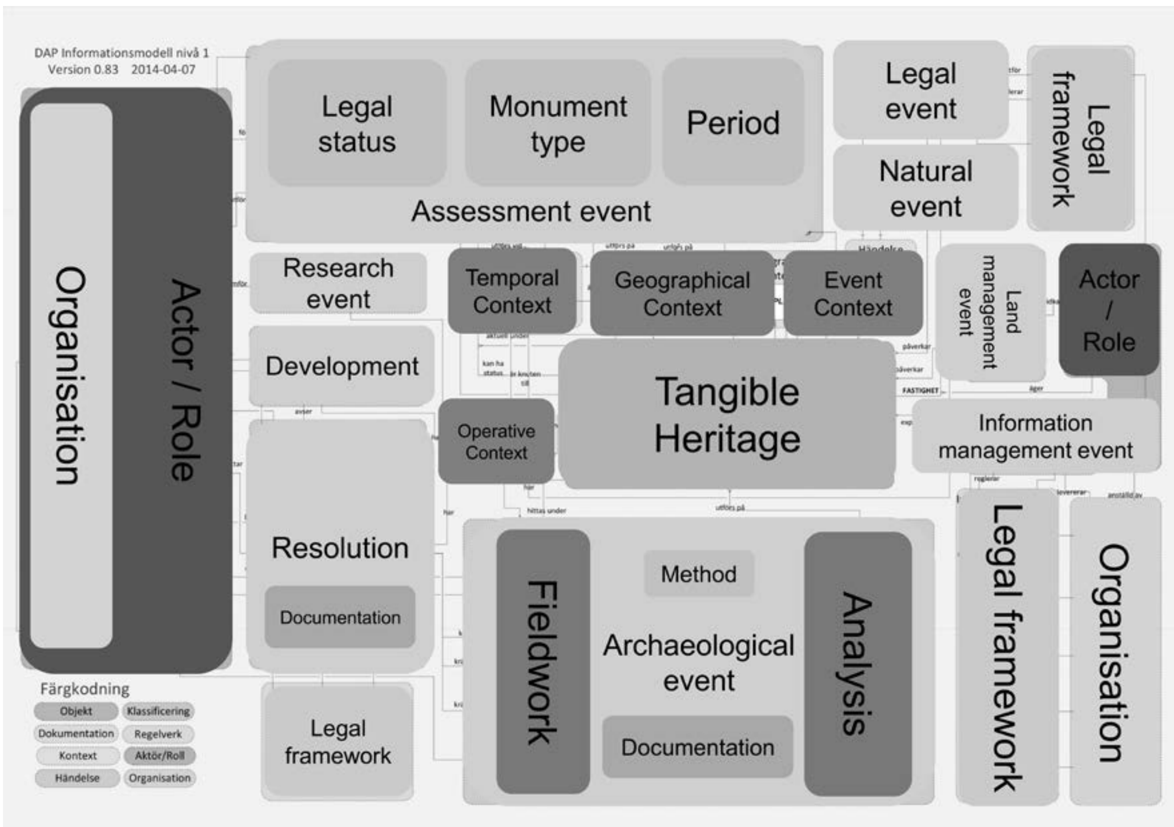


Figure 2: A rough outline of the top-level information model for DAP. (CC-BY 2014 Riksantikvarieämbetet).

Heritage's MIDAS (English Heritage, 2012) model, which will hopefully facilitate the subsequent process of mapping Swedish concepts to international standards such as CIDOC CRM (2013) in order to make Swedish data more widely interoperable.

In addition to the early and ongoing modelling work, a couple of elements of a future DAP-based infrastructure are already in use. In 2010 the Swedish National Heritage Board launched SOCH (Swedish Open Cultural Heritage) (Riksantikvarieämbetet, n.d.a), an aggregator and web-service which harvests meta-data describing ancient monuments, artefacts, and historic buildings, events and personages into a central index, creates semantic connections between these objects (fig. 3), and then publishes the data as RDF linked open data queryable via a web service. The metadata and objects are drawn from the collections of just over 40 museums and local societies in Sweden, as well as from the Heritage Board's own datasets (primarily the sites and monuments database [Riksantikvarieämbetet, 2014c] and the database of built heritage [Riksantikvarieämbetet, 2014a]), and SOCH also acts as Sweden's national aggregator for Europeana (n.d.), delivering the semantically enriched data on to be linked into an ever broader international context. SOCH forms a platform linking together objects from disparate data sets, upon which third party applications may be built. Foremost among these is the National Heritage Board's own 'Kringla' site (Riksantikvarieämbetet, n.d.b), which forms the main end-user interface to SOCH; but there are a number of other third-part mobile and web applications built around the SOCH interface.

More recently, in 2013, the National Heritage Board launched 'Samla' (Riksantikvarieämbetet, n.d.d), an 'open archive' web site for the Board's own publications, but crucially also for fieldwork units to upload and publish their own site reports as PDFs. While not a digital archive in a preservation sense, Samla nonetheless facilitates the digital dissemination of reports in a way that was previously not available, and has had a very positive reception from the archaeological community with over three thousand reports uploaded to the service to date.

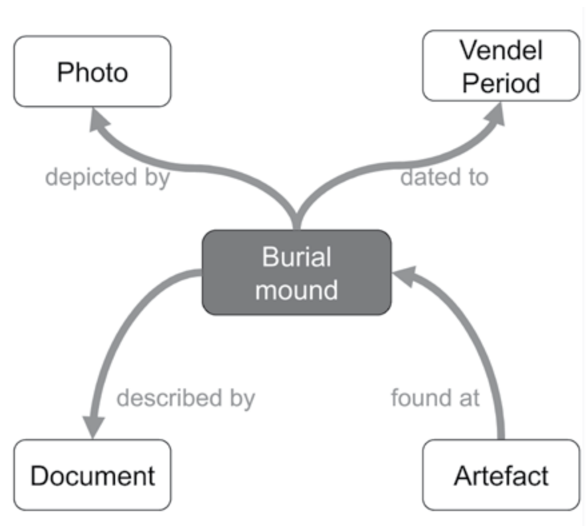


Figure 3: A schematic diagram showing examples of the kinds of semantic relations supported by SOCH for cultural heritage objects in the index. (CC-BY 2014 Riksantikvarieämbetet).

5. Cooperation and Collaboration

The new digital archaeological workflow which the DAP programme aims to establish, while based on the existing analogue processes, will mean a new way of working for archaeologists and local government employees in Sweden: it will not merely be a digitised version of the same process that exists now, and will involve changes to the information systems not only of the National Heritage Board, but also of local government bodies and of fieldwork units. This anticipated change in working practices across the sector underscores the importance of close collaboration and cooperation with other stakeholders from an early stage. While there will undoubtedly be technical challenges in the implementation of a future DAP-based infrastructure, by far the strongest potential barriers to adoption and success are the rather those of mindset (resistance to change and to a more open way of working) and of licensing (resistance to open data in particular).

The importance of cross-sector cooperation and collaboration in this work cannot be overstated. Any future DAP-based distributed infrastructure will have users across many different organisations, creating, modifying, and accessing resources and data from just as many different organisations, each with their own responsibilities and datasets, but linked together as open data: local authorities with resolutions pertaining to ancient monuments, the Forest Agency with information about monuments in forested areas, museums with finds in their collections, the National Land Survey with geospatial data, researchers… the list goes on. It is therefore imperative that we not only secure engagement from other stakeholders with the sector, but also that we plan the wider infrastructure carefully, ensuring not only that the National Heritage Board is able to manage the data we're responsible for, but also that we're able to connect to (and fetch) data that external bodies are responsible for, react when they change, and ensure that those other bodies are also able to make use of DAP data in their own systems.

6. Planned work

The DAP programme remains at an early stage of development, and the majority of the practical work remains to be done.

In order effectively digitise the archaeological process in Sweden and ensure that information flows smoothly between different stakeholders, we will need to define (or, preferably, specify an existing) set of protocols and formats for data publication and interchange. These will have to be open standards, preferably compatible with our goals of publishing data as linked open data, and DAP will aim to re-use existing solutions where possible.

In order to ensure that digital archaeological data remains accessible in the long-term, we will need to specify processes and practice for a digital preservation workflow, as well as either creating or designating one or more digital repositories. And in order to ensure that all the disparate datasets that will be made accessible – of different types, on different topics, and from different systems – we will need to ensure that there exists a common ontological reference model, with mappings for the different data sources and shared taxonomies.

7. Current status

Despite the scale of the DAP programme, and the large proportion of the work that will require slow and careful planning before ultimately seeing fruition, there are nonetheless a number of activities we can begin with immediate effect immediately and which can hopefully deliver important results in a short space of time. Chief among these is to take an inventory of what, exactly, is out there in terms of existing digital records held by fieldwork units, and the scale of the risk that it may be lost if not rescued and preserved promptly. This is important firstly to get an idea of the scale of the problem – which we suspect to be large – and to plan for its imminent collection and preservation, but also in order to see what types of information are recorded and how they are structured, so that this can inform our future data modelling and mappings. We are also undertaking a detailed study of user-needs within the sector, in order to ensure that the results of DAP actually solve the problems that professional users of Swedish heritage data face in the course of their work.

Two other activities which can be carried out immediately for swift results bear mentioning. Firstly, a maquette 'events register' to aid in modelling archaeological events (fieldwork, monuments inventories, etc) to get a feel for what sorts of data are involved, how such a register might work in practice, and how it might be connected to other information sources. This would not be a finished product, but rather a proof-of-concept mock-up to inform future development and architectural decisions.

Secondly, in parallel with the modelling of archaeological concepts mentioned above, it is necessary to create and adapt a series of shared controlled vocabularies for Swedish archaeological concepts. Some of these already exist, albeit not in machine-readable form (e.g. the Swedish Monuments Types Thesaurus [Riksantikvarieämbetet, 2014d]) while others will have to be created from scratch or adapted from existing standards from other countries (e.g. an events types thesaurus). Such thesauri will have to be created and released in digital form irrespective of the final form the DAP infrastructure ultimately takes, so there is nothing to be lost in beginning work early. Ideally we would hope to be able to release these structured vocabularies as linked open data SKOS (Miles & Bechhoffer, 2009) authorities (cf. the SENESCHAL (2014) project in the UK), although how these will be stored and managed internally remains to be seen. We foresee the need for, among others, common structured vocabularies for the following terms in Swedish archaeology:

- Monuments types
- Legal status
- Events
- Periods
- Materials
- Built heritage
- Evidence types
- Techniques
- Artefact types

Note, however, that it is far from clear that all of these should be produced or maintained by the National Heritage Board – some lie firmly within the remit of other heritage bodies.

8. Challenges Ahead

There are of course still a number of outstanding challenges which we have yet to satisfactorily resolve, given the early stage the programme is at. Among these are how the data within a distributed linked data infrastructure such as the one DAP proposes should be properly versioned, and how long-term preservation is going to work in practice. We very much welcome suggestions and feedback on our plans for DAP; we're still finding our way as we go!

The DAP programme is a massive and ambitious undertaking, and we'll be lucky to deliver even half the infrastructural components in its vision within the first five-year period. Nonetheless, we hope to produce deliverables on a fairly regular basis that simplify and digitise parts of the Swedish archaeological process when taken alone, and when taken together contribute to a broader interconnected infrastructure. DAP does not aim to reinvent the wheel, but rather to make use of existing interoperability standards, existing datasets and data structures, 'joining the dots' by linking them together into a rich infrastructure of open Swedish heritage data.

References

ADS. (n.d.). Archaeology Data Service. [Online] Available from: http://ads.ahds.ac.uk/ [Accessed: 17th July 2014].

BERNERS-LEE ,T. (2009). Linked Data – Design Issues. [Online] Available from: http://www.w3.org/DesignIssues/LinkedData.html [Accessed: 17th July 2014].

CIDOC CRM. (2013). The CIDOC Conceptual Reference Model. [Online] Available from: http://www.cidoc-crm.org/ [Accessed: 17th July 2014].

CREATIVE COMMONS. (n.d.). Creative Commons. [Online] Available from: http://creativecommons.org/ [Accessed: 17th July 2014].

DANS. (n.d.) e-depot voor de Nederlandse archeologie. [Online] Available from: http://www.edna.nl/ [Accessed: 17th July 2014].

ENGLISH HERITAGE. (2012). MIDAS Heritage: The UK Historic Environment Standard. [Online] Available from: http://www.english-heritage.org.uk/publications/midas-heritage/midas-heritage-2012-v1_1.pdf [Accessed: 17th July 2014].

EUROPEANA. (n.d.). Europeana. [Online] Available from: http://europeana.eu/ [Accessed: 17th July 2014].

MILES, A., BECHHOFER, S. (2009). SKOS Simple Knowledge Organization System Reference. [Online] Available from: http://www.w3.org/TR/skos-reference/ [Accessed: 17th July 2014].

OPEN KNOWLEDGE FOUNDATION. (n.d.). Open Data Commons: Legal Tools for Open Data [Online] Available from: http://opendatacommons.org/ [Accessed: 17th July 2014].

REGERINGSKANSLIET / LAGRUMMET. (2014). Kulturmiljölag (1988:950) SFS 2014:694. [Online] Available from: http://www.riksdagen.se/sv/Dokument-Lagar/Lagar/Svenskforfattningssamling/Lag-1988950-om-kulturminnen_sfs-1988-950/ [Accessed: 17th July 2014].

RIKSANTIKVARIEÄMBETET. (2014). Bebyggelseregistret – BeBR. [Online] Available from: http://www.raa.se/hitta-information/bebyggelseregistret/ [Accessed: 17th July 2014].

RIKSANTIKVARIEÄMBETET. (2014b). DAP – digital arkeologisk process. [Online] Available from: http://www.raa.se/dap [Accessed: 17th July 2014].

RIKSANTIKVARIEÄMBETET. (2014c). Fornsök. [Online] Available from: http://www.raa.se/hitta-information/fornsok-fmis/ [Accessed: 17th July 2014].

RIKSANTIKVARIEÄMBETET. (2014d). Lista med lämningstyper och rekommenderad antikvarisk bedömning. Version 4.1 [Online] Available from: http://www.raa.se/app/uploads/2014/07/Lämningstypslistan_ver-4_1_20140626.pdf [Accessed: 17th July 2014].

RIKSANTIKVARIEÄMBETET. (n.d.a). K-samsök – Swedish Open Cultural Heritage (SOCH). [Online] Available from: http://www.ksamsok.se/ [Accessed: 17th July 2014].

RIKSANTIKVARIEÄMBETET. (n.d.b). Kringla. [Online] Available from: http://www.kringla.nu/ [Accessed: 17th July 2014].

RIKSANTIKVARIEÄMBETET. (n.d.c). Riksantikvarieämbetet. [Online] Available from: http://www.raa.se/ [Accessed: 17th July 2014].

RIKSANTIKVARIEÄMBETET. (n.d.d). Samla – Riksantikvarieämbetets öppna arkiv. [Online] Available from: http://samla.raa.se/ [Accessed: 17th July 2014].

SENESCHAL. (2014). Heritage Data: Linked Data Vocabularies for Cultural Heritage. [Online] Available from: http://www.heritagedata.org/ [Accessed: 17th July 2014].

Exploring Time and Space in the Annotation of Museum Catalogues: The Sloane Virtual Exhibition Experience

Stephen STEAD,[1] Dominic OLDMAN,[2] Jonathan Whitson CLOUD[2]

[1]University of Southampton and Paveprime Ltd, UK
[2] British Museum, UK
steads@paveprime.com, doldman@britishmuseum.org, JWhitsonCloud@britishmuseum.org

Abstract
Modern digital museum catalogues differ from historic manual catalogues in that the record boundary is virtual, in the digital and physically on the page, for the manual. The history of the edits in digital systems can only be determined if the change history is explicitly recorded or if periodic snapshots of the data are preserved (although this only gives periodic aggregates of edits). In contrast the pre and post edit states of each edit and annotation are visible on the page in a physical catalogue volume. Card index systems vary in their completeness in this respect, as in some circumstances the card is replaced rather than updated. This has the effect of producing an aggregate of edits similar to a periodic digital snapshot.
Within the catalogue entries there are many possible interpretations of the use of geographical place names. For instance they can refer to the place of manufacture, the region of use, the place of collection, the region of a particular culture or the origin of the manufacturer. Each such interpretation has a different semantic meaning and consequently a different mapping to the CIDOC Conceptual Reference Model (CIDOC-CRM). Similarly dates and other temporal appellations can have many meanings: date of manufacture, date of collection, date of accession into the museum, date of record compilation and/or editing or significant dates in the history or provenance of the object. Again these different meanings have different mappings to the CIDOC-CRM. In addition to these content elements, there are spatial and temporal relationships between the catalogue entries and annotations in historic manual catalogues and between annotations in card index systems. These too have particular mappings to the CIDOC-CRM. This paper addresses these different CIDOC-CRM mappings in the context of catalogues of the Sloane collection.

Keywords: CIDOC-CRM Digital Catalogue, British Museum

1 Historical Background

Forming part of the collection bequeathed to the British Museum by Sir Hans Sloane are the catalogues of the objects (artificial and natural) themselves, mostly handwritten by Sloane himself. Not an inconsiderable undertaking given the size of the collection, particularly for someone maintaining a successful medical practice and also holding public positions, including Secretary and President of the Royal Society, and the Royal College of physicians.

As new acquisitions were added to the collection, and for a period of around 50 years, Sloane maintained the catalogues himself, updating, cross referencing and amending and ever expanding record of his personal collection and passion. The reason for Sloane's diligence would have been threefold; partly as the natural motivation of a fully fledge member of the collecting class who inevitably developed a personal and 'mysterious relationship to ownership'(Benjamin, 1999) but moreover the scientific necessity on which Sloane, in correspondence, summarises by saying that the, 'accurate arrangement of these curiosities constituted my major contribution to the advancement of science' (Clarke, 1980), perhaps an early indication of the different cultures and divisions between the museum and the academy. His diligence perhaps contributed to a critique that he lacked the same level of skills of his contemporaries (Ultee, 1988), but also, more straight forwardly and perhaps also in keeping with a collectors relationship with his/her collection, he didn't trust others to do it. However, other hands have at various stages, contributed or made additions to the catalogues including annotations and references.

Clearly some of the catalogues are brief in their descriptions being more of acquisition registers rather than catalogues. However, the personal touch of Sloane means that they have a particular style and incorporate certain idiosyncrasies or methods of documentation including, for example, an alphabetical system that coded the size of books and folios (De Beer, 1953), a method of coding still employed in the British Museum and other museums (called 'association codes' today). Despite the sometimes brief descriptions the fact that they are handwritten by Sloane, and are a permanent record of the terminology and rationales of the day, mean that they provide a dimension that has been lost in modern day cataloguing. The development of computerised records, which has long passed through its 'hay day' period (an activity that perhaps may never by repeated with the same level of resources), lack a sense of ownership and the more personal touch that paper and pen provide. The computer record of today, although evolving and the result of different hands, including curator, documentation specialist, and intern, often lacks a sense of time and history as old records are overwritten and snapshots or versions are unavailable. Even attribution is often not considered necessary - a situation, with versioning, that we may live to regret as documentation evolves from an inventory to a representation of an institutions collective knowledge and history.

Sloane's catalogues provide us with perspective encapsulated by his language, structure and particular methods of documentation. They provide insight into the relationship that he and others had with the collection, and the society of the time – like today, an age of discovery but also one of exploitation. In the Pitt Rivers museum in Oxford, original labels are still displayed with the objects, again to provide an insight into the Museum's past philosophy and objectives. The process of update, without the opportunity to simply erase, means that Sloane's progress and methods (over 50 years of cataloguing) can be studied in a way that is often lost forever in more modern times.

There are many different reasons for studying Sloane's work. By comparing the techniques and language used by Sloane, we can make comparisons between, not just techniques, but the systems of knowledge that operated in the eighteenth century and today. The different priorities, language and other documentation conventions provide information about the systems of 'rationality' that existed in 18th century society, that Sloane lived and practiced in (Greenhill, 1992, p.9). We may find some interesting differences but we may also find that there are many things that we have in common albeit materialised in different forms and guises. We can determine how the practices of Sloane and his contemporaries influenced the present British Museum and its current culture, and we can possibly determine how far we have come, or how little we have developed. This type of analysis we may not be able to do with other periods of the Museum's documentation history making Sloane's catalogues even more important.

How should they be compared with modern records? Not just an analysis of their structure and extent, but also their approach, perspective and context. Only by using the CIDOC CRM can the modern day record be fully compared to the record of the 18th century – either that, or another 50 years of someone's time.

As the first phase of the preparation of the Sloan catalogues for such analysis a mapping of the content, physical layout and preparation history of the catalogues to the CIDOC CRM provides a stable, reusable information resource that can underpin the analysis. Subsequent phases will provide the mechanism to record the use and interpretation of this elemental mark-up in both successful and unsuccessful inference making.

2 The CIDOC CRM mappings

The mapping of the association between the objects and the places mentioned in catalogue entries produces a number of possible relationships:

A] The place where the object was made: E24 Physical Man-Made Thing P106i was produced by E12 Production P7 took place at E53 Place P87 identified by E44 Place Appellation.

B] The place where the object was used: E24 Physical Man-Made Thing P16i was used for E7 Activity P7

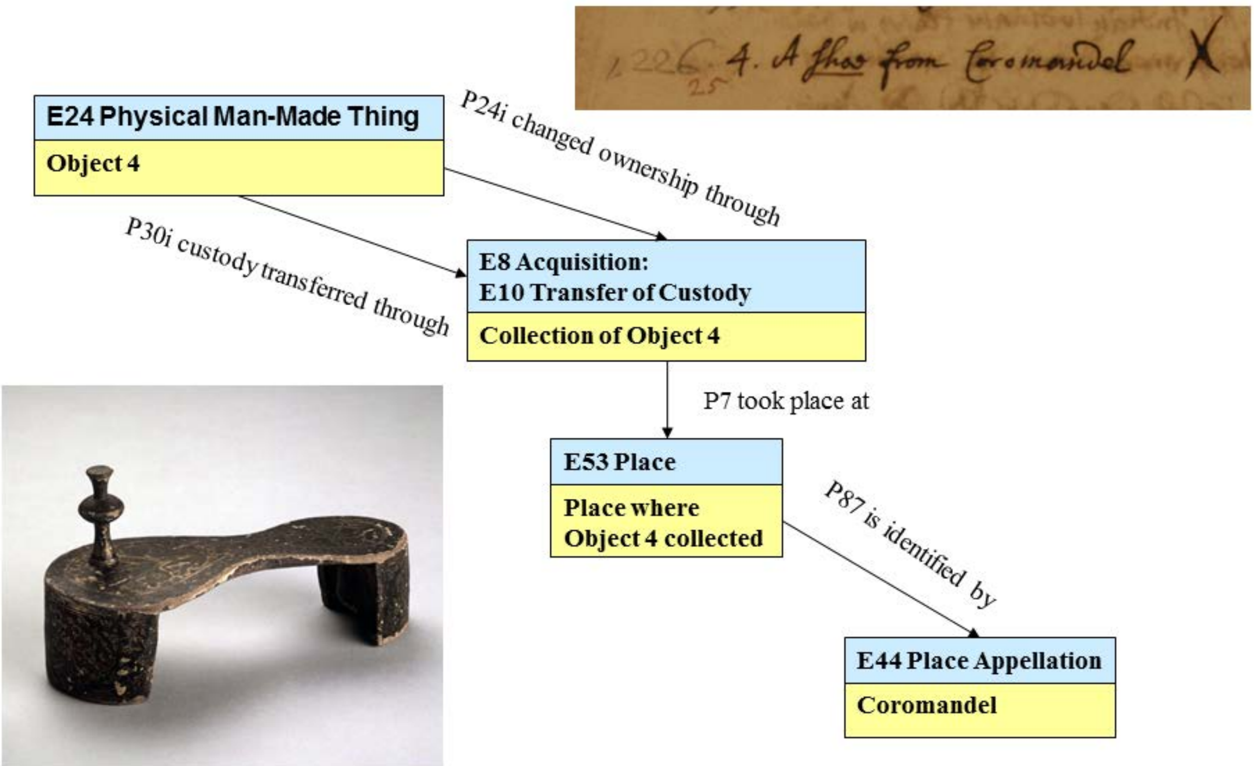


Figure 1: Place of Collection.

took place at E53 Place P87 identified by E44 Place Appellation

C] The place where the object was collected: E24 Physical Man-Made Thing P24i changed ownership through and P30i custody transferred through an event that is multiply instantiated as simultaneously an instance of E8 Acquisition and an instance of E10 Transfer of Custody P7 took place at E53 Place P87 identified by E44 Place Appellation (see Figure 1)

In fact all three of these mappings could be simultaneously true if the object was made, used and collected in the same place.

In other catalogues, entries can refer to the nationality of the manufacturer and this spatial relationship may refer to places that object has never itself been associated with. In this case we would take the spatial relationship to be to the place that the company was formed.

D] E24 Physical Man-Made Thing P106i was produced by E12 Production P14 carried out by E40 Legal Body P95i was formed by E66 Formation P7 took place at E53 Place P87 identified by E44 Place Appellation

A further potential relationship is the one between the object and the spatial extent of the culture that made it. In the CRM the spatial extent is bound to the temporal extent of the culture in an instance E4 Period.

E] E24 Physical Man-Made Thing P106i was produced by E12 Production P10 falls within E4Period P7 took place at E53 Place P87 identified by E44 Place Appellation

In addition the same instance of E4 Period is related to its temporal extent by:

E1] E4Period P4 has time-span E52 Time-Span P78 identified by E49 Time Appellation

Similarly the instance of E12 Production could be linked to time:

A1] E12 Production P4 has time-span E52 Time-Span P78 identified by E49 Time Appellation

As can the E7 Activity representing the use of the object:-

B1] E7 Activity P4 has time-span E52 Time-Span P78 identified by E49 Time Appellation

This construct can also be applied to the simultaneously instantiated E8 Acquisition and E10 Transfer of Custody representing the collection of object:-

C1] E8 Acquisition/E10 Transfer of Custody P4 has time-span E52 Time-Span P78 identified by E49 Time Appellation

In the Sloane Catalogues case there is another E8 Acquisition/E10 Transfer of Custody when the collection was transferred to the newly formed British Museum, which was created specifically to house Sloane's Collection. This can also be referred to as the date of accession.

F] E8 Acquisition/E10 Transfer of Custody [Accession by British Museum] P4 has time-span E52 Time-Span P78 identified by E49 Time Appellation [1753]

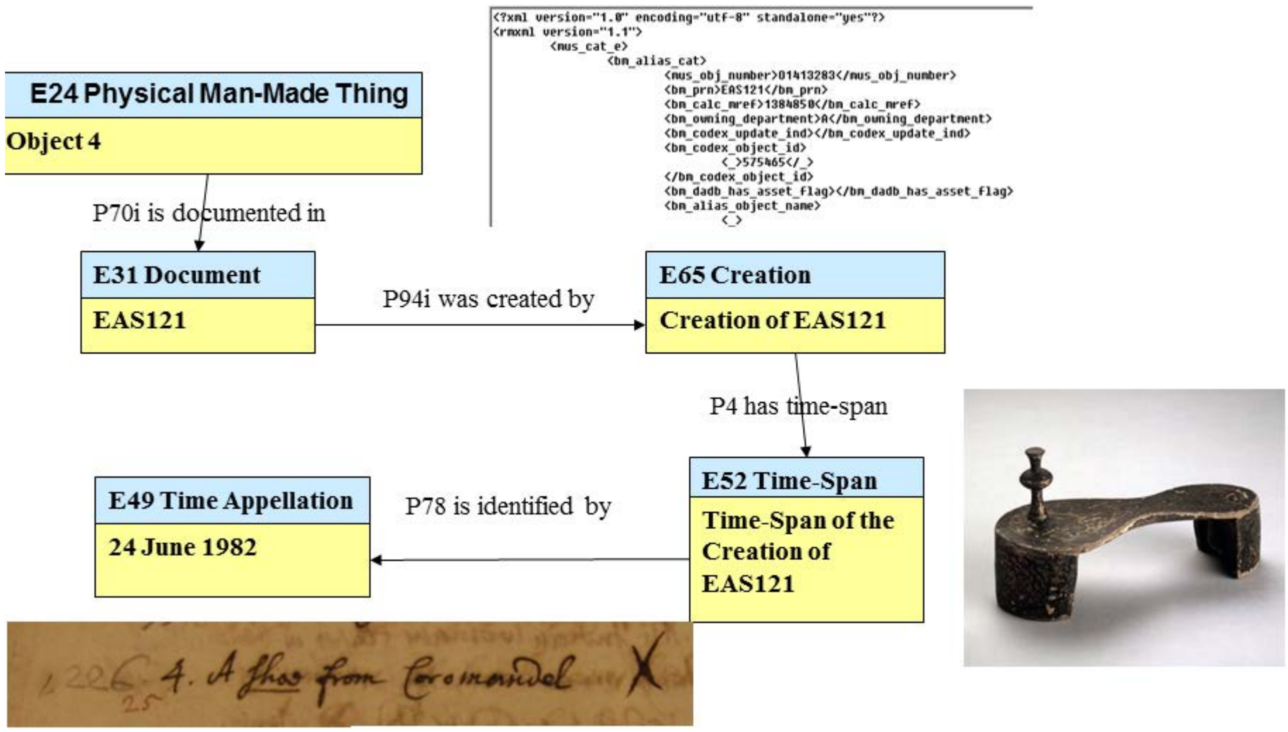


FIGURE 2: DATE OF RECORD COMPILATION.

F1] E8 Acquisition/E10 Transfer of Custody [Accession by British Museum] P23 transferred title from and P28 custody surrendered by E39 Actor P1 is identified by E41 Appellation [Hans Sloane]

F2] E8 Acquisition/E10 Transfer of Custody [Accession by British Museum] P22 transferred title to and P29 custody received by E40 Legal Body P1 is identified by E41 Appellation [British Museum]

The modern computer record for the object also has a date associated with its creation:

G] E24 Physical Man-Made Thing 710i is documented in E31 Document 94i was created by E65 Creation P4 has time-span E52 Time-Span P78 identified by E49 Time Appellation (see Figure 2)

There are also spatial and temporal relationships between the catalogue entries as written in the original Sloane Catalogues. So if two entries are written one after the other in the catalogue it can be inferred that they were created one before the other. These temporal and spatial relationships can be documented in the CRM using the property P120 occurs before (occurs after) for the temporal relationship and the property P122 borders with for the spatial relationship. In figure 3 the temporal and spatial relationships between two catalogue entries for shoes are illustrated.

The rich annotation history of the Sloane Catalogues also produces spatial relationships between the original entries and the annotations and between the annotations. The recording of these relationships allows the documenting of inferences about the sequences of annotations. This inference documentation is part of the ongoing research on the Sloan Catalogue.

In figure 4 the spatial relationships between an original catalogue entry (4. A shoe from Coromandel) and three annotations (an ‘X’ placed after the original entry; the cabinet number ‘L226’ written in the left margin in front of the original entry and the text ‘25’ written in the left margin and tucked under the ‘L226’ and the original entry) are documented. Some of the properties have been omitted for clarity. This will allow inferences about the likely annotation sequence to document the spatial evidence that they used. So for example the inference that the sequence was Original then X then L226 then 25 can point to these spatial relationships as part of the evidence that was used during inference making.

3 Conclusion

Using the CIDOC CRM allows the temporal and spatial relationships that are implicit in the text and annotations of the Sloane Catalogues to be made explicit. This will support good documentation practice and allow the reuse of the documentation by other scholars secure in the knowledge that semantics of the original interpretation have been captured and made available for both human and machine use. Future work will build on this explicit foundation to document the inferences made during the reconstruction of the original content of Sloane’s display cabinets. It is

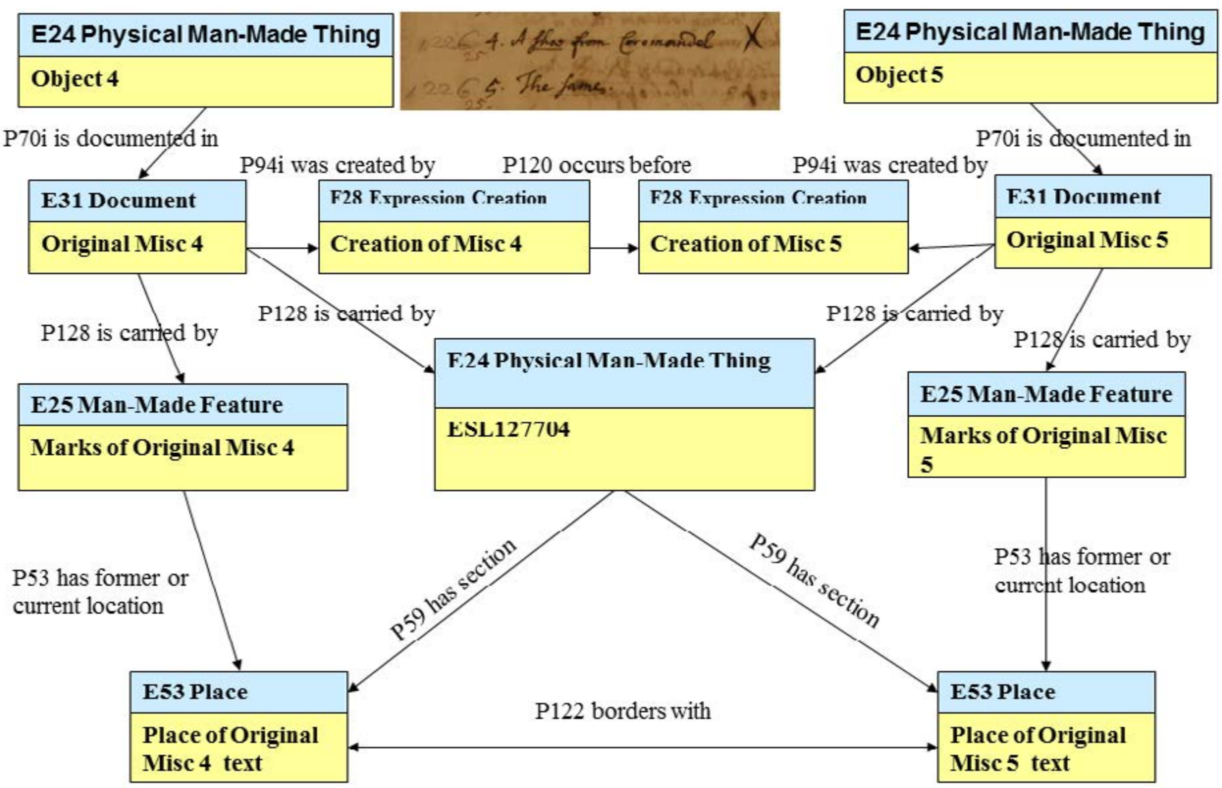


FIGURE 3: RELATIVE RECORD COMPILATION.

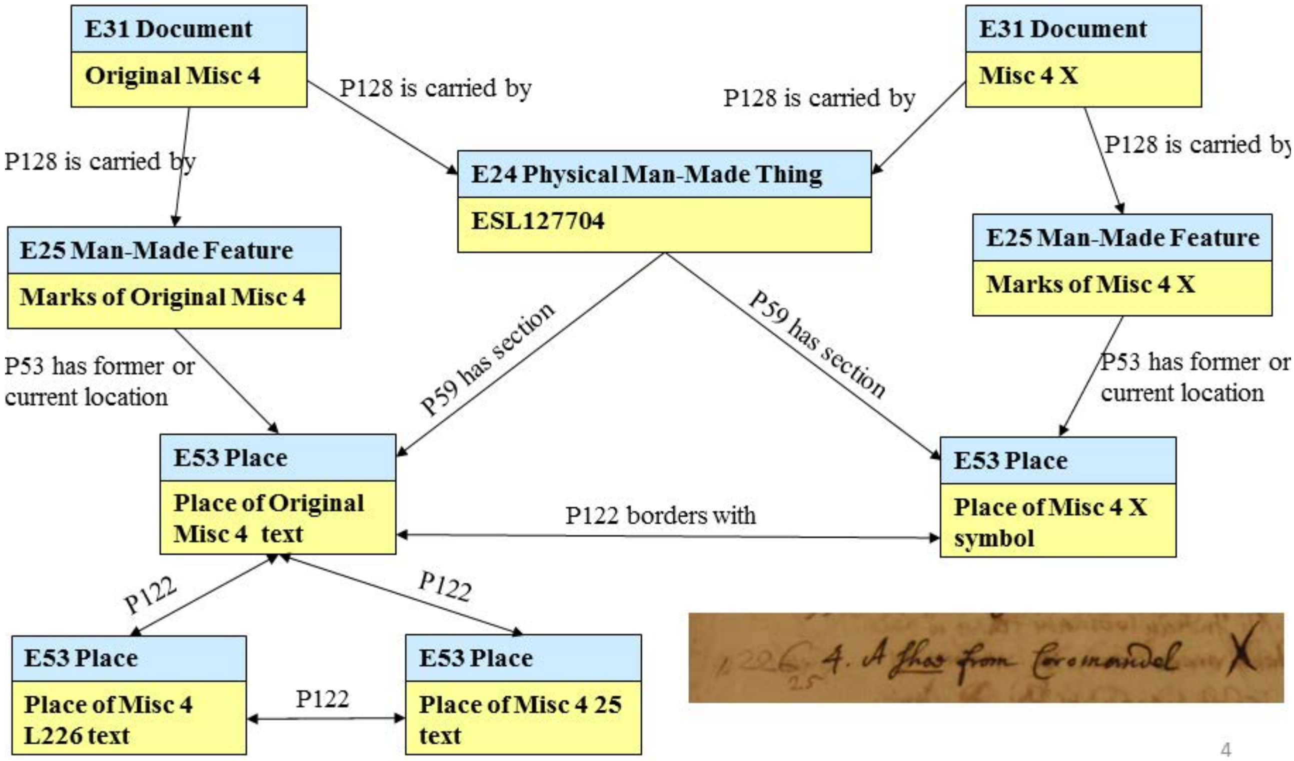


Figure 4: Spatial Relations Between Catalogue Entry Versions.

hoped that exposing such fundamental building blocks will encourage good practice and allow scholars to have confidence in the results of such interpretative projects in the future.

References

DE BEER, G. (1953). *Sir Hans Sloane and the British Museum. Pl. 10.* Oxford University Press: London.

Le Boeuf, P., Doerr, M., Emil Ore, C., Stead, S. (eds). 2013. 'Definition of the CIDOC Conceptual Reference Model, 5.1.2. October 2013.' ICOM/CIDOC CRM Special Interest Group. http://www.cidoc-crm.org/docs/cidoc_crm_version_5.1.2.pdf

BENJAMIN, W. (1999). *'Unpacking My Library: A Talk About Book Collecting', in Illuminations*. London: Pimlico. p. 267.

CLARKE, J. A. (1980). Sir Hans Sloane and Abbé Jean Paul Bignon: Notes on Collection Building in the Eighteenth Century. *The Library Quarterly* 50 (4), pp. 475-482.

GREENHILL, E. H. (1992). *Museums and the Shaping of Knowledge.* New York: Routledge.

ULTEE, M. (1988). Sir Hans Sloane, Scientist. *The British Library Journal*. 1–20.

STANLEY, T., (1994). Papyrus Storage at Princeton University. In The Book and Paper Group Annual 13. Washington, D.C.: American Institute for Conservation of Historic and Artistic Works (AIC).

TAYLOR, J., LEACH, B. & SHARP, H., (2011). The history and conservation of the papyrus of Tuy. *Technical Research Bulletin* 5, pp.95–104.

THEMELIS, P. & TOURATSOGLOU, I., (1997). *Tafoi tou Derveniou.* Athens: TAPA.

VANDERPOOL, E., 1962. News letter from Greece. American Journal of Archaeology, 66(4), pp.389–391. Available at: http://www.jstor.org/stable/502026.

WARE, G. A. *et al.*, (2000). Multispectral document enhancement: ancient carbonized scrolls. IGARSS 2000. IEEE 2000 International Geoscience and Remote Sensing Symposium. Taking the Pulse of the Planet: The Role of Remote Sensing in Managing the Environment. Proceedings (Cat. No.00CH37120), 6, pp.2486–2488. Available at: http://ieeexplore.ieee.org/lpdocs/epic03/wrapper.htm?arnumber=859615.

WRIGHT, M., (1983). A Method of Extracting Papyri from Cartonnage. *Studies in Conservation* 28(3), pp.122–126.

Building Comprehensive Management Systems for Cultural – Historical Information

Chryssoula BEKIARI,[1] Martin DOERR,[1] Dimitris ANGELAKIS,[1] Flora KARAGIANNI[2]

[1] Institute of Computer Science, Foundation for Research and Technology - Hellas (FORTH), Greece
[2] European Centre for Byzantine and Post-Byzantine Monuments (EKBMM), Greece
bekiari@ics.forth.gr, martin@ics.forth.gr, agelakis@ics.forth.gr, karagianni@ekbmm.gr

Abstract

The role typically held in society by museums and cultural institutions has changed in recent years mainly because of the advancements in information technologies and the internet. The new role that emerges for these institutions is that of knowledge organization. The Internet provides them with an opportunity to create new and interesting information services that utilize scholarly work in order to create new forms of information for presenting historical and archaeological research results and for supporting educational activities and entertainment. A new information flow and a new kind of systems emerge. This paper comments on the current situation and describes the needs and the requirements for supporting this new role of the museums and cultural organizations.

Keywords: Information systems, documentation practice, knowledge representation

1. Introduction

In the digital age, a new role is emerging for cultural institutions, that of knowledge organization.

This is due to two main reasons. The first is related to the recent advances in information technology that allow for the production and presentation of born digital material utilizing various new methods and devices. The second is related to the ever greater financial constraints due to reductions of public funding and/or the need to find new financial resources. In trying to compensate for the latter, museums and archives of monuments (also called “SMRs”) have been propelled into a world of economics, which is fundamentally foreign to them (Tobelem, 1997). Not only are they subjected to the rigors of theoretical economic analysis, but they may also have to envisage their operating budget, manpower requirements, commercial returns, visitor numbers, value of acquisitions in narrow financial terms. A new kind of cultural industry is emerging that allows us to view/regard cultural institutions as ‘cultural enterprises’. In the light of the above changes, cultural institutions employ new types of management, actions and activities for exploiting the cultural heritage resources (Low & Doerr, 2010). Many of these actions involve new digital services to attract visitors (general public, specialists, students and children) to the actual museums or monuments, or on the digital space of the institution (portals).

Despite the fact that large budgets have been invested in producing and disseminating new digital material on the web, little care is given about the quality, curation and preservation of this material. In most cases, the newly produced digital material is stored in repositories and is documented using very simplified data models such as the Dublin Core. Also, the majority of museum information systems are not designed to manage and handle the amount of generated data. In most cases, museum systems are designed only for internal use with curatorial notes and administrational data, much like the old card index systems, and are not capable for supporting historical research, knowledge accumulation or the new forms of education and entertainment. At the same time, the openness of the Web and the ease to combine linked data from different sources creates new opportunities to utilize the new data in novel and innovative ways.

This situation imposes many requirements on the quality of data; a new challenge that has to be addressed is the evaluation of the quality of the data retrieved from the Web, that is, the evaluation of the accuracy, timeliness, reliability, and trustworthiness (Gil & Artz, 2007) of the data. A major source for data quality assessment is the analysis of the provenance of knowledge. “Information about provenance constitutes the proof of correctness [...] and [...] determines the quality and amount of trust [...]” (Tan, 2007). Provenance information about a data item is information about the history of the item, starting from its creation and the material evidence or authentic sources it is based on (Pitzalis *et al.,* 2010).

The simplified data models that most cultural institutions use for declaring the metadata, are not capable of storing and maintaining such information. Also, the new information services that are based on such simplified models cannot be expanded or updated and are therefore incapable to follow the progress of the science and the research about the cultural heritage sources. Finally, there are no standard systems to support the information flow and knowledge lifecycle required for the new digital born material and services (Oldman *et al.* 2014).

In the light of this new situation, many issues are emerging relating to the handling of 'documentation' procedures, the assignment of new roles to curators, the engagement of new professional personnel and the acquisition of appropriate IT tools. These issues are discussed in the following sections.

2. The new environment

Under this new framework we distinguish the following procedures:

• *Administration procedure*
This procedure facilitates the management of the actual museum objects and includes activities for identifying, acquiring, keeping, organizing exhibitions, plans for preservation/restoration and funding etc.

• *Scientific documentation procedure*.
This procedure concentrates on collecting and organizing the evidence about the cultural assets, and on describing all the results that are derived both from the actual physical observation of the cultural assets and through historical research. This procedure involves activities for documenting the current state of a cultural asset, the historical events that might have affected it, the historical sources, bibliographic references, records of the excavations, observations, measurements, and any other evidence that support the documentation of the asset.

• *Knowledge exploitation procedure*
This procedure facilitates the exploitation of the results of the scientific documentation procedure and aims at understanding contexts, causes, regularities at different temporal and spatial scales and with different thematic orientation since the outcome of the scientific documentation procedure is usually not appropriate for the general public. Curators, scholars and professionals should formulate scenarios for the digital services, choose themes or create stories that are based on their current research which they wish to reveal, that allow to present things in context and order. Sometimes, additional, related information and new evidence from other knowledge stakeholders and the bibliography are needed in order to validate a scenario or a storyline. The intention of this procedure is to create new information products that will populate new information services. Sometimes this procedure may also enrich the scientific data.

• *Dissemination procedure*
Current technology provides a variety of media, mobile devices, pcs, infokiosks etc. This procedure focuses on publicizing the new information services to target audiences and on facilitating access to the new products through various target media (mobile device, PC etc.), e.g. an information service on the Web that provides digital essays for students or for disable people on a PC or on a smart phone etc.

Some of the people involved in this new environment are described below.

- Registrars that manage the inventory, acquisition, loan-in, loan-out and insurance of the object. They mainly work on the administration procedures.
- Curators that research physical objects and their relevance in context (history). They work on the scientific documentation procedures.
- Conservators that investigate material properties conserve and restore objects. They work on the scientific documentation procedures
- Curators that organize exhibitions and object loans. They record the scientific documentation of the objects and produce exhibition catalogues and academic publications.
- Scientists and scholars from various disciplines, teachers, graphic designers, film directors, and other experts that work on the knowledge exploitation procedures.
- Computer scientists that work on the dissemination procedure and they maintain the information systems, services and products.

New types of information systems are required to support these 'cultural enterprises'. The data structures used in such systems should be based on an ontology that integrates all the information and procedures into a coherent logical model. The employment of ontology is particularly useful for designing suitable data structures, since ontologies provide effective views of the world of a specific domain or ways of projecting reality on information elements and constructs. They provide consistent models of possible states of affairs in some "universe of discourse". The CIDOC – CRM (CIDOC CRM, 2011) is recommended as the most appropriate for documenting cultural entities since it is derived from the underlying semantics of database schemata and document structures used in cultural heritage and museum documentation. It does not define any of the terminology typically appearing as data in the respective data structures; however it foresees the characteristic relationships for its use. It does not aim at proposing new things about what cultural institutions should document; rather it explains the logic of what they actually use to document. It is intended to facilitate the integration, mediation and interchange of heterogeneous cultural heritage information and it aims to be monotonic in the sense of Domain Theory. That is, the existing CRM constructs and the deductions made from them should remain valid and well-formed, even as new constructs are added by extensions to the CRM. Finally the CIDOC CRM is the culmination of more than a decade of standards development work by the International Committee for Documentation (CIDOC) of the International Council of Museums (ICOM) based on a wide sectoral engagement. Its use is now wide spread and various modelling activities in research projects have produced more specific extensions of the CIDOC CRM, relating to the scientific documentation of excavations, of observation events etc.

Another very important characteristic is the use of XML technology. This technology facilitates the overall view of the required data and it has a beneficial effect on the input process. It supports deep, exhaustive and extensible data structures, which can be easily customized to case-specific data entry forms and (multilingual) representational structures. It supports document boundaries which allow the management of jointly created units of knowledge and meta-metadata. The use of XML also supports data exchange with other devices and tools, dynamic exploitation of the web and additional translations of XML documents in other languages.

Furthermore, permanent mappings can be created from an xml document based on an XML Schema which is compliant with an ontology, i.e., sets of equivalence statements between data structure parts and paths in the ontology, which allow for the definition of a deterministic data transformation algorithm from the data structure into an instance of the ontology, e.g., in RDF encoding. The nature of the RDF representation is that of a network. RDF propositions "thrown" into one pool of data integrate automatically via URLs into one large network of "the latest stage of knowledge" of an "aggregation service". Whereas the latter allows for querying combined knowledge across all contents, it is not as effective as XML for handling administrational and knowledge production procedures. We therefore advocate, for the time being, the combined use of both technologies.

Documentation processes take place in all the above mentioned types of documentation procedures. In these 'cultural enterprises' the documentation processes may employ more than one editors or professionals. This is especially true for the exploitation procedure. Three basic stages are distinguished in this type of documentation process. These are:

(a) The editing stage; this is when one or more archaeologists or scientists cooperate in order to collect all the evidence or document their data. In this stage only authorized users (editors) should be able to add or modify the data of a group of documents or in one document.

(b) The stage of requesting publication; this is when archaeologists or scientists request the publishing of their data after having finished the documentation process. The curator or the responsible for a documentation process will check, correct and confirm the output of the documentation process.

(c) The publication stage. The authorities of a cultural institution decide to disseminate to the public the result of a documentation process.

All the above characteristics create requirements which should be satisfied by the new type of cultural information systems. These systems should provide functions for:

a) creating/editing/copying an xml document, request the publication of a document, retract the publication of a document etc.
b) authoring xml documents which follow a schema and facilitate text formatting and internal or external data interlinking
c) supporting a variety of user types. Controlled access should be provided for editorial work on xml document basis.
d) versions administration
e) translation
f) export / import in xml/rdf

3. System design and functions

The Centre for Cultural Informatics of ICS-FORTH has developed a comprehensive management information system in close collaboration with a number of cultural institutions in Greece and abroad.

The design and development of this system started in 2006 in co-operation with the European Center for Byzantine and Post-Byzantine Monuments. This centre is a non-profit organization supervised by the General Secretariat for Culture. It conducts research on the restoration, maintenance and promotion of Byzantine and post-Byzantine monuments at a European and international level and it is active in the Middle East, North Africa, the Balkans and the Black Sea, where there survive monuments that owe their existence to the Byzantine culture and influence.

The goal of the first version of the system was to validate the benefits gained from employing the CIDOC CRM to design the data structures, XML technology, open source native XML databases, Web Services. This first version aimed to facilitate the administration and scientific documentation procedure, to support multiple languages, data exchange and interoperability. Particularly, this version was used for the documentation of the restoration data of Byzantine and Post Byzantine Monuments. The dissemination of this data to the public was done through the "Anne Komnene" site, which was created especially for this reason. This version employed entities in the form of XML schemas compatible to CIDOC CRM and suitable for both the basic and the scientific documentation of the monuments. These entities are:

1. *Museum object*: This entity holds information about the museum object and specifically about (a) the identity of the Object and the XML document itself, (b) description of the present day state of the object (c) historical description, (d) administrative data and finally (e) references.
2. *Monument*: as above but for monuments
3. *Archives* (photos, designs etc.): These entities hold information about photographs, slides, drawings, videos and recordings, documents, cards and catalogues both digitized and on paper.

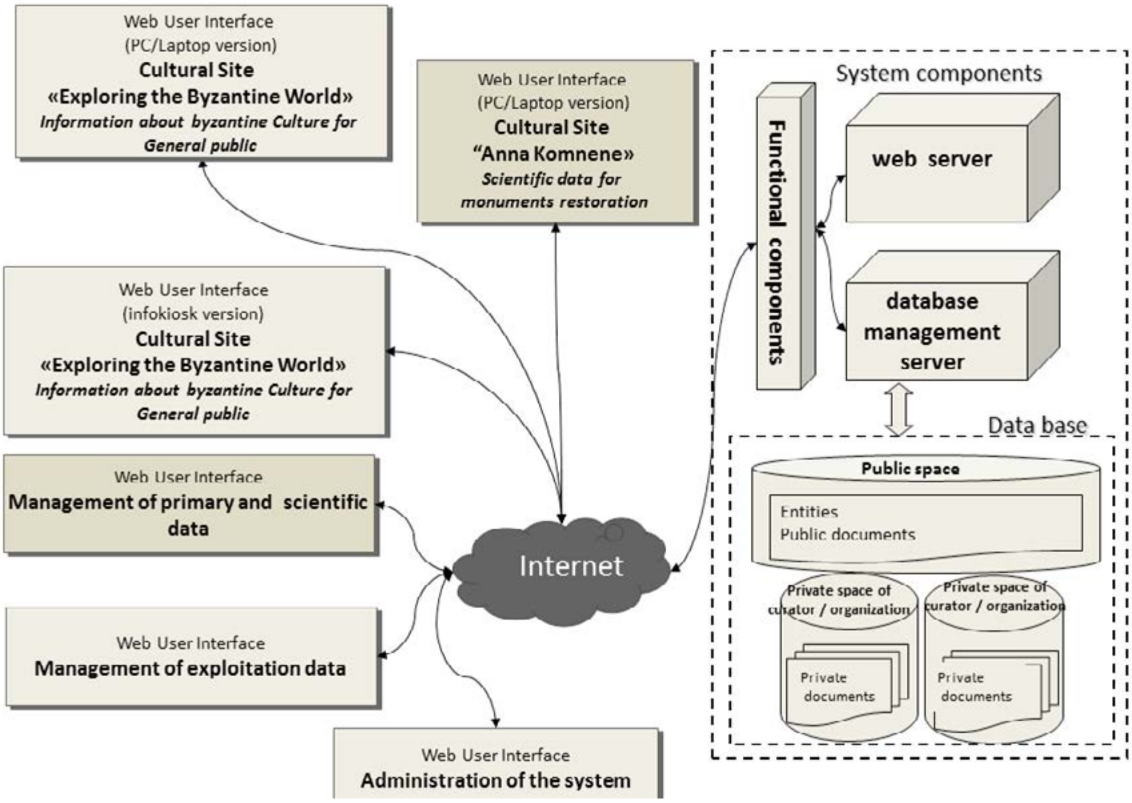


FIGURE 1: THE OUTLINE OF ARCHITECTURE.

4. *Evidences*: It holds information about the evidences (oral or written) referring to the monument or museum object.
5. *Persons*: It holds information about an individual person.
6. *Organizations / Departments*: It holds information about a group, an institution or a department of an institution.
7. *Places*: It holds information about specific places that are somehow related to Objects and Monuments.
8. *Materials*: It holds information about specific materials.
9. *Events*: It holds information about historical events.
10. *Bibliography*: It holds information about published materials and it is compliant with the bibtex standard.

The above XML schemas have many components and repeated structures and/or include links to other entities. A detail account and description of the data model for museum objects and monuments based on the CIDOC CRM is given in (Bekiari *et al.*, 2008).

In 2013, we developed the second version of the system for facilitating the exploitation and dissemination procedures. The goal was to design and implement e-services to inform target audiences (teachers, researchers, schoolchildren and the general public) about monuments, persons, historical events, achievements and aspects of the daily life in Byzantium.

The following figure shows the outline of the resulted infrastructure.

The second version includes new information services for exploiting the historical and architectural research results produced by the European Center for Byzantine and Post-Byzantine Monuments which were entered in the system during the scientific documentation procedure.

These services provide information about the Byzantine Culture in the form of cultural trails, thematic essays, educational activities and podcasts. Using the CIDOC CRM, FRBRoo (FRBRoo, 2012), CRMdig (CRMdig, 2013) as guides, we designed suitable xml data structures (xml schemas) for the above services. The general public can access these services through the "*Exploring the Byzantine World*" site (www.explorinbyzantium.gr). Authorized access is provided for the editors, curators, managers, readers etc by signing in.

The additional entities of the second phase are:

1. *Digital photo*: This entity holds information about born digital material. (maps to *E38 Image from CIDOC CRM* and *D1 Digital Object from CRMdig*)
2. *Information text:* This entity holds the literary text created for presentation or storytelling (maps to *E33 Linguistic Object from CIDOC CRM*). Usually this is part of a Trail or an Essay.
3. *Glossary*: This entity holds the specialized glossary (maps to *E33 Linguistic Object from CIDOC CRM*).

4. *Essay*: This entity holds all the information about a thematic essay (maps to E*33 Linguistic Object from CIDOC CRM* and *D1 Digital Object from CRMdig*)
5. *Trail*: This entity holds all the information about a cultural trail (maps to E*33 Linguistic Object from CIDOC CRM* and *D1 Digital Object from CRMdig*)
6. *Game*: This entity holds all the information about a game and the game itself (maps to *E33 Linguistic Object from CIDOC CRM* and *D1 Digital Object from CRMdig*)
7. *Podcast*: This entity holds all the information about a podcast and the podcast itself (maps to *E33 Linguistic Object from CIDOC CRM* and *D1 Digital Object from CRMdig*)
8. *Comment*: This entity holds users comments about an information object (maps to *E73 Information Object from CIDOC CRM*)

A cultural trail is stored in the database of the system 'Synthesis' as a valid xml document that is described by a certain xml schema. Authorized users may edit or modify the content of a cultural trail using an XML editor. The cultural trail information service provided to the general public through the above site, presents the data of the xml file of a cultural trail. Similarly, the essay service, the educational service and the podcast service present the data of the respective valid xml files that are described by suitable xml schemas. The above mentioned entities employ xml schemas that have many components and repeated structures or include links to other entities. Most of the above entities include links to "Information text" Entity.

The "information text" entity is mainly designed to store literary or scholarly text suitable for dissemination through information services along with the metadata for this kind of text. The basic elements of this entity and their mapping to the above standards are:

(1) ***Identification number***: holds the identification number of the literary text (*CIDOC: E33 Linguistic Object. P1 is identified by (identifies):E42 Identifier*)

(2) ***Title***: The title of the literary text (*CIDOC: E33 Linguistic Object. P102 has title (is title of):E35 Title*)

(3) ***Type***: holds the type of the literary text e.g. "text about a monument of a cultural trail" (*CIDOC: E33 Linguistic Object.P2 has type (is type of):E55 Type*)

(4) ***Short description***: holds a short synopsis of the literary text (*CIDOC: E33 Linguistic Object.P3 has note: E55 Type. P3.1 has type*: "short description")

(5) ***Information text***: holds the formatted literary text (*CIDOC: E33 Linguistic Object.P3 has note: E55 Type. P3.1 has type*: "literary text")

(6) ***Editing details***

(6.1) ***Language***: holds the language of the text, e.g. English etc. (C*IDOC: E33 Linguistic Object.P72 has language (is language of):E56 language*)

(6.2) ***Creation date***: holds the date the text was created (*CIDOC: E33 Linguistic Object.P94 was created by: F28 Expression Creation.P4 has time-span:E52 Time-Span*)

(6.3) ***Author/creator***: holds the name of the author of the text (CIDOC: E33 Linguistic Object. P94 was created by: F28 Expression Creation. P14 carried out by (performed): E39 Actor.P14.1 in the role of "author")

(6.4) ***Last revision date***: holds the date of last revision of the text (*CIDOC-FRBRoo: E33 Linguistic Object.P94 was created by: F28 Expression Creation. P134 continued (was continued by):E7 Activity.P4 has time-span: E52 Time-Span*) and (*E33 Linguistic Object.P94 was created by: F28 Expression Creation. P134 continued (was continued by):E7 Activity.P2 has type:E55 Type{= review}*)

(6.5) ***Reviewer***: holds the name of the reviewer of the text (C*IDOC-FRBRoo: E33 Linguistic Object. P94 was created by: F28 Expression Creation P14 carried out by (performed): E39 Actor. P14.1 in the role of "reviewer"*)

(7) ***General Purpose***: holds information about the target audience type e.g. "for public", "student use" etc. (*CIDOC: E33 Linguistic Object.P103: E55 Type*)

(8) ***Intended use***: holds the type of the device that is most suitable for presenting the text. (*CIDOC:E33 Linguistic Object. P101 had as general use (was use of):E55 Type*)

(9) ***Dimensions***: holds the text size. (*CIDOC: E33 Linguistic Object. P43 has dimension (is dimension of):E54 Dimension*). These are:

9.1) ***Unit name***: holds the name of the measured unit, e.g. "words"(*CIDOC:E33 Linguistic Object. P43 has dimension (is dimension of):E54 Dimension. P91 has unit: E58 Measurement Unit*)

(9.2) ***Value range***: holds the count of the measured unit, e.g. 500 (*CIDOC:E33 Linguistic Object. P43 has dimension (is dimension of):E54 Dimension. P92 has value: E60 Number*)

(10) ***Components***: holds the links for the visual or literary supplementary components of a literary text. (*CIDOC: E33 Linguistic Object. P148 has component (is component of): E33 Linguistic Object*) & (*CIDOC: E33 Linguistic Object .P106 is composed of (forms part of): E33 Linguistic Object*)

Besides the link, the following information is stored for each component:

(10.1) ***Component type***: Stores the type of the component, e.g. “Digital Photo”. (*CIDOC:E33 Linguistic Object.P2 has type (is type of):E55 Type*)

(10.2) ***Component title***: Stores the title of the component (*CIDOC:E33 Linguistic Object.P102 has title*)

(11) ***Relations with other entities***: Stores links to other entities, especially to those created and populated during the scientific documentation procedure and are related to this particular literary entity, either as proof of knowledge provenance or motivation or as the outcome of a scientific observation event etc. Besides the link to a specific entity, the following information is also stored:

(11.1) ***Comment***: Stores the justification of the relation (*CIDOC-CRMdig: E33 Linguistic Object. L44 (is extracted from): D32 Knowledge Extraction*)

(11.2) ***Relation type***: Stores the type of the relation with the specific Information text Entity. (*CIDOC-CRMdig: E33 Linguistic Object. L44 (is extracted from): D32 Knowledge Extraction. P101 had as general use (was use of) :E55 Type*)

(11.3) ***Identification number***: Stores the identifier of the related entity. (*CIDOC-CRMdig: E33 Linguistic Object. L44 (is extracted from): D32 Knowledge Extraction. P67 refers to (is referred to by):E1 CRM Entity. P1 is identified by (identifies): E41 Appellation*)

(11.4) ***Title of entity***: Stores the title of the related entity. (*CIDOC-CRMdig: E33 Linguistic Object. L44 (is extracted from): D32 Knowledge Extraction. P67 refers to (is referred to by):E1 CRM Entity. P102 has title (is title of):E35 Title*)

(12) ***Legal status / intellectual property rights***: Stores the intellectual property rights and the name of the rights' owner. (*CIDOC: E33 Linguistic Object. P104 is subject to (applies to): E30 Right) & (CIDOC: E33 Linguistic Object .P105 right held by (has right on):E39 Actor*)

(13) ***Bibliographic references***: Stores links to bibliographic references. (*CIDOC: E31 Document. P70 documents (is documented in): E1 CRM Entity*)

A curator wishing to add more literary or scholarly texts for dissemination to the information services has only to create the appropriate information text entities.

A Cultural trail is modelled as a CIDOC CRM *E33 Linguistic Object* which is composed of other linguistic objects. Its structure consists of repeated structures of Linguistic Objects that represent the stops of a cultural trail and the monuments of each stop. A cultural trail may have many versions in the form of valid xml documents depending on the type of the devise used for presenting it, for example a version for a PC/Laptop, another for Infokiok, and yet another version that it is accessible by all. The elements of a cultural trail and their mapping to the above standards are the following:

(1) ***Identification number***: The Identification number assigned to a cultural trail (*CIDOC:E33 Linguistic Object. P1 is identified by (identifies):E42 Identifier*)

(2) ***Title***: The title of the cultural trail (*CIDOC: E33 Linguistic Object. P102 has title (is title of):E35 Title*)

(3) ***Audience level:*** This tag is used for storing information about the knowledge level of the audience that this cultural trail is appropriate for. (*CIDOC: E33 Linguistic Object. P103 was intended for (was intention of): E55 Type*)

(4) ***Accessible through***: This tag holds information about the device type through which this version of cultural trail is accessible. (*CIDOC-CRMdig: E33 Linguistic Object. P128 is carried by: D8 Digital Device*)

(5) ***Language***: This tag holds the language in which the literary text is written (*CIDOC:E33 Linguistic Object. P72 has language (is language of):E56 language*)

(6) ***Responsible***: This element holds the information about the person or organization which is responsible for documenting the cultural trail. (*CIDOC: E33 Linguistic Object.P94 was P105 right held by (has right on)): E39 Actor*)

(7) ***Creation date***: This element holds information about the creation date of the literary text. (*CIDOC: E33 Linguistic Object.P94 was created by: F28 Expression Creation.P4 has time-span:E52 Time-Span*)

(8) ***Status***: This tag holds information about the status of the literary text in relation to its publication ie. Published /un published/ under publication etc. (*CIDOC: E33 Linguistic Object.P2 has type:E55 Type*)

(9) ***Characteristic icon***: This field stores a representative thumbnail for representing this cultural trail on the web. (*CIDOC-CRMdig: E33 Linguistic Object. P67 is referred to by: D39 Thumbnail*)

(10) ***Short description***: This field holds a short description text about the cultural trail. (*CIDOC:E33 Linguistic Object.P3 has note: E55 Type. P3.1 has type: “short description”*)

(11) ***Description***: This field holds the link to a literary text for the specific cultural trail. Specifically it contains the following fields: (*CIDOC: E33 Linguistic Object. P148 has component (is component of): E33 Linguistic Object*)

(11.1) ***Identification number of the information text***: The identification of the Information text entity which describes the particular literary text. (*CIDOC: E33 Linguistic Object. P1 is identified by (identifies):E42 Identifier*)

(11.2) ***Title of the information text***: The title of the Information text entity which describes the particular literary text (*CIDOC: E33 Linguistic Object.P102 has title*)

(12) ***Stop***: This is a repeated structure that holds information about a tourist or historic site (town, castle etc.) someone might visit or stop, as well as all the monuments located in the nearby area. It links an information text entity that describes the stop to the specific trail (CIDOC: E33 Linguistic Object. P148 has component (is component of): E33 Linguistic Object)

It contains the following sub elements:

(12.1) ***Identification number of the information text*** ((*CIDOC: E33 Linguistic Object. P148 has component (is component of): E33 Linguistic Object. P1 is identified by (identifies):E42 Identifier*)

(12.2) ***Title of the information text*** ((*CIDOC: E33 Linguistic Object. P148 has component (is component of): E33 Linguistic Object. P102 has title (is title of):E35 Title*)

(12.3) ***Representative thumbnail for the stop***: (*CIDOC: E33 Linguistic Object. P148 has component (is component of): E33 Linguistic Object.P138 has representation: E36 Visual Item*)

(12.4) ***Monument***: It is a repeated structure used for describing the monuments of a stop. For each monument there is a link to an information text entity which contains the literary text for the monument (*E33 Linguistic Object. P148 has component (is component of): E33 Linguistic Object*)

(12.4.1) ***Identification number of the literary text:*** (*E33 Linguistic Object. P148 has component (is component of): E33 Linguistic Object). P1 is identified by (identifies):E42 Identifier*)

(12.4.2) ***Characteristic icon (thumbnail)***: (*E33 Linguistic Object. P148 has component (is component of): E33 Linguistic Object). P138 has representation: E36 Visual Item*)

(12.4.3) ***Short description***: (*E33 Linguistic Object. P3 has note: E55 Type. P3.1 has type: "short description"*)

(12.4.4) ***Title of the literary text***: ((*E33 Linguistic Object. P148 has component (is component of): E33 Linguistic Object). P102 has title (is title of):E35 Title*)

An editor wishing to add stops to a specific trail has only to repeat the above "stop" sub element as many times as needed. Respectively, adding monuments to a particular stop is done by adding "monument" sub element as many times as needed. Games, podcasts, essays have been modelled as xsd schemas in the same way.

Finally the system "synthesis" supports a generic workflow for documenting cultural entities and provides functions for creating, editing, navigating and retrieving documents, for data migration, for document translation, for import from and export to xml/RDF and for associating documents of cultural entities with documents of additional entities like events, persons, organizations, places, materials, photos etc.

4. Conclusions

The development and implementation of version 2.0 of the system were done in the space of a month, while the information services of "*Exploring the Byzantine World*" were completed in two months. The development work mainly comprised the addition of new xml schemas for describing the appropriate XML documents for exploiting the scientific knowledge. Also, some marginal fields of the existed schemas were added or changed, but the database design and function remained mainly unaffected by these changes.

During the exploitation procedure a number of actions took place including: bibliography selection, creation of literary or scholarly texts for exploiting the research results about the byzantine culture, text translations, curation of photographs, shooting of new digital photographs, creation of scenarios for educational activities, development of graphical designs, correlation of documented events with monuments and museum objects etc. These actions entailed the cooperation of a number of experts from different disciplines such as archaeologists, byzantinologists, museologists and museum educators, linguists/ editors of scientific texts, graphic and project designers. The information services, developed for the "*Exploring the Byzantine World*", are disseminated through a web that is accessible by PCs/Laptops and through infokiosks:

(http://www.exploringbyzantium.gr/EKBMM/Page?lang=en&template=infokiosk).

The content provided by cultural trails and thematic essays can be downloaded in the form of a "pdf" file for off line access. An "accessible version" is also provided by the information services.

In conclusion we may say that the adoption of this new kind of information system by cultural institutions resulted in the following:

(a) Integration, mediation, adaptation and interchange of heterogeneous information became possible without loss of information

(b) New digital material (texts, photographs, etc.) was created that has been well documented, is manageable and compliant with international standards meaning that this material can very easily be reused and that the requirements about data provenance have been satisfied.

(d) Existing data can be updated and new added to the information services by the users themselves.

(e) The dissemination of data in linked open repositories is trivial

(f) The creation of new information services is very fast

Acknowledgements

Synthesis 2.0 is the outcome of the collaboration of the authors of this paper with many other members of the Centre of Cultural Informatics of Information Systems Laboratory. These are: Lida Charami, Korina Doerr, Christos Georgis, Konstantina Konsolaki, Athina Kritsotaki, Eirini Maravelia, Giorgos Samaritakis, Ionas Perakis.

Also this work was supported by "Digital Convergence" Operational Programme and Regional Operational Programmes for Macedonia & Thrace, Crete & the Aegean Islands, Thessaly, Central Greece & Epirus, and Attica, Co-funded by Greece and the European Union

References

BEKIARI, C. CHARAMI, L. DOERR, M. GEORGIS, C. KRITSOTAKI, A. (2008). *Documenting Cultural Heritage in Small Museums. in Proceedings of the Annual Conference of CIDOC*, Athens, Greece, (September)

CIDOC CRM (2011), *Current Official Version of the CIDOC CRM. Definition of the CIDOC Conceptual Reference Model. The version 5.0.4 of the reference document*. [online]. Available from: http://www.cidoc-crm.org/docs/cidoc_crm_version_5.0.4.pdf [Accessed, 30th July 2014]

CRMdig (2013), CRMdig An Extension of CIDOC-CRM to s*upport provenance metadata, version 3.0*, [online]. Available from http://www.ics.forth.gr/isl/CRMext/CRMdig/docs/CRMdig3.0.pdf [Accessed 30th July 2014]

FRBRoo (2012), FRBRoo Model version 1.0.2[online]. Available from http://www.cidoc-crm.org/docs/frbr_oo/frbr_docs/FRBRoo_V1.0.2.pdf [Accessed, 30th July 2014]

GIL, Y. & ARTZ, D. (2007) Towards Content Trust of Web Resources. Journal of Web Semantics. 5(4). (December). p. 227–239

LOW, J.T., & DOERR, M. (2010). *A Postcard is Not a Building - Why we Need Museum Information Curators: In Proc. of the CIDOC 2010 Conference: Museums in intercultural dialogue - New Practices in Knowledge Sharing and Information Integration*. Shangai, China, (November)

OLDMAN, D. DOERR, D. DE JONG, G. NORTON, B. WIKMAN, T. (2014) *Realizing Lessons of the Last 20 Years: A Manifesto for Data Provisioning & Aggregation Services for the Digital Humanities*. D-Lib Magazine 20(7/8) (July/August), doi:10.1045/july2014-oldman, [online]. Available from : http://www.dlib.org/dlib/july14/oldman/07oldman.html [Accessed: 30th July 2014]

PITZALIS, D. NICCOLUCCI, F., THEODORIDOU, M. DOERR, (2010). *LIDO and CRMdig from a 3D Cultural Heritage Documentation Perspective. Proceedings of VAST 2010: The 11th International Symposium on Virtual Reality, Archaeology and Cultural Heritage*, Eurographics, Palais du Louvre, Paris, France, 21-24 (September) p. 87-95. Aire-La-Ville, Geneva, Switzerland: Eurographics.

TAN, W.C. (2007) Provenance in Databases: Past, Current, and Future. IEEE Data Engineering Bulletin.30 (4). (December). pp. 3–12

TOBELEM J.. (1997), The Marketing Approach in Museums, *Museum Management and Curatorship*. 16 (4). pp. 337–354.

Managing Time Dimension in the Archaeological Urban Information System of the Historical Heritage of Rome and Verona

Alberto BELUSSI,[1] Sara MIGLIORINI,[2] Piergiovanna GROSSI[1]

[1] University of Verona, Department of Computer Science, Verona, Italy
[2] University of Verona, Department of Time Space Image Society (TeSIS), Verona, Italy
alberto.belussi@univr.it, sara.migliorini@univr.it, piergiovanna.grossi@univr.it

Abstract
Time and space are two important characteristics of archaeological data. As regards to the first aspect, in literature many time dimensions for archaeology have been defined which extend from the excavation time, to the dating of archaeological objects. Standard ISO 19018 describes temporal characteristics of geographical information in terms of both geometric and topological primitives. The first aim of this paper is to analyze the applicability of such Standard for representing archaeological data, referring to the model adopted by the cities of Verona and Roma (Italy) as case study. However, since archaeological dates are often subjective, estimated and imprecise, one of the main lacks in the Standard is the inability to incorporate such vagueness in date representation. Therefore, the second contribution of this paper is the extension of the Standard in order to represent fuzzy dates and fuzzy relationships among them.

Keywords: time, vagueness, archaeology, dating

1. Introduction

Geographical Information Systems (GISs) are usually adopted for managing archaeological data, since one of their main characteristics is an absolute or relative location in 3D space. This information allows to derive important relations between findings, in particular concerning stratigraphic analysis. Besides to spatial location, temporal dimension is of considerable interest for archaeological research. For this reason, some attempts can be found in literature in order to define a 4D GIS tailored for archaeological data (De Roo *et al.*, 2014), where the fourth dimension is time.

Standard ISO 19108 (ISO, 2002) describes the temporal characteristics of geographical information. In particular, it applies the concepts of geometry and topology, which are typical of the spatial domain, to the description of temporal aspects. The main observation is that a point in time occupies a position in a temporal reference system and can be connected with other points through ordering relations. Topological structures can be used to explicitly describe relations between time points, even when they cannot be directly derived. These structures are particularly useful in the archaeological domain, where relative time relations between objects are frequently better known than their exact location in time. For all these reasons, the first contribution of the paper is an evaluation of the Standard's applicability for modeling archaeological data (Sec. 4). This evaluation is done by considering two existing information systems, called SITAR (Sistema Informativo Territoriale Archeologico di Roma) and SITAVR (Sistema Informativo Territoriale Archeologico di Verona), which collect and manage the archaeological data of Rome and Verona (Italy), respectively (Basso *et al.*, 2013). SITAVR development started in 2011 through our collaboration with the Archaeological Agency of Veneto Region and a cooperation agreement with the Archaeological Special Agency of Rome, which was developing its information system SITAR for the Italian capital since 2007. In the following the term SITA*R is used to denote both information systems.

From this preliminary evaluation has emerged that the time dimensions of archaeological data are typically vague. Due to this inherent vagueness, many dates are wrongly described as periods instead of instants with the aim to provide a possibility interval for its value. For instance, the construction date of a building can be located with more confidence between 1830 and 1850, plus/minus 10 years. This kind of date specification suggests the use of a fuzzy approach for representing time. Moreover, also the ordering relations between time points can incorporate a certain degree of possibility. Therefore, the second contribution of the paper is the extension of the Standard with fuzzy types and relations, and the application of them to the SITA*R model.

The overall aim of this paper is to propose a model based on existing standards and consolidated reasoning techniques, for representing temporal dimensions of archaeological data. Such framework can become an invaluable help for archaeologists during the dating and interpretation processes, and can be applied in other contexts with similar characteristics, such as geology.

The remainder of the paper is organized as follows: Sec. 2 summarizes some results about the representation of time in archaeology. Sec. 3 presents the Standard ISO 19108, while Sec. 4 applies its concepts to the modeling of the SITA*R information system. Sec. 5 describes how the Standard can be extended in order to incorporate fuzzy time dimensions; such extended concepts are used in Sec. 6 for representing SITA*R concepts. Finally, Sec. 6 summarizes the obtained results and discusses future.

2. Related Work

A first investigation about the applicability of Standard ISO 19108 for the representation of archaeological data is proposed in (De Roo et al., 2014). The authors conclude that the standard can be successfully applied in this context, but they also highlight the lack of constructs for describing the inherent vagueness of such data.

In (Katsianis et al, 2008) the authors identify six potential time categories for archaeological finds which includes: excavation time, database time, stratigraphic time, archaeological time, site phase time and absolute time. The SITA*R model includes many of this time categories. In particular, it includes the excavation time, the stratigraphic time (in terms of relative temporal positions between finds), the archaeological time (e.g. Roman Time or Middle Age), the site phase time (i.e. the distinction of different phases during an object life), and the absolute time.

In (Li et al, 2002) the authors discuss the possibility of incorporating a fuzzy approach into a particular spatio-temporal processing framework in which temporal information is stored through a series of snapshots associated to particular instants in time and relationships regard the relative ordering among events. In this framework spatial objects are temporally located into a specific time layer (snapshot) associated to a particular instant in time. The authors define the concept of fuzzy time layer which is an imprecise time interval within initial and final time points and possibility distribution functions. The proposed model is applied to a wildlife migration modeling analysis.

3. Standard ISO 19108

Standard ISO 19108 (ISO, 2002) describes the temporal characteristics of geographical information.

The schema consists of two packages: *Temporal Objects* and *Temporal Reference System*. Package *Temporal Objects* defines temporal geometric and topological primitives that shall be used as values for the temporal characteristics of features and datasets. The temporal position of a primitive shall be specified in relation to a temporal reference system. For this purpose, package *Temporal Reference* System provides elements for describing temporal reference systems.

Package *Temporal Objects* is illustrated in Fig.1. It includes primitive and complex objects: *TM_Primitive* is an abstract class that represents a non-decomposable element of time geometry (*TM_GeometricPrimitive*) or topology (*TM_TopologicalPrimitive*), while *TM_TopologicalComplex* is an aggregation of connected *TM_TopologicalPrimitives*. Similarly to the corresponding spatial concepts, *TM_GeometricPrimitive* provides information about temporal positions, while *TM_TopologicalPrimitive* provides information about connectivity in time. Both *TM_Primitives* implement the interfaces *TM_Order* and *TM_Separation*: the first one provides an operation for determining the relative position of a primitive with

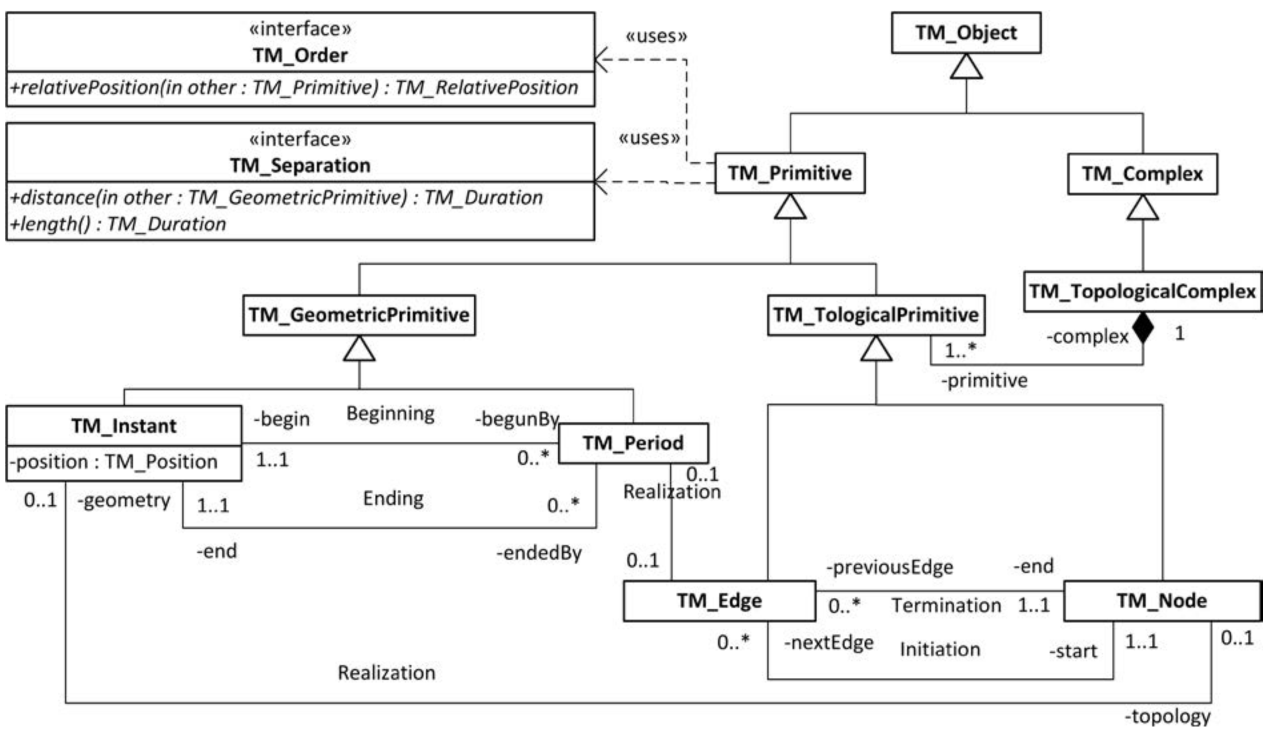


FIGURE 1: PACKAGE TEMPORAL OBJECT OF THE STANDARD ISO 19108.

Relation	Types	Condition
a before b	any	$\exists$ a sequence $S \in \mathcal{C}$ such that: a is earlier than b in S $\wedge$ $(a,b) \notin \mathcal{I} \wedge (b,a) \notin \mathcal{I} \wedge$ $(a,b) \notin \mathcal{T} \wedge (b,a) \notin \mathcal{T}$.
a meets b	(E, E)	$\exists n \in \mathrm{N}((a,n) \in \mathcal{T} \wedge (n,b) \in \mathcal{I})$
a begins b	(N, E)	$(a,b) \in \mathcal{I}$
a begunBy b	(E, N)	$(b,a) \in \mathcal{I}$
a equals b	any	a and b are the same primitive
a ends b	(N, E)	$(a,b) \in \mathcal{T}$
a endedBy b	(E, N)	$(b,a) \in \mathcal{T}$
a metBy b	(E, E)	$\exists n \in \mathrm{N}((n,a) \in \mathcal{I} \wedge (n,b) \in \mathcal{T})$
a after b	any	$\exists$ a sequence $S \in \mathcal{C}$ such that: a is later than b in S $\wedge$ $(a,b) \notin \mathcal{I} \wedge (b,a) \notin \mathcal{I} \wedge$ $(a,b) \notin \mathcal{T} \wedge (b,a) \notin \mathcal{T}$

TABLE 1: ALLEN'S TEMPORAL RELATIONS THAT CAN BE DERIVED FROM THE STRUCTURE OF A TOPOLOGICAL COMPLEX C. IN THE TABLE I IS THE SET OF INITIATION ASSOCIATIONS AND T IS THE SET OF TERMINATION ASSOCIATION IN THE TOPOLOGICAL COMPLEX. IN THE SEQUEL, (A, B) E I STANDS FOR A E TM_NODE AND B E TM_EDGE AND THERE EXISTS AN INITIATION ASSOCIATION BETWEEN THEM. SIMILARLY FOR (A, B) E T.

respect to another one, while the second interface provides operations for computing the length (duration) of a primitive and the distance from another one.

In the temporal context there are two geometric primitives: instant (*TM_Instant*) and period (*TM_Period*), and two corresponding topological primitives: node (*TM_Node*) and edge (*TM_Edge*) which can be realized as an instant and a period respectively. Each edge starts and ends in nodes, while a node can also exists without being associated with edges. When a node has a realization on the time axis as instant, its temporal position is determined, otherwise it can be qualitative described by means of the temporal relations represented by the edges that starts and ends in the node. When an edge has a realization on the time axis as period, its temporal position is determined, otherwise it simply represents a temporal relation between two nodes and its corresponding period can be only qualitative described by means of its start and end nodes.

In order to deal with a connected set of nodes and edges, the *TM_TopologicalComplex* class has been introduced. A topological complex is a set of connected topological primitives. Each edge of a topological complex has its start and end nodes inside the complex. It can be represented as a graph in which a set of *TM_Primitives* are contained where the above described constraint on edges is satisfied. A *TM_TopologicalComplex* allows to compactly represent

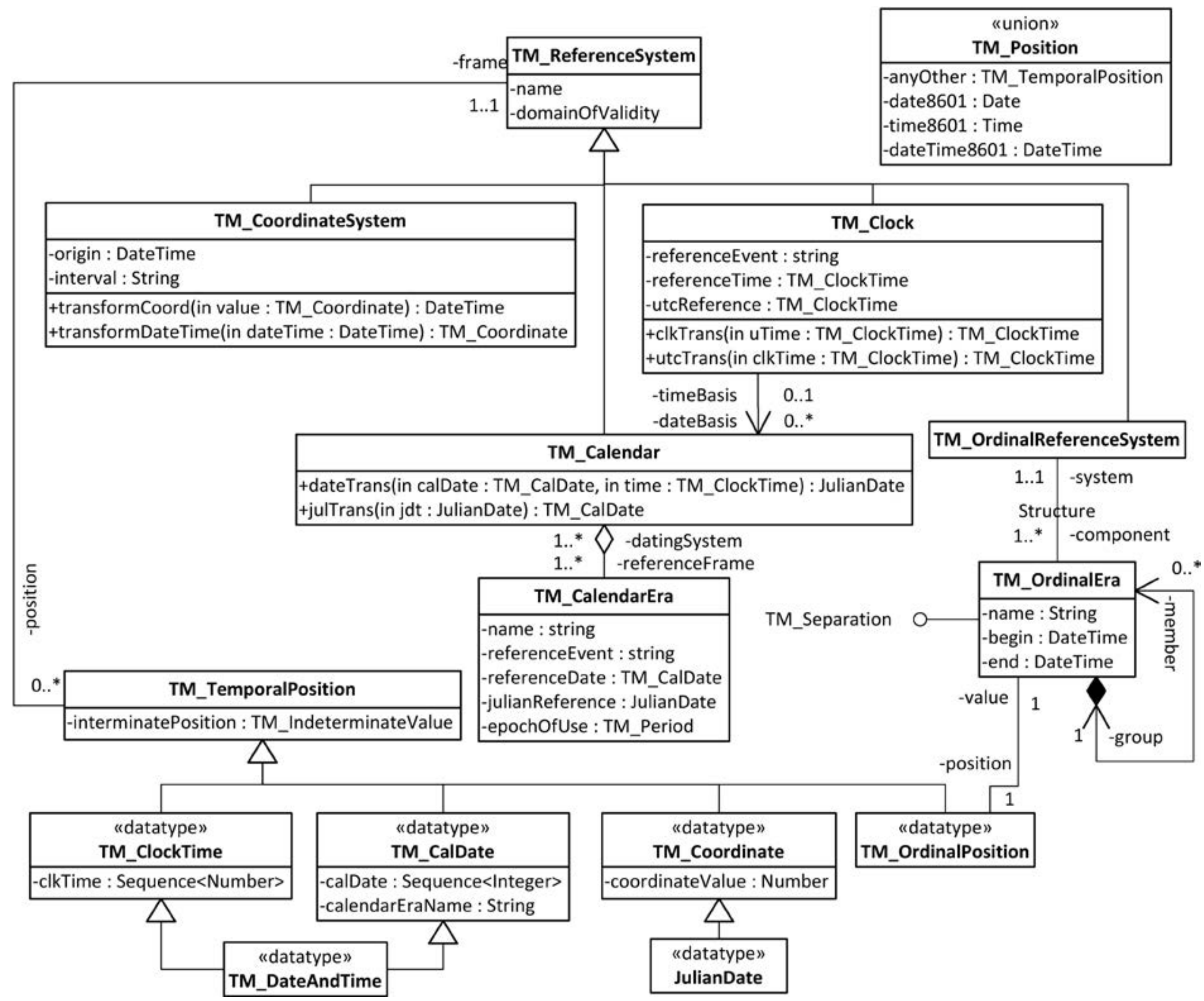


FIGURE 2: PACKAGE REFERENCE SYSTEMS OF THE STANDARD ISO 19108.

relations among objects. In particular, Allen's relations (Allen, 1983) can be derived as illustrated in Table 1.

Package *Temporal Reference Systems* is illustrated in Fig. 2. Standard ISO 8601 (ISO, 2004) specifies the use of the Gregorian Calendar and 24 hour local, or Coordinate Universal Time (UTC) for information interchange. This last one shall be the primary temporal reference system for geographical information. Anyway, when more than one temporal reference system is used in a single feature catalogue, application schema, or dataset, the definition of each temporal characteristic shall identify the temporal reference system that is used. Package *Temporal Reference Systems* includes three common types of temporal reference systems: calendars (used with clocks for greater resolution), temporal coordinate systems, and ordinal temporal reference systems.

A *TM_Calendar* is a discrete temporal reference system that provides a basis for defining temporal position with a resolution up to one day. A calendar has a hierarchical structure in which a specific type of time interval is used at each level. In other words, the number and types of temporal granularities provided by a calendar depends on the number and types of provided hierarchical levels. A *TM_Clock* can be used with a calendar in order to provide a complete description of a temporal position within a specific day.

A *TM_CoordinateSystem* is a temporal coordinate system based on a continuous interval scale defined in terms of a single time interval: all dates are defined as a multiple of the standard interval associated with the reference system and with respect to a chosen origin. It eases the computation of distances between points and the description of temporal operations which can be complicated when temporal positions are described in terms of calendar dates and times in a day.

A *TM_OrdinalReferenceSystem* is based on an ordinal scale. In its simplest form, it is an ordered series of events. It is particularly appropriate in a number of applications of geographic information (e.g., archaeology) in which relative position in time is known more precisely than duration. In such applications, the order of events in time can be well established, but the magnitude of the intervals between them cannot be accurately determined. An ordinal temporal reference system consists of a set of *TM_OrdinalEras*. They can be often hierarchically structured such that an ordinal era at a given level of the hierarchy includes a sequence of coterminous shorter ordinal era.

A *TM_Position* is a union class that consists of a set of data types. In particular, it includes a *TM_TemporalPosition* that can be used when it is necessary to explicitly define the adopted reference system.

A *TM_CalDate* is a data type that shall be used to identify temporal position within a calendar. It has a fundamental property *calDate* which provides a sequence of positive integers where: the first one identifies a specific instance of the unit at the highest level of the calendar hierarchy, the second one identifies a specific instance of the unit used at the next lower level in the hierarchy, and so on. The format defined by Standard ISO 8601(ISO, 2004) for dates in Gregorian calendar may be used for any date that is composed of values for year, month and day. In particular, besides to the complete format *YYYY-MM-DD*, the standard allows to specify dates at reduced precision: *YYYY-MM* which refers to a particular month of a year, *YYYY* which refers to a particular year, and *YY* which refers to a particular century (e.g., 19 stands for the century from 1900 to 1999). To represent years before 0000, the standard also permits an expanded year representation [±*YYYY*] which uses an extra digits beyond the four-digit minimum: each year must be prefixed with a + or - sign instead of the common AD or BC notation; by convention 1 BC is labelled +0000, 2 BC is labeled -0001, and so on (ISO, 2002).

A *TM_ClockTime* is a data type that shall be used to identify a temporal position within a day. Similarly, a *TM_DateAndTime* is a subclass of both *TM_CalDate* and *TM_ClockTime* which provides a single data type for identifying a temporal position with a resolution of less than a day.

A *TM_Coordinate* is a data type that shall be used for identifying temporal position within a temporal coordinate system. A specialization of this type is JulianDate which identifies a position with respect to the Julian proleptic calendar.

Finally, *TM_OrdinalPosition* is a data type that shall be used for identifying temporal position within an ordinal temporal reference system.

4. Modeling SITA*R Time Dimensions with Standard ISO 19108

In the archaeological context time dimension may be specified using different reference systems and different calendars. For this reason, this paper considers only *TM_TemporalPosition* objects as possible instances for *TM_Position*, while it does not consider *DateTime*, *Date*, or *Time*. In other words, it assumes that the reference system and the used calendar are always explicitly declared.

In SITA*R three main objects of interest can be recognized: *ST_InformationSource*, *ST_ArchaeoPart* and *ST_ArchaeoUnit*, which are also characterized by some temporal dimensions discussed in this section. An *ST_ArchaeoUnit* is a complex archaeological entity obtained from an interpretation process performed by the responsible officer. Such interpretation is performed based on some finds retrieved during an excavation process or a bibliographical analysis, which are represented by *ST_ArchaeoPart* instances that are selected during the interpretation process. Given an *ST_ArchaeoUnit* object a set of possible temporal phases of its evolution are identified, then the selected *ST_ArchaeoPart* objects are assigned to one of the phases. This assignment process is one of the fundamental tasks in archaeology (Katsianis *et al.*, 2008). For instance, examples of phases in the existence of an archaeological entity are: installation/foundation, life/use, and renovation/reuse.

In SITA*R the sequence of phases describing the evolution of an *ST_ArchaoeUnit* object is defined as an *ST_Sequence* object, which in turn is a composition of *ST_Phase objects* (Figure 3).

In order to link these classes of the SITA*R model to the ISO Standard (ISO, 2002), we can observe that, since the relative order between each pair of phases is typically known with more certainty than their absolute position, the collection of phases in the sequence of an *ST_ArchaeoUnit* object can be modeled using a topological approach, as also suggested in the Standard. Fig. 6 illustrates the final result obtained by applying this solution. More specifically,

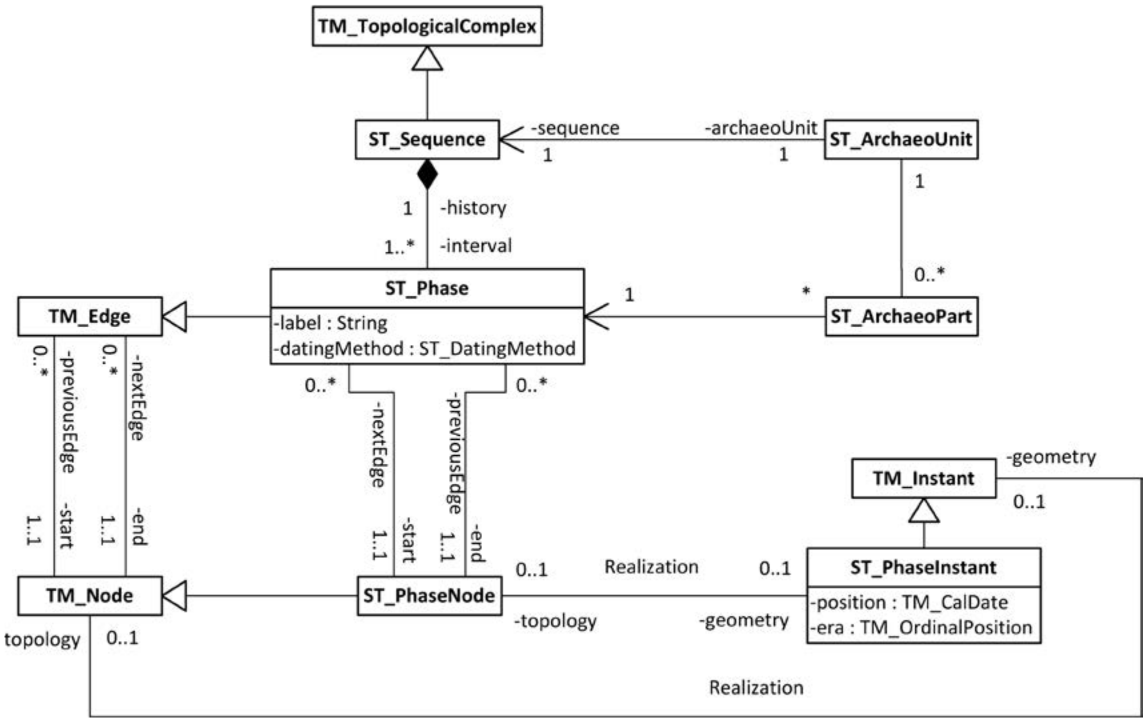


FIGURE 3: REPRESENTATION OF THE TIME ASPECTS CHARACTERIZING AN ARCHAEOLOGICAL UNIT IN SITA*R.

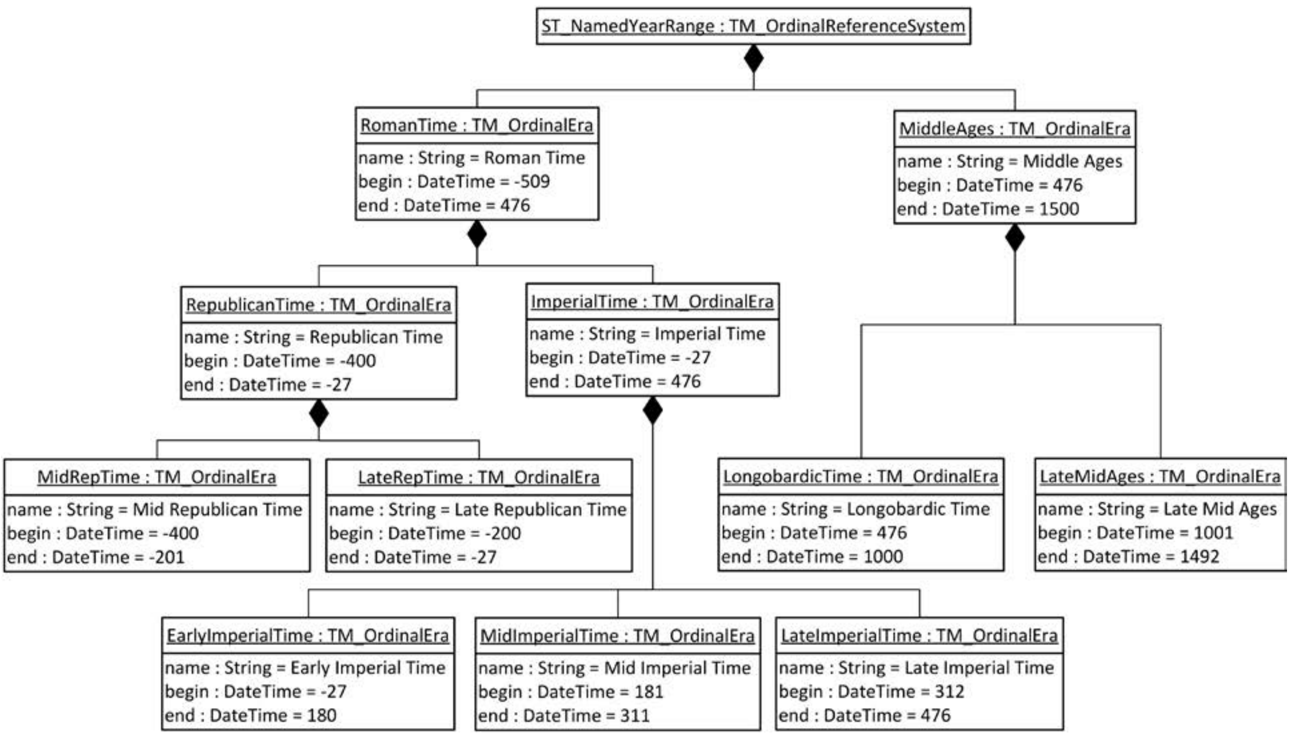


FIGURE 4: EXAMPLES OF ORDINAL ERAS USED IN SITA*R.

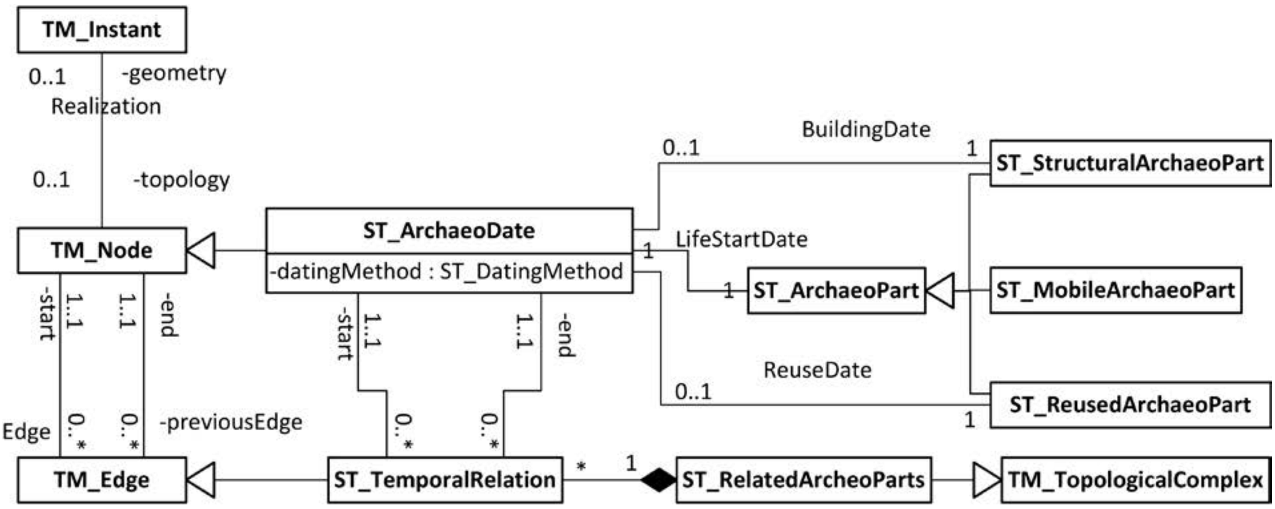


FIGURE 5: REPRESENTATION OF THE TIME ASPECTS CHARACTERIZING AN ARCHAEOLOGICAL PARTITION IN SITA*R.

the *ST_Sequence* object of an *ST_ArchaeoUnit* can be described as a topological complex, thus the *ST_Sequence* class can be declared in the model as a specialization of the *TM_TopologicalComplex* class of the Standard. Moreover, an *ST_Sequence* object is composed of *ST_Phase* objects; therefore, the *ST_Phase* class has to be declared in the model as a specialization of *TM_TopologicalPrimitiv*e class (i.e., *TM_Edge*, since it represents a period). SITA*R adds two additional properties: a meaningful label (e.g., foundation, use, etc.) and the specification of the dating method (e.g., stratigraphic analysis). Also the *Initiation* and *Termination* associations are specialized, because they connect a *ST_Phase object* with particular nodes (instances of the class *ST_PhaseNode* specializing *TM_Node*) which can be realized with a specialization of *TM_Instant*, called *ST_PhaseInstant*. Each *ST_PhaseInstant* has two attributes: a position (inherited from *TM_Instant*), which here can be only of type *TM_CalDate*, and a new attribute, called era, which is a *TM_OrdinalPosition*; at least one of them has to be not null. The value of the era attribute is a *TM_OrdinalEra* object defined with reference to a particular *TM_OrdinalReferenceSystem*, which is called *ST_NamedYearRange* in SITA*R and is exemplified in Fig. 4.

Each *ST_ArchaeoUnit* is connected to one or more constituent *ST_ArchaeoPart*, each one representing a single result of an excavation or other investigation processes denoted by the associated information source. For instance, it can be a structural element, a mobile element or a reused element. Each *ST_ArchaeoPart* is dated in some way and is assigned to a certain phase of the associated *ST_ArchaeoUnit*. In particular, any *ST_ArchaeoPart* is characterized by a *LifeStartDate* role which identifies the beginning of the object life. Moreover, if the partition identifies a structural element, it has also a *BuildingDate* role which denotes the date of its construction completion, while if the partition is a reused element, it is also characterized by a *ReuseDate*. An implicit constraint exists between the life-start date assigned to an archaeological partition and the possible additional dates: both *BuildingDate* and *ReuseDate* have to be after the *LifeStartDate*. Moreoever, constraints can also be defined regarding the phase assignment related to the association of the partition with an archaeological unit: for instance, the life-start date of a mobile or structural partition shall be contained in the assigned phase, while those of a reused element shall precede the phase start node (Figure 5).

The date assigned to an *ST_ArchaeoPart* object is described in the model by the *ST_ArchaeoDate* class, called partition chronology, which is related to the ISO Standard as it is a specialization of the *ST_Node* class and has consequently a realization in the *ST_Instant* class. An additional attribute describing the applied dating method characterizes the *ST_ArchaeoDate* class. Exploiting the ISO classes, the chronology of a partition can also be represented by topological primitives, since a relative order between related partitions is better known, than their absolute location. Some edges, called *ST_TopologicalRelation*, can be placed between nodes representing *ST_ArchaeoDate* objects, in order to define temporal relations between related archaeological partition dates. A set of temporal relations related to some connected partitions constitute a topological complex, called *ST_RelatedArchaeoParts*. In accordance with the Standard (ISO. 2002), the relative positions of two *TM_TopologicalPrimitives* depend upon the positions they

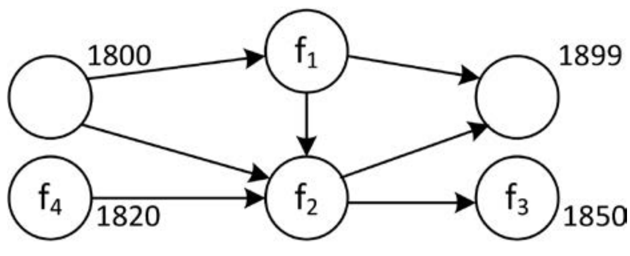


FIGURE 6: EXAMPLE OF TOPOLOGICAL COMPLEX REPRESENTING ORDINAL TEMPORAL RELATIONS BETWEEN CHRONOLOGIES OF ARCHAEOLOGICAL PARTITION.

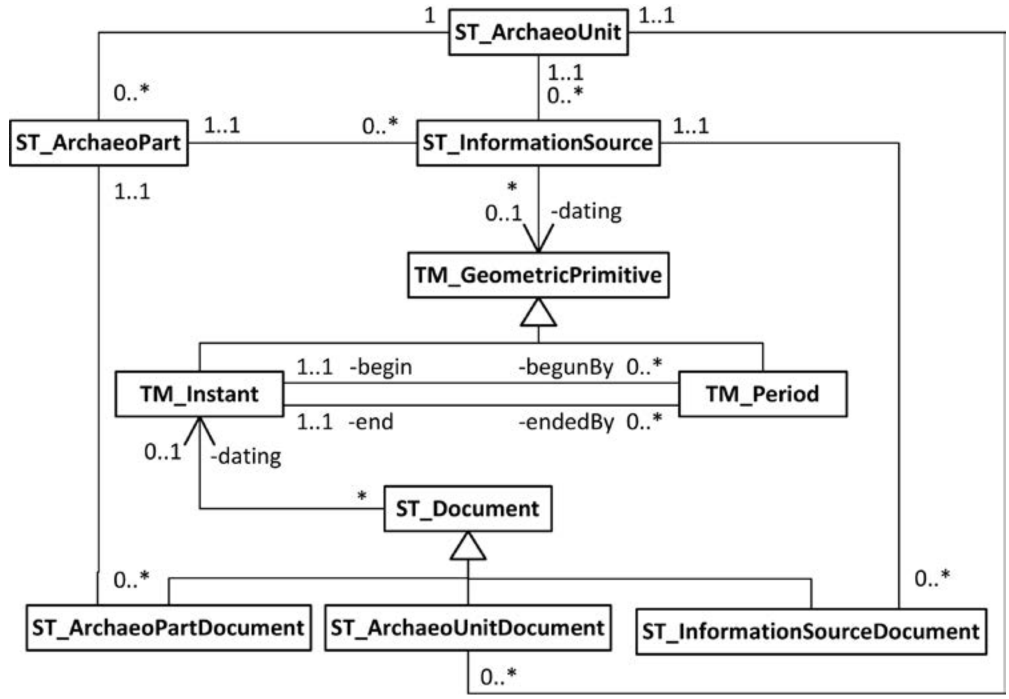


Figure 7: Representation of the time aspects characterizing an information source and a document in SITA*R.

occupy within the sequence of *TM_TopologicalPrimitives* that make up a *TM_TopologicalComplex*, as discussed in Fig. 2 of Sec. 3. The following example illustrates a possible topological structure composed of a set of related archaeological partitions.

Example 1 - Let us consider four archaeological finds labeled as f1, f2, f3 and f4 which are coarsely dated as follows: f1, f2 are located in the 19th century, while f3 is dated 1850 and f4 is dated 1820. Besides these geometrical values, the following temporal relations have been detected: f1 before f2 and f3, while f2 before f3 and after f4. This knowledge can be represented by the topological complex in Fig. 6. Dates associated to nodes f3 and f4 are realized as the years 1850 and 1820, respectively. Conversely, dates related to nodes f1 and f2 are not realized, but they are located between two dummy nodes representing the years 1800 and 1899. Given such topological structure some automatic reasoning techniques can be applied in order to realize also such dates. In particular, all dates between 1820 and 1850 could be consistent realizations for f2, while all dates between 1800 and 1820 could be consistent realizations for f1 (Figure 6).

Each *ST_ArchaeoPart* and each *ST_ArchaeoUnit* refers to an instance of *ST_InformationSource*. An *ST_InformationSource* represents the way used to start collecting information about an archaeological object: for instance, it can be an excavation, a bibliographical study, a construction work, and so on. Each *ST_InformationSource* is characterized by a time dimension that, in accordance to (De Roo, 2013), is represented as a geometric primitive, since it is a generally known and documented in some way, as illustrated in Figure 7. This geometric primitive can be instantiated with both a *TM_Instant* or a *TM_Period* depending on the particular type of information source and the available information.

Finally, each of these three main entities can be connected with another SITA*R object characterized by a time dimension: *ST_Document*, which represents a generic collected document; for instance, an excavation report, a cartographic product, and so on. Three specializations of documents are defined: one for information sources, one for archaeological partitions and one for archaeological units. Similarly to the time dimension of an information source, the dating of this object is also represented by a geometric primitive, since it refers to a modern-age date which is usually well-documented.

5. Extending the Standard for Modeling Vague Time Dimensions

The main lack of the Standard ISO 19108 in the representation of archaeological time is the absence of constructs for expressing vagueness. This section analyses how fuzzy concepts can be incorporated into the model presented in the Sec. 3. In particular, this paper concentrate on trapezoidal fuzzy distributions, since they are computationally less expensive, while they provide a sufficient representation of the time knowledge generally provided by archaeologists, as stated in (Belussi & Migliorini, 2014). In the archaeological domain, temporal knowledge is generally characterized by a level of vagueness and dates are usually expressed as periods of great confidence together with a safety additional interval. For instance, the construction date of a building can be expressed as: between 1830-1850 with more confidence plus or minus 10 years of safety. Fuzzy set theory and

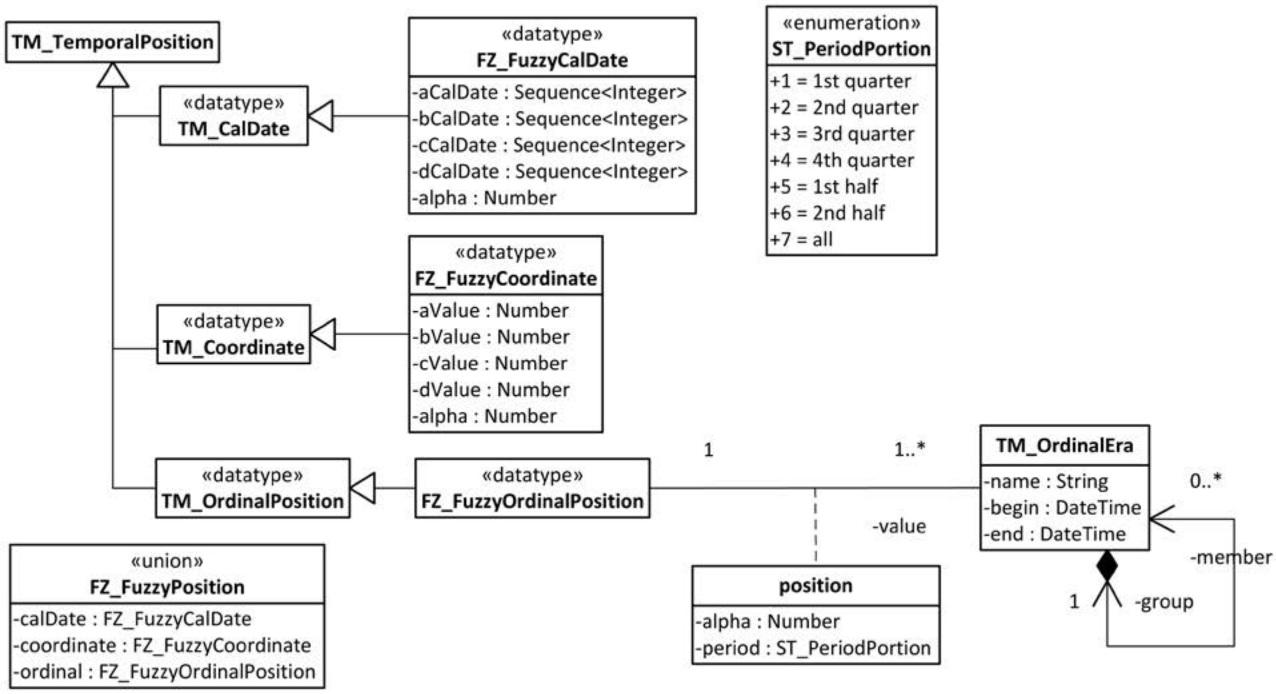


Figure 8: Fuzzy extension of the temporal positions in Standard ISO 19108.

probability theory are two related but different ways for modeling uncertainty. In Sanjaa & Tsoozol (2007), the authors discuss the similarities and differences between these two theories. In particular, they state that probability statements are about the likelihoods of outcomes: an event either occurs or does not, and you can choose on it. This theory is typically used to make predictions and is characterized by only two outcomes: true and false. Conversely, fuzzy set theory was introduced as a mean to model the uncertainty of natural language and is extended to handle the concept of partial truth (or degree of truth). It cannot say clearly whether an event occurs or not and is usually applied for describing happened events. For these reasons, a fuzzy representation of time seems to be the more appropriate solution.

Definition 1 - A fuzzy set F is a pair (U, μ), where U is a set and μ is a membership function $\mu : U \rightarrow [0, 1]$, such that for all $u \in U$ the value μ(u) represents the grade of membership of u in F. In particular, μ(u) = 1 reflects full membership of u in F, while μ(u) = 0 express the absolute non-membership in F. The set $\{ u \in U \mid \mu(u) = 1\}$ is called *core* of *F*, while the set $\{ u \in U \mid \mu(u) > 0\}$ is called *support* of *F*.

Definition 2 - A trapezoidal membership function can be encoded by a 4-tuple (*a, b, c, d*), where the intervals [*b, c*] and [*a, d*] represent the core and the support of the fuzzy set, respectively. A value $\alpha \in [0,1]$ can be added to the tuple in order to represent a maximum grade of membership when it is different from 1.

As a general idea, each possible *TM_TemporalPosition* will be extended in order to express a possibility membership function instead of a certain date. In particular, in the archaeological context time granularity is never more fine-grained than a day, thus we can safely omit to consider the *TM_ClockTime* and *TM_DateAndTime* datatypes. The fuzzy extension of the temporal position is illustrated in Fig. 8.

Each calendar date is represented in a fuzzy form using the *FZ_FuzzyCalDate* datatype which contains a trapezoidal tuple (a, b, c, d)[α], where a, b, c, d are sequences of integers representing dates. Similarly, a fuzzy ordinal positions inside an ordinal temporal reference system is represented with a specialized class *TM_FuzzyOrdinalPosition*, which has a qualified association with the corresponding *TM_OrdinalEra* enriched with a degree of possibility $\alpha \in [0,1]$ and a period. The period attribute allows one to specify a portion (e.g., the first quarter) of the selected era which is more possible. Moreover, the cardinality on the era side is changed from 1 to 1..*, since different positions can be defined each one with a possibility value. These positions can be interpreted as a disjunction of positions. Finally, each coordinate inside a coordinate reference system is extended by the datatype FZ*_FuzzyCoordinate* which contains four numeric values representing the trapeze extremes, and the value α.These datatypes can be used as value for the union *FZ_FuzzyPosition*. A *FZ_FuzzyPosition* is the type of the position attribute of a generic *FZ_FuzzyInstant* which is a fuzzy specialization of a temporal instant.

The last aspect to be considered regards the relative ordering between topological primitives inside the same topological complex. In particular, the Standard establishes how to determine the relative ordering between topological primitives, based on their position in the sequence that makes up the topological complex. However, in a fuzzy environment such relations cannot be certain but are characterized by a possibility value.

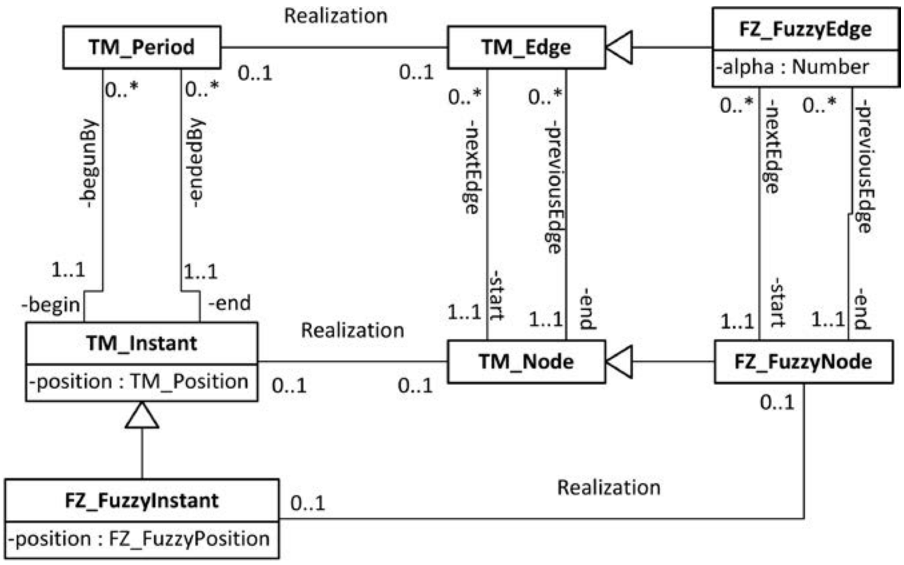


FIGURE 9: FUZZY EXTENSION OF THE TOPOLOGICAL PRIMITIVES IN STANDARD ISO 19108.

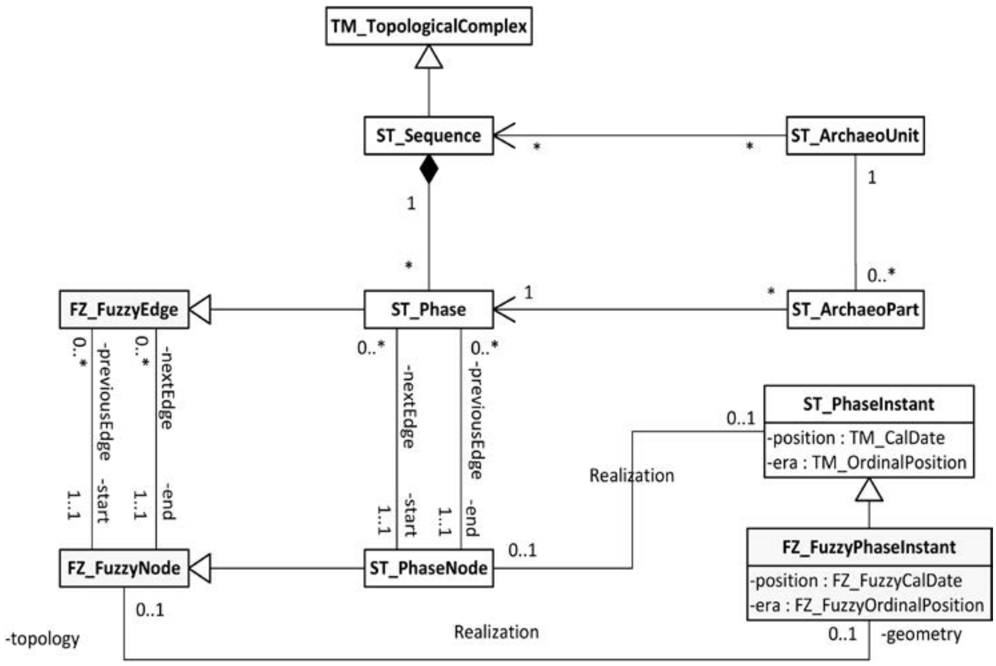


FIGURE 10: FUZZY REPRESENTATION OF A ST_ARCHAEOUNIT.

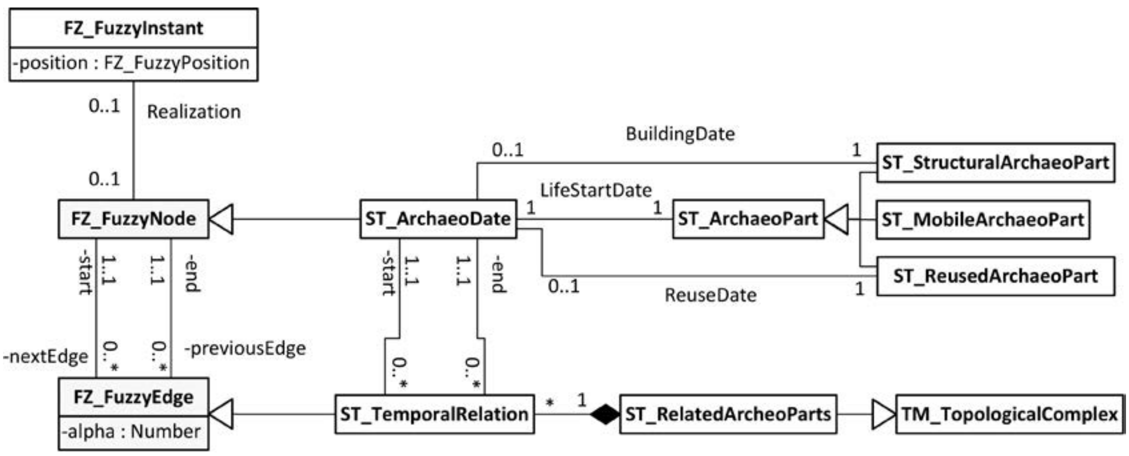


FIGURE 11: FUZZY REPRESENTATION OF A ST_ARCHAEOPART.

Therefore, a specialization of *TM_Edge* is defined which is called TM_FuzzyEdge and is enriched with a possibility value $\alpha \in [0,1]$, as illustrated in Fig 9. When the a *FZ_FuzzyEdge* is not realized, it simply represents an uncertain relation between two nodes, while when it is realized the corresponding period is characterized by two fuzzy extremes, as illustrated in Fig. 9.

6. Modeling Vague SITA*R Time

This section illustrates how the fuzzy datatypes presented in the previous section can be used for modeling vague time aspects in SITA*R.

As regards to *ST_ArchaeoUnit*, the extension of its time aspects is illustrated in Fig. 10. First of all, a phase instant is represented by a *ST_FuzzyPhaseInstant* whose position and era attributes are redefined to be typed with the corresponding fuzzy datatypes. Similarly, the topology is represented by the corresponding fuzzy types making uncertain the relation between phases.

An *ST_ArchaeoPart* has three time dimensions: its own dating, the relation with a phase of the corresponding archaeological unit, the definition of a set of time relationships between archaeological partitions. Each *ST_ArchaeoDate* can be realized through a *FZ_FuzzyInstant* in order to express the vagueness of the dating process. Conversely, the assignment to a particular phase remains unchanged even in presence of vagueness, while the relation between two archaeological partitions is represented by a *FZ_FuzzyEdge* in order to assign a possibility value to each relation. This new representation of *ST_ArchaeoPart* is illustrated in Fig. 11.

7. Conclusion

This paper proposes a model for representing and managing time dimensions in archaeological data. In particular, it evaluates the applicability of the Standard ISO 19108 by considering two real-world information systems, called SITA*R, which has been developed for the archaeological data of Roma and Verona. From this preliminary analysis has emerged that the Standard is unable to represent the inherent vagueness of archaeological data. Therefore, an extension of the Standard concepts has been defined which is based on a fuzzy representation of dates and of ordering relations about time points. Such extension has been successfully applied to the SITA*R case.

Time knowledge about particular findings and relations among them are typically used by archaeologists to derive new knowledge or during the interpretation process. In literature many techniques have been proposed for automatically deriving new knowledge from available data. The model proposed in this paper can be easily translated into formalisms suitable for applying these reasoning techniques. They allow one to answer two main questions: check the network consistency and compute the minimal network in order to reduce some vagueness. The answers to these questions can be used to guide archaeologists in the complex dating and interpretation process.

References

ALLEN, J. F. (1983). Maintaining Knowledge About Temporal Intervals. *Communications of the ACM*, 26(11). pp. 832–843.

BASSO, P., BELUSSI, A., GROSSI, P., and MIGLIORINI, S. (2013). Towards the Creation of an Archaeological Urban Information System: Data Modeling of the Historical Heritage of Verona. In W*orkshop on Integrating 4D, GIS and Cultural Heritage* (AGILE 2013), pp. 1–3.

BELUSSI, A., and MIGLIORINI, S (2014). Modeling Time in Archaeological Data: the Verona Case Study. *Department of Computer Science, University of Verona*, Tech. Rep. RR 93/2014.

DE ROO, B., VAN DE WEGHE, N., BOURGEOIS, J. and DE MAEYER, P. (2014). The Temporal Dimension in a 4D Archaeological Data Model: Applicability of the GeoInformation Standard. In: *Innovations in 3D Geo-Information Sciences*, Lecture Notes in Geoinformation and Cartography. Springer International Publishing, pp. 33–55.

ISO. (2002). 19108 Geographic Information – Temporal Schema.

ISO. (2004). 8601 Data elements and interchange formats – Information interchange – Representation of dates and times. KATSIANIS, M., TSIPIDIS, S., KOTSAKIS, K., and KOUSOULAKOU, A. (2008). A 3D Digital Workflow for Archaeological Intra-Site Research using GIS, J*ournal of Archaeological Science* 35(3), pp. 655–667.

LI, J., WU, S., and HUANG, G. (2002). Handling Temporal Uncertainty in GIS Domain: A Fuzzy Approach. *International. Arch. of Photogrammetry, Remote Sensing and Spatial Inf. Sciences* 34(4/w4), pp. 83–87.

SANJAA, B., and TSOOZOL, P. (2007). Fuzzy and probability. In: *Proceedings of International. Forum on Strategic Technology (IFOST 2007), October 2007, Ulaanbaatar, Mongolia*. IEEE, pp. 141–143.

Towards an Archaeological Information System: Improving the Core Data Model

Muriel van RUYMBEKE,[1] Cyril CARRÉ,[2] Vincent DELFOSSE,[3] Pierre HALLOT,[2] Michelle PFEIFFER,[4] Roland BILLEN[3]

[1] ULg Library, University of Liege, Belgium
[2] Geomatics Unit, Department of Geography, University of Liege, Belgium
[3] Laboratory for User Cognition & Innovative Design, University of Liège, Belgium
[4] Cluster of Excellence Asia and Europe in a Global Context, University of Heidelberg, Germany
mvanruymbeke@ulg.ac.be, Cyril.Carre@ulg.ac.be, Vincent.Delfosse@ulg.ac.be, p.hallot@ulg.ac.be, michelle.pfeiffer@asia-europe.uni-heidelberg.de, rbillen@ulg.ac.be

Abstract

Geographical Information Systems demonstrated their utility in numerous archeological studies. However, such systems do not encompass all the characteristics of the archaeological data. Following an interdisciplinary approach, the Geomatics Unit (University of Liege) develops a robust, flexible and dynamic Archeological Information System. This research focusses on the general archeological data characteristics namely the spatio-temporal, heterogeneous, multimodal and imperfect components. In this paper, we present the latest outputs of this project and more specifically the integration of archeological data ambiguity. We depict the ambiguity management by introducing two new concepts; Lifemap and Interpretative sequence that make it possible to handle ambiguous hypothesis regarding historical objects within our system.

Keywords: Cultural Heritage, Archaeological Information System, Data model, Spatio-temporal database, Data Imperfection

1. Introduction

Since 2009, notably through the project *Virtual Leodium*, the Geomatics Unit of the University of Liège and their partners (Billen et al., 2009; Billen et al., 2012) work on the development of a robust, flexible and dynamic Archaeological Information System. Such system should provide solutions to the increasing needs of archaeologists and historians for a powerful set of tools for managing, archiving and exploring large archaeological data sets to encompass archeological problems (Burrough and McDonnell, 1998).

Due to the archaeological data complexity, namely its spatio-temporal, heterogeneous, multimodal and imperfect character, the development of an Archaeological Information System turns out to be a highly complex challenge that has been faced in the past resulting in numerous and versatile data models. Despite the high number of models, data imperfection is often ignored (Gonzalez-Perez, and Parcero-Oubina, 2011). Only few authors consider imperfection (namely for modelling purposes) (De Runz 2008, Pillot and Saligny 2011, Desachy 2012, Desjardin et al. 2012). These works distinguishes four different (but non-exhaustive) types of archaeological data imperfection: uncertainty, imprecision, ambiguity and incompleteness.

This paper presents the latest improvements of the *Virtual Leodium* core data model. In the previous archaeological information modeling works (Billen et al., 2012, Pfeiffer et al., 2013), we proposed solutions to handle geometrical ambiguity and incompleteness. Here, we present an improved model handling time and function's imperfection. A new concept called *Episode* allows dealing with *Events* in a wider understanding than in the previous models. Moreover, new classes have been created. *Interpretative Sequence* organizes episodes into ordered path. *Life Map* amounts to *Historical Object* timeline, gathering all episodes even if they are contradictory. Finally, *Agent* and *Figure* classes complete the model.

The first part of the paper outlines the research context and our case study *Virtual Leodium*. We focus on the workflow, the archaeological information model and the current prototype of the archaeological information system. In the second part, the improvements of the core data model are raised theoretically and illustrated through a new case study, the archaeological modelling of the Weser (B) drainage basin.

2. The *Virtual Leodium* project

The initial purpose of *Virtual Leodium* was the development of an Urban Archaeological information system based on a city scale model. The scale model represents the city of Liège at the beginning of the 18th century on a scale of 1/1200 on around 1m². Its creator Gustave Ruhl, an amateur of archaeology took more than ten years to model this masterpiece of art. The model dates back to the beginning of the 20th century and is an important historic source for the city of Liège before some major changes occurred in the city due to the industrialization period. The access to the masterpiece of art, accommodated since 1907 by the University of Liège, has for long been restricted due to three reasons: its location (in a manuscripts reading room), its fragility (impossible to move), and its conservation condition (requiring special thermo-hygrometric conditions).

When the project set up, its aims were twofold; creating a 3D model of the scale model to preserve the city model and hence guarantee access for a larger public in order to valorize its documentary and didactic value; and furthermore enriching the city model semantically by iconographic, textual and archaeological sources documenting the scale model and in a broader sense the city of Liège itself (Pfeiffer, 2011).

To do so a workflow composed of three different phases, namely the 3D data acquisition that is subdivided in the 3D data acquisition and the semantic data acquisition, the development of an archaeological data model and finally the implementation of the archaeological information system.

2.1. Data acquisition

The model has been scanned in 3D using an optical inspection technique (Billen et al., 2009). This result in a 3D points cloud composed of 650 scans that is cleaned up and merged with the 3D software Geomagic Studio. The data reconstruction is performed manually using the out coming data surface models and photographs taken from all over the scale model.

Before using the photographs as draping texture within 3D graphics software Maya, the photographs are rectified with the software Photoshop. The 3D modelling phase is still in process, for now about 25% of the entire model is modelled in 3D. The 3D data acquisition workflow is proposed in Fig. 1.

In the second sub-phase, the semantic data acquisition is performed. For this phase numerous documents (more than 600) left by Ruhl that includes notes plans, sketches, drawings, and photographs documenting the scale model regarding, i.e. its material, scale, and sources of inspiration, are collected and studied (Pfeiffer, 2011). In general, sources documenting the eighteenth century of Liège's framework are also studied and in a further step integrated into the spatial database.

2.2. The archaeologically-based data model

The archaeological object is placed in the center of our model and is the smallest unit of our system. An

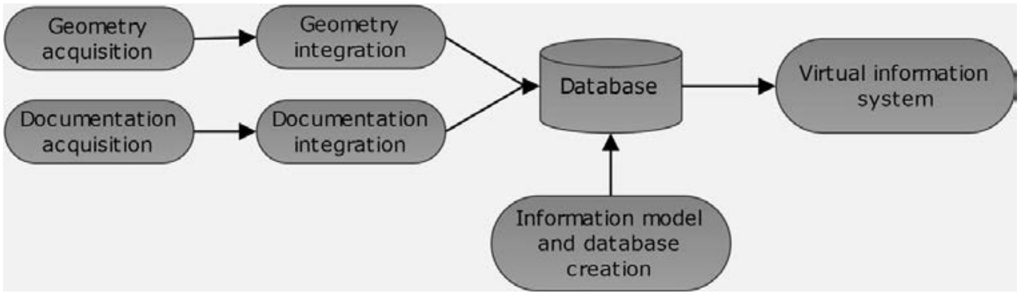


Figure 1: Workflow of the Virtual Leodium project.

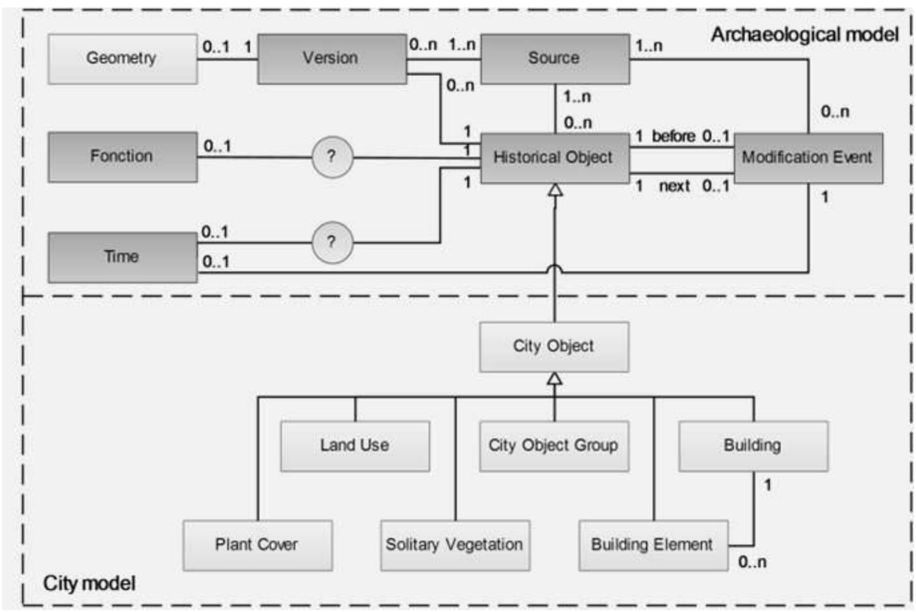


Figure 2: *Virtual Leodium information UML model.*

historical object is defined by its spatiality, temporality and its functionality. This definition relies on the object identity definition proposed by Peuquet (Peuquet, 1994) for geographical object and re-used for archaeology by Rodier (Galinié *et al.*, 2004; Rodier *et al.*, 2011; Rodier *et al.*, 2012). However, the definition used within our data model differs from the one proposed by Rodier. For Rodier, a change of one of the three constitutional components of an object implies a change of the object identity. We opted for a more flexible identity definition, which means that a change impacting one of the three constitutional components does not necessarily imply an object identity change. This definition relies on the proposition of Hornsby and Egenhofer (Hornsby and Egenhofer, 2000). The second difference concerns the data imperfection that is not handled in Rodier's approach. We adopt imperfection types of De Runz (De Runz, 2008). First, the presence of the three constitutional components (spatiality, temporality and functionality) is made optional in order to manage the incompleteness of the object. Then, we create a concept called *Version*, which allows representing multiple spatial representations for the same object and consequently is a way to manage spatial ambiguity. The *Modification Event* is a concept based on the event management, a quite common concept within the field of computation. With the *Modification Event* it becomes possible to keep track of the changes of an archaeological object in a large temporal framework. In such way that changes between different states of an archaeological object are recorded.

The model is developed according to the multimodal character of the archaeological data. This means that it is supposed to handle the three most common data sources within archaeology, namely: the written sources, the iconographic sources and the archaeological remains, although for now the implemented prototype of the AIS handles only written and iconographic sources. The *Source* can document a building, a building element or a group of building but also a *Modification Event* related to an archaeological object.

The other objects of the model are picked up from the OGC standard CityGML (Gröger *et al.*, 2012) for sake of interoperability with 3D geographical information model.

2.3. Archaeological Information System Prototype

The archaeologically-based data model is mapped into an ArcGIS Geodatabase; it is the core of the Archaeological Information System (AIS) prototype developed during the *Virtual Leodium* project. The AIS relies on client-server architecture and is developed in Java. The ArcGIS Engine provides the geo-database functionalities. The ArcScence API provides a visualization platform for the 3D data. The previously collected geometric and semantic information is thereby stored in an ArcGIS Geodatabase. Currently only the simplified version of the archaeologically-based data model is implemented. Interested readers are invited to refer to Billen (Billen *et al.*, 2012) to find out more information about the simplified data model.

AIS deals with a two-component interface. On one hand the user can visualize and navigate through the 3D model, on the other hand the user can query the related semantic database. Due to the synchronization of both components, the system provides semantic information about the different buildings by clicking or selecting them, concurrently querying the semantic database highlights the related 3D objects in the 3D scene. The system handles partly the concept of ambiguity and incompleteness by supporting multiple 3D versions of the same object.

3. The Weser Project

The archaeological modelling of the Belgian Weser drainage basin project aims at studying settlements and territorial management of the area in Ancient and Early Medieval time. The research objective is to perform spatio-temporal analysis and predictive modelling of the related archaeological finding sites. This requires recording all archaeological structure evolutions and changes over a long term period and also all related scientific opinions or restitutions. We decide to apply the *Virtual Leodium*'s data model and improve it in order to resolve remaining critical modelling issues. First of them is the *Historical Object* identity and discriminant component; how can we determinate what causes a change in an object or, on the contrary, what causes the replacement of an object by another one? The second one concerns both time and function's imperfection. Finally the understanding of the event management in its computational signification ensures that the changes between versions are managed. This prospect is insufficient for the new case study that needs a huge management of events impacting one (or two or even the three) components of the *Historical Object*.

3.1. The improved data model

The UML class diagram (Fig. 3) of the model has been designed using ArgoUml software. Just like *Virtual Leodium*'s model, the core data model links to external models or ontologies for geometry, function and time and to authority bases or data RDF for *sources*, *agent* and *figures*. With the exception of inheritance relationship between *episode* and *version* and *event*, all the relationships of the model are composition relationships.

The following explains the main characteristics of the model. The first improvement relates to *Historical Object* definition. We suggest to define the *Historical Object* as follows: a consistent group of elements belonging to the same body from its emergence until its disappearance. The body in question can be an architectural body, a professional corporate body, a human body, etc. In an archaeological sense the body can be seen like a consistent group of facts and units of stratigraphy. This definition is either close to the 'structure' definitions of M. Bats (Bats *et al.*, 1986) and P. Van Ossel (Van Ossel *et al.*, 1988). Indeed, adapting for the Lattes (F) excavations the Harris principles of archaeological stratigraphy (Harris, 1979), Bats defines a structure as a group of contemporaneous facts sharing a same function. Two years later, Van

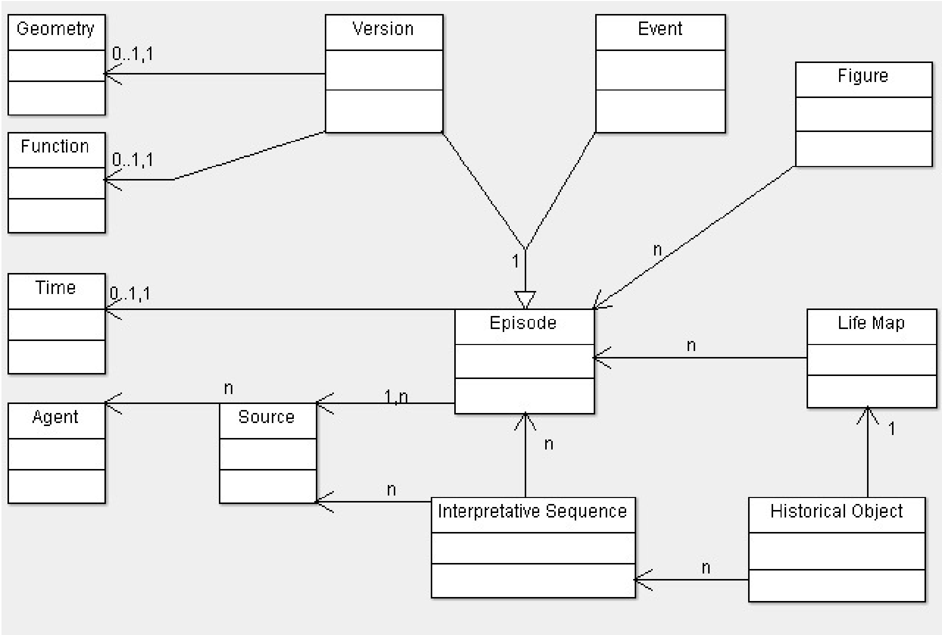


Figure 3: Improved core data model.

Ossel gives up the contemporaneousness to keep only the function's uniqueness. To us, neither time, space nor function is predominant to discriminate the identity of an object. What matters is the belonging to a same body. According to Emmanuel Kant's Theory of knowledge (Kant, 1781), the subject constructs its knowledge object, not the contrary. Following this opinion, we think it is the user's decision and responsibility to establish what type of body he wants to work with in a single project taking into consideration its specific requirements. Besides object identity definition, we wanted also to encompass both function and time lack of fulfillment, incoherency, uncertainty, imprecision, contradiction, like we did for geometry in the *Virtual Leodium* model. For that reason, the version concept has been generalized to the function component.

The version is a *Historical Object* state, documented by *sources* from which we know something about the function and – or the geometry and the chronology of a *Historical Object* state. Taking into account the fact that chronology is linked with the generalized class *episode*, we can consider that *version* is composed by geometry, function and time. But we consider also that each of these components may stay empty because of a lack of documentation. A *version* exists from the moment that a source tells something about the state of a *Historical Object*. If different sources describe the same geometry, with the same function and time, it is the same version. If there are differences between geometry, function or time, there are new *versions* (therefore new *episodes*) of the *Historical Object*. Two versions are thus used to express the fact that there are two different functions associated with the same geometry and chronology, or two different geometries associated with the same function and chronology or two different chronologies for a same association function – geometry. A new *version* will occur for a geometrical information improvement for example.

Compared to the *Virtual Leodium* model, the event concept is here expanded. We leave the computational vision exclusively dedicated to record changes between two different states of the system and we adopt the cognitive approach where events are used as phenomena descriptor (Pfeiffer *et al.*, 2013). Therefore, an *Event* is a documented fact, possibly interacting with an HO state but without necessarily generating a change. Nevertheless, it can act as versioning for tracking changes between versions as well. The more interesting property of *Event* is to manage *Historical object*'s important data devoid of geometry. *Event* has neither geometry nor function. It has a more historical nature. It can be seen like the Harris 'feature interfaces' units of stratigraphy. However all historical facts recorded for an HO are not necessarily *events*; most of them are just function's expression and should be recorded as such.

Episode class generalizes *Version* and *Event* classes. It links them not only with *Time* but with *Source* as well. It centralizes the whole model around these two most specific classes of historical domains. Although *source* class must absolutely be filled to allow episode's existence, time can remain empty. With these characteristics and with the help of *version* and *event*, *episode* ensures management of a proportion of imperfect data, notably the lack of fulfillment, imprecision and incoherency.

Interpretative sequence is an ordered series of episodes. It has one or more sources. The later has one or more authors, most often historians, archaeologists or architects, all of them involved in studying the historical object story and recording or publishing their interpretations or restorations. However, some sources may be anonymous. Information they provide must be recorded as well. Interpretative sequence has only one direction. It displays a linear or a cyclic time but doesn't admit flashbacks or contemporaneousness. An interesting aspect of interpretative sequence is that it is the way to deal with relative chronology. It is the way to encompass contradictory information, incoherency or doubt as well: if two people don't agree about a building history it is not an issue; two different interpretative sequences would be set up.

LifeMap amounts to the *Historical Object* Timeline. It organizes all its episodes according to a chronological grid despite their involving, or not, into an interpretative sequence. Indeed it is possible to record floating episodes without integrating them into interpretative sequences. *LifeMap* can be understood like the historical object's path, or its route built by the amount of versions and events stored in the system. Besides *time* linked with *episode*, *LifeMap* constitutes therefore a second temporal index that facilitates temporal queries.

In the model, people belong to two different categories: historical figures, associated to historical objects episodes and present (or at least not too far away in the past) professional sources' authors. Two different classes have been created. *Agent* is the class storing sources and interpretative sequences authors while *Figure* assembles historical people, implicated in events.

Sources deal with documentation: direct or indirect, digital or printed, textual or figurative, artistic or scientific, sources are raw material for professional studying past. In historical domains, nothing exists without being documented. Sources provide information about episode and interpretative sequences as well. Their rereading and the re-examination of their metadata lead often professionals to suggest new historical interpretations.

3.2. Case study

A specific case study of the Weser project, the evolution of a single building (currently a church): ST-0101 (Fig. 4), will help understanding the use of the proposed model. In this example, the body type selected to define object's identity is building. Although most of the buildings can be described with names, the identifier of each structure follows the same model: ST (abbreviation for structure) and one number. This choice avoids affecting the data with some dominant interpretation; the case of ST-0101, described below, through light on the importance of semantic neutrality. Nowadays, this cultural heritage building is known under the name 'Saint Hermes and Alexander church' or 'Theux's church'. However, it is not that certain this building has always been a Place of Worship. Moreover, it had not always been dedicated to Saint Hermes and Alexander.

Studied since the nineteenth century, this listed monument has a long story. As the researches and publications went on, interpretations regarding its development grew in

Figure 4: Saint Hermes and Alexander Church, North side. ©KIK-IRPA, Bruxelles.

number. To feed the global study on the Weser's region archaeology, all of these interpretations must be recorded and managed. To illustrate the model described above and in order to facilitate the comprehension only few of them (the last ones) will be taken into consideration.

In the church's story state of art proposed by Bertholet and Hoffsummer (Bertholet and Hoffsummer, 1986), the authors give new scientific data (i.e. archaeological excavations or dendrochronological analysis). They used them as a starting point for their new building's evolution reconstruction divided into nine steps (Fig. 5). They date back to fifth century the very first building of the structure, a house or a pagan Place of Worship (Fig. 5: no. 1). Secondly, during the sixth or seventh century the authors think that the building was extended to become a Christian church, probably dedicated to Saint Pierre (Fig. 5: no. 2). In the second half of the ninth century this building was replaced by a bigger one (Fig. 5: no. 3): a long single-nave building ended by an oriental choir topped by a tour. According to Bertholet and Hoffsummer two events could explain this new building: the receipt of St Hermes' relics (around 860) and the new church's dedication, or a fire caused by a Norman's incursion. Around 1091 a new church, three times bigger than the previous one, was erected and also dedicated to Saint Alexander. It is a three naves hall-church covered by flat ceiling (Fig. 5: no 4). More than one century later in the beginning of the thirteenth century, a tower is added against the north wall (Fig. 5: no 6). From that moment, the church acts as fortified church. During the fourteenth century, important fortification works are

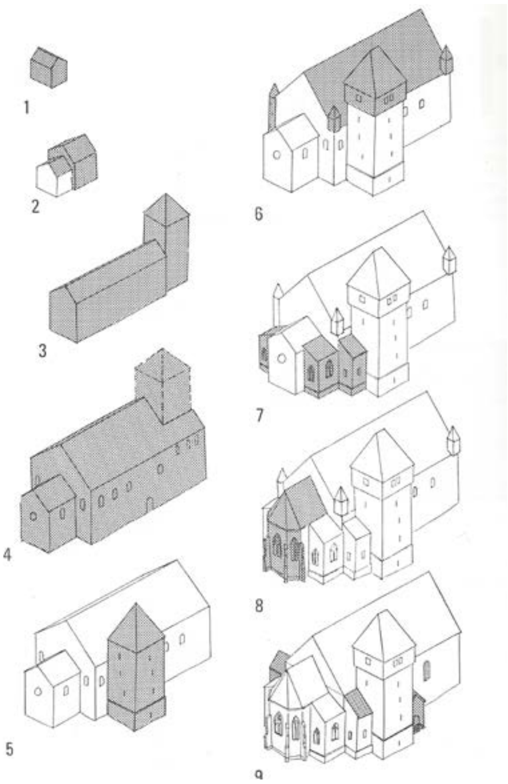


Figure 5: General Evolution of ST-0101 building (Bertholet and Hoffummer 1986, p. 256).

achieved: tower is equipped with hoardings (currently still preserved), naves roof are repaired and wooden turrets surmount the angles and the entrance (Fig. 5: no. 6). At the beginning of the sixteenth century, two transformation phases (Fig. 5: no. 7 and 8) modify the east part of the building to give it a gothic shape (Fig. 6). The fortified church use decreases until the eighteenth century phase (Fig. 5: no. 9) when the defensive elements are dismantled.

Using the proposed model, these nine steps scenario can be translated into one life map composed by eleven episodes (nine versions and two events) (Fig. 7). On that figure, light rectangles are versions while dark ones are events. Their first line gives their identifier, the second one their function (for version) or their category (for events), the third line gives time information and the fourth line (for versions only) points to geometrical data.

Four concurrent interpretative sequences organize the episodes series (Fig. 8) given that the first building is assumed to be either a Place of Worship or either a house and given also that two different events may have caused the building of third church.

Figure 6: Saint Hermes and Alexander Church, South side. ©KIK-IRPA, Bruxelles.

Moreover, this model allows the management of Interpretative sequences in large numbers. It ensures notably the integration of all scientific opinions related to a same structure. For example, the structure ST -0101 illustrated above has been re-examined on the occasion of artefacts analysis undertook at the European Center of Archaeometry of the University of Liège published in 2009 by Van Wersch (Van Wersch *et al.*, 2009). In her study, the author suggests to revise the dating of two first buildings. Consequently two new versions and four new interpretative sequences appear in the HO Life-Map (Fig. 9).

4. Conclusion and future developments

In this paper, we present the latest improvements of our Archaeological Information System. First, we exposed the general research context of the *Virtual Leodium* project while insisting on the general workflow and the current prototype. We outlined our current data model by insisting on the main concepts, namely the concept of *Version* that enables partly the handling of data imperfection and the object identity that we based on the main characteristics of the archaeological data. However, some unresolved issues remained in the model, notably the data imperfection management which is not yet completed. Furthermore, the object identity definition still needed to be improved to encompass its vague definition and the event management that was too restricted in regard to the archaeological data complexity.

In respect to the Weser project it turned out that the model had to be improved. In this research, we focused on the object identity definition which we enlarged by introducing the concept of *Body*. Thanks to this new modeling

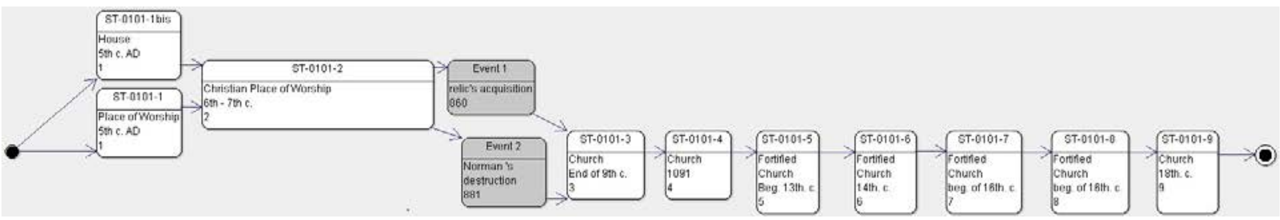


Figure 7: St-0101 first LifeMap.

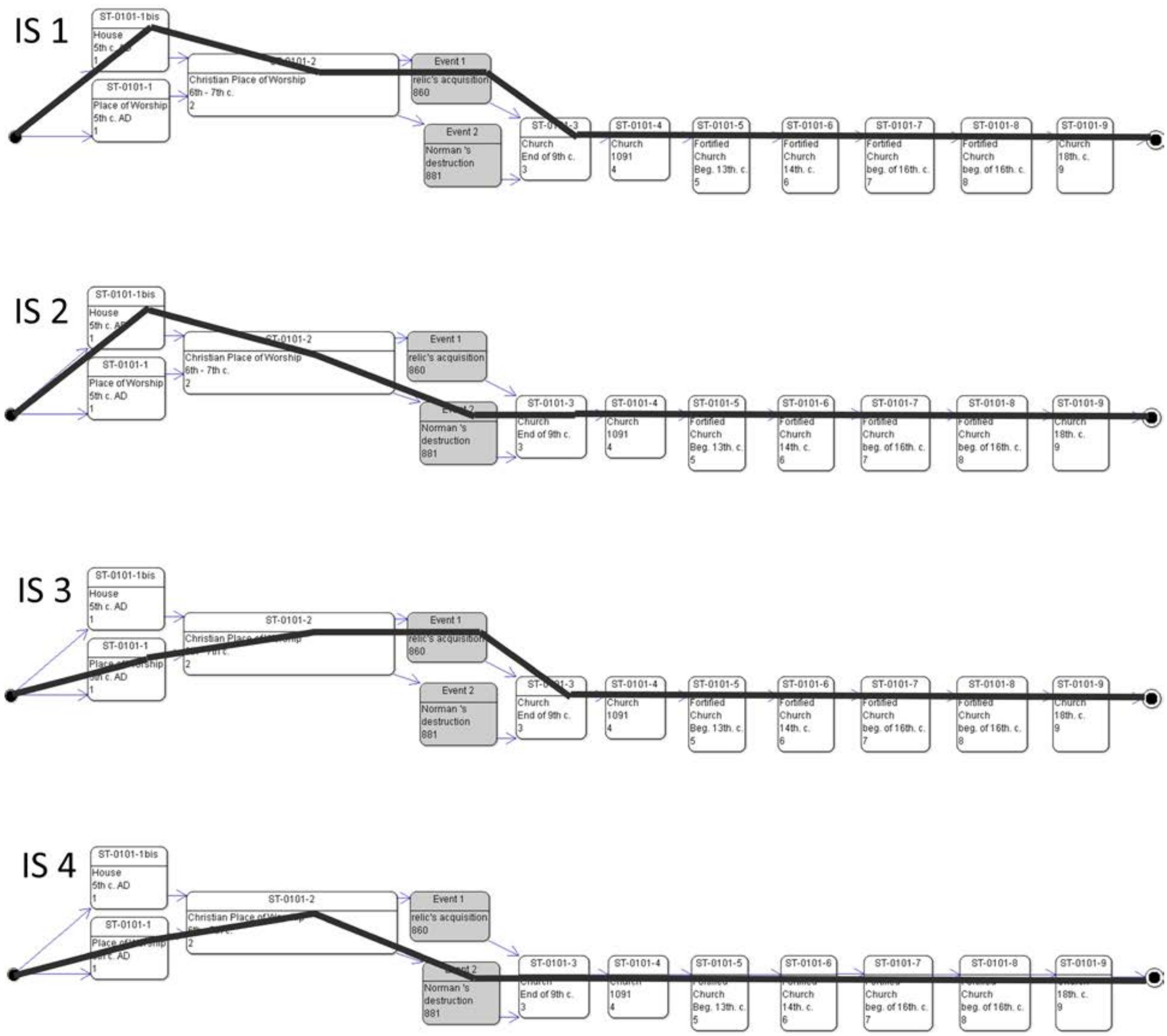


Figure 8: ST-0101 four Interpretative sequences.

approach the HO identity integrates henceforth not only the archaeological data characteristics but also the related contextual information in form of scientific interpretations by the use of *Interpretative sequence* and the concept of *Lifemap*. Besides we improved the definition of the event management by replacing our computational oriented vision by a phenomena descriptor vision. Due to these improvements and the addition of the *Episode* concept, a temporal indexing becomes possible.

Applying the model to the case study of the Weser showed that the model manages complex archaeological data. Nevertheless, further case studies need to be performed in order to test the consistency of the model. In the near future, we plan to implement the model. The database will be populated with multimodal data coming from the Weser project. Moreover, we plan to comply at existing modelling standards and to compile our system with Open Source software mainly for interoperability and persistence purposes.

References

BATS, M., BESSAC, J.- C., CHABAL, L., De CHAZELLES, Cl.-A., FICHES, J.-L., POULET, P., PY, M., *et al.* (1986). *Enregistrer la fouille archéologique : le système élaboré pour le site de Lattes (Hérault)*. Lattes: ARALO.

BERTHOLET, P. and HOFFSUMMER, P. (1986). *L'église-halle des Saints Hermès et Alexandre à Theux. Histoire et Archéologie d'un édifice singulier*. Dison.

BILLEN, R., BLAIN, P., DONNEAU, O., HABRAKEN, S., RENOTTE, Y., VAN RUYMBEKE, M. (2009). Virtual model of the city of Liège in the eighteen century - 'Virtual Leodium'. In: Verly, J. (ed.). *Proceedings of 3D Stereo MEDIA 2009. International 3D Stereo Film and Technology Festival (3D Stereo MEDIA*). *Liège, 1-3 December 2009.*

BILLEN, R., HEVY, B., LAROCHE, F., CARRÉ C., SERVIÈRES M., VAN RUYMBEKE, M., TOURRE,

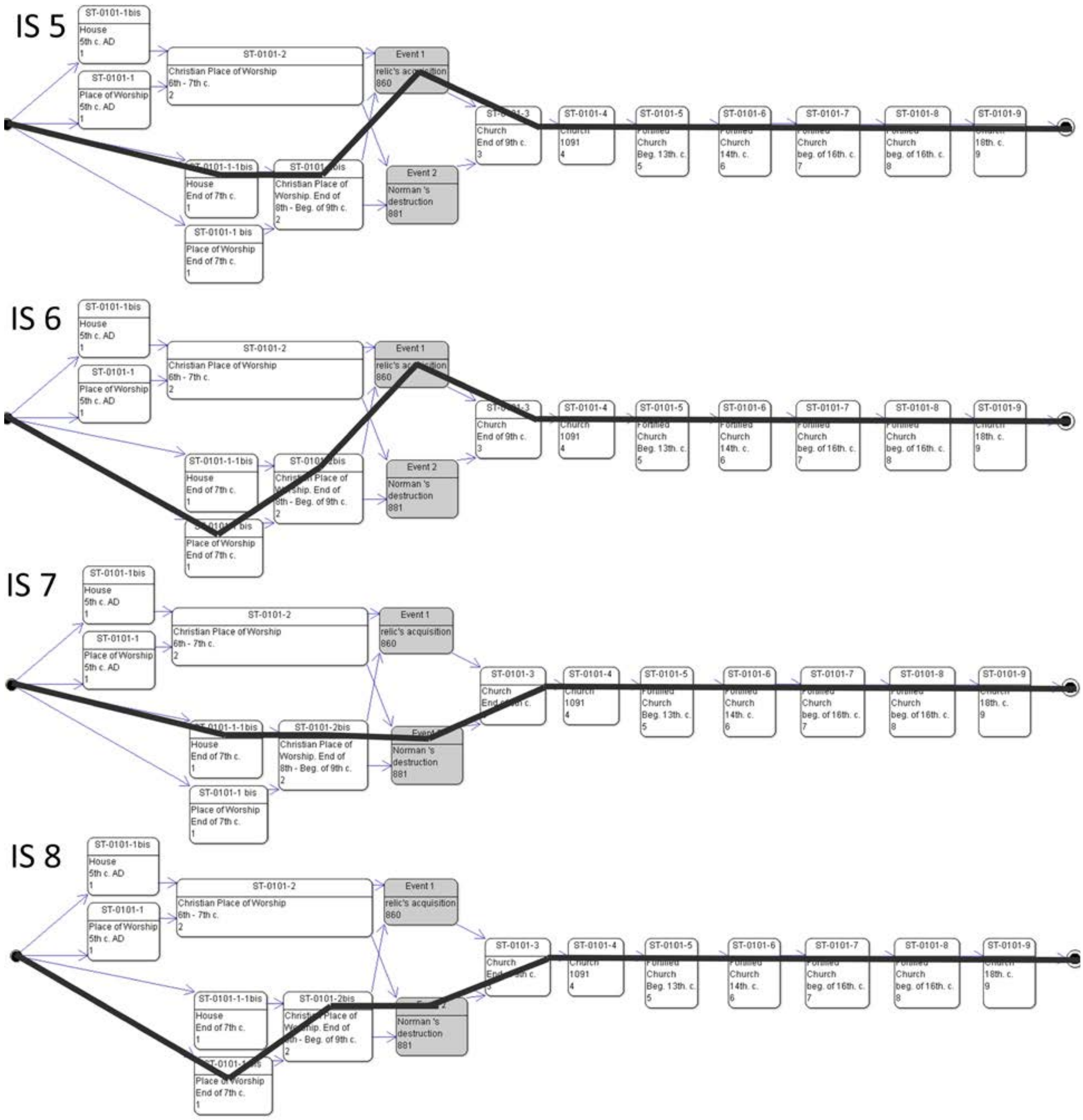


Figure 9: ST-0101 last LifeMap and last new interpretative sequences.

V. (2012). A generalized approach for historical mock-up acquisition and data modelling: towards historically enriched 3D city models. In: Leduc, T., Moreau, G., Billen, R. (eds.). *Usage, Usability, and Utility of 3D City models*, *Nantes 29-31 October 2012*. EDP Sciences DOI: http://dx.doi.org/10.1051/3u3d/201202009

BURROUGH, P. A. and MC DONNELL, R. (1998). *Principles of geographical information systems*. Oxford, United Kingdom: Oxford University Press.

DESACHY, B. (2012). *Formaliser le raisonnement chronologique et son incertitude en archéologie de terrain*. Cybergeo : European Journal of Geography. [Online] Available from: http://cybergeo.revues.org/25233. [Accessed : 12th July 2014.]

DESJARDIN, E., NOGENT, O., DE RUNZ, C. (2012). Prise en compte de l'imperfection des connaissances depuis le saisie des données jusqu'à la restitution 3D. *Archeologia e Calcatori* Supplemento. (3). p. 385 – 395.

DE RUNZ, C. (2010). *Imperfection, temps et espace : modélisation, analyse et visualisation*, France: Éditions universitaires européennes.

GALINIÉ, H., RODIER, X. and SALIGNY, L. (2004). Entités fonctionnelles, entités spatiales et dynamique urbaine dans la longue durée. *Histoire & Mesure*. XIX (3/4). pp. 223-242.

GONZALEZ-PÉREZ, C., PARCERO-OUBINA, C. (2011). A Conceptual Model for Cultural Heritage Definition and Motivation. In: *Revive the past: proceedings of the 39th Annual Conference of Computer Applications and Quantitative Methods in Archaeology (CAA)*, Beijing, 12-16 April 2011, Amsterdam: Amsterdam university Press, pp. 234 – 244.

GRÖGER, G. *et al.* (2012). *OGC City Geography Markup Language (CityGML) Encoding Standard*. Open Geospatial Consortium.

HARRIS, E.C. (1979). *Principles of Archaeological Stratigraphy*. London: Academic Press.

HORNSBY, K., EGENHOFER, M.J. (2000). Identity-based change: a foundation for spatio-temporal knowledge representation. *International Journal of Geographical Information Science* 14 (3). pp. 207-224.

KANT, E. (1869) *Critique de la raison pure. translated from german by Jules Barni*. [Online] tome 1 and 2, Paris. Available from: http://babel.hathitrust.org/cgi/pt?id=hvd.hwajhz;view=1up;seq=8

PILLOT, L., SALIGNY, L. (2011). The evolution of territorial occupation: exploratory spatial data analysis. Uncertainty and heterogeneity of data in Archaeology, In: *Revive the past: proceedings of the 39th Annual Conference of Computer Applications and Quantitative Methods in Archaeology (CAA)*, Beijing, 12-16 April 2011, Amsterdam: Amsterdam university Press, pp. 234 – 244.

PEUQUET, D. (1994). It's about time: a conceptual framework for the representation of temporal dynamics in geographic information systems. In: *Annals of the Association of American geographers*. 84, pp. 441 – 461.

PFEIFFER M., CARRÉ, C., DELFOSSE, V., HALLOT, P., BILLEN, R. (2013). *Virtual Leodium: from an historical 3D city scale model to an archaeological information system*. In: *International Committee for Documentation of Cultural Heritage (CIPA) CIPA. Symposium: Recording, Documentation and Cooperation for Cultural Heritage, Strasbourg 2-6 September 2013.*

PFEIFFER, M. (2011). *Approche critique de ' La noble Cité de Liège en 1730 ' à travers l'étude du collège des Jésuites en Île. Évaluation par application d'un outil archéomatique*. Liège : University of Liege.

RODIER, X., BARGE, O., SALIGNY, L., NUNINGER, L., BERTONCELLO, F. (2011). *Information spatiale et archéologie*. Paris : Errance.

RODIER, X., SALIGNY, L., LEFEBVRE, B. (2012). *La modélisation de l'information spatio-temporelle pour l'étude de la fabrique urbaine sur le temps long*. École thématique MoDys 2012, Fréjus.

VAN OSSEL, P., ARROYO-BISHOP, P.-D., LANTADA ZARZOSA, M.-T. (1988). *Manuel d'utilisation des fiches d'enregistrement informatisées du système ArchéoDATA pour les fouilles des Jardins du Carrousel*, Paris.

VAN WERSCH, L., MATHIS, F. and HOFFSUMMER, P. (2009). *Verres plats du haut Moyen Âge: église des Saints-Hermes-et-Alexandre à Theux* (Belgique). In: *Vitrail, verre et archéologie entre le Ve et le XIIe siècle, Actes de la table ronde réunie au Centre d'études médiévales d'Auxerre*, Paris, pp.161-170.

Used software:

Adobe System:
Adobe Photoshop, graphics editing program, http://www.adobe.com/en/products/photoshop.html.

ArcGIS Engine:
Esri, ArcGIS Engine, collection of GIS components and developer resources, http://www.esri.com/software/arcgis/arcgisengine.

ArgoUML:
UML diagramming application, http://argouml.tigris.org/.

Autodesk:
Autodesk Maya, 3D software, http://www.autodesk.com/products/autodesk-maya/overview/.

Geomagic:
Geomagic Studio, 3D software, http://www.geomagic.com/en/products/studio/overview/.

Chapter 4. Internet and Archaeology

Archaeological Open Access Journals: The Case of 'Archeologia e Calcolatori'

Alessandra CARAVALE, Alessandra PIERGROSSI

Istituto di Studi sul Mediterraneo Antico (ISMA) - CNR, Italy
alessandra.caravale@isma.cnr.it, alessandra.piergrossi@isma.cnr.it

Abstract

Our paper intends to provide an overview of archaeological open access journals, with particular reference to Italy, starting from the experience of 'Archeologia e Calcolatori'. Alongside published journals, on-line open access journals are increasing considerably, and are acquiring an important role in the publication of scientific results. 'Archeologia e Calcolatori' is included among the Italian journals in DOAJ (Directory of Open access Journals). This journal began in 1990 in paper format only, and since 2005 has joined the Open Archives Initiative (OAI) and is also published on-line. All articles published since 1998 are available in PDF format. The articles are shared in the circuit of the Open Archives, allowing harvesting from OAI service providers.

Keywords: Open access, Archeological on-line journals, 'Archeologia e Calcolatori'

In the twentyfirst century, the Internet has become the common environment for knowledge exchange even in Humanities and the possibility of sharing has been implemented in recent years through different channels. Compared to the traditional means, the *Innovation and Communication Technology* (ICT) is, both quantitatively and qualitatively, an added value in the knowledge, conservation, enhancement and enjoyment of the cultural heritage. As for other areas of research, also in archaeology the use of modern technology has greatly expanded the possibilities for research and lines of investigation and has led to new forms of learning and culture dissemination (Filippi, 2004, p. 235).

The first of these forms is that of open data, i.e., the sharing and use of data, such as files, catalogs, plans, photographs, inventories, excavation reports, which are produced in large quantities in the course of archaeological investigations, which are often stored for a long time in the archives, and therefore often remain unpublished for a long time. The on-line publication and the free reuse of the primary data is, now more than ever, an essential tool for the development of research and protection strategy of the cultural heritage and the territory.

Another area that immediately understood the great potential of the network was that of large computerized *corpora*. The network in this case inherited and amplified the work of computerization started in some cases since the early Seventies that only through its wide dissemination would find a suitable route. Among the large *corpora*, one can not fail to mention the Beazley Archive Project at Oxford University (http://www.beazley.ox.ac.uk/index.htm), which began in 1979 as the computerization of the archives of Sir John Beazley on Attic pottery, which has expanded over the years also to other classes of materials. Another major project is the LIMC France, which is now available through a website containing three different databases (http://www.limc-france.fr/).

The objective of this paper is to give a brief synthesis of one of the rapidly increasing uses of the web for sharing and spreading archaeological data and achievements: that of open access scholarly journals among which an important position is occupied by 'Archeologia e Calcolatori'.

The journals that adopt the open access editorial policy must respond to precise specifications laid down by the *Berlin Declaration on Open access to Knowledge in the Sciences and Humanities*, expressed in 2003 at the Max Plank Society (http://openaccess.mpg.de/286432/Berlin-Declaration), agreed over the years by nearly 300 scientific institutions and research organizations (Fig. 1):

1) Firstly, there must be free and universal access to data and their transmission, ensuring the correct attribution of intellectual authorship.

2) Moreover, this data should be housed in an on-line repository linked to an academic institution, a research institute, a government agency or a recognized organization that conforms to adequate technical standard.

This is a strategy aimed at removing every economic, legal, or technical barrier to the access to scientific information. The open access solution appears, ethically, as the most natural destination of the results obtained by the scientific research supported by public fundings from a perspective of democratisation of the scientific information and of control of the communication channels ownership by the academic and research world. It is also a winning response to the increasing prices of scientific journals, whose production and distribution are concentrated in the hands of few and powerful commercial publishers, often multinationals.

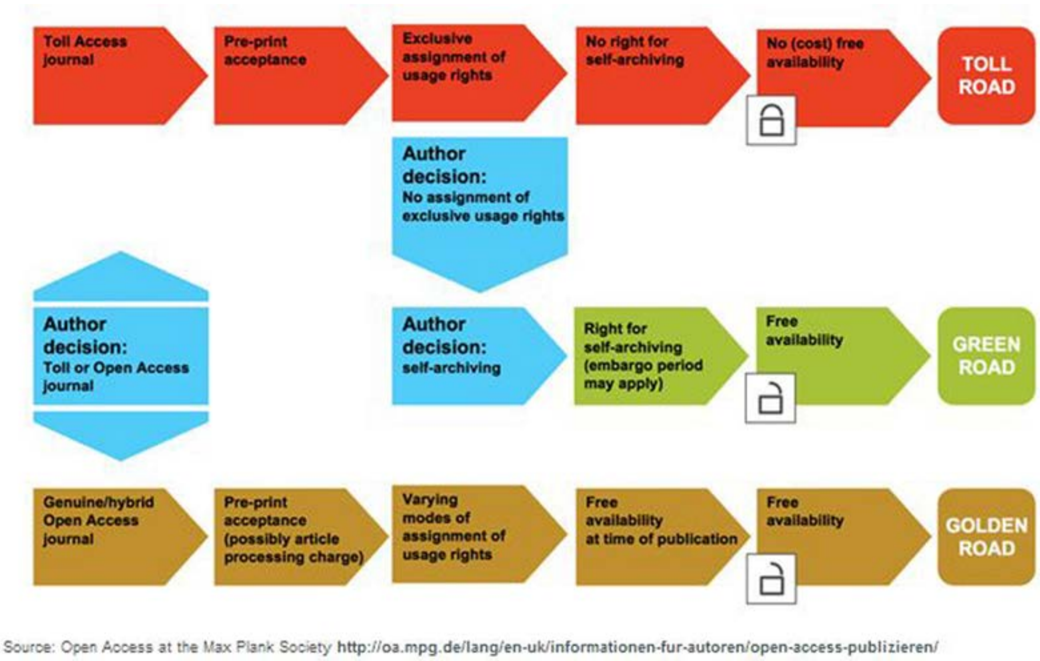


Figure 1: Open access scheme (from http://www.clib-jena.mpg.de/lis/?page_id=138/).

There are two options to open access publication: the golden road and the green road. In the first case, scholars publish their work, while keeping the copyright, in new electronic peer reviewed journals or already existing periodicals converted from the 'toll access' model, without charging readers for access. By following the green road, the scholars can instead deposit in institutional or disciplinary repositories pre-print versions (a practice named self-archiving), without the need of any authorization, or even post-print articles, that appear in paper in parallel and in accordance with the publishers policies of copyright, at times subjected to a period of embargo. It is important that these repositories are built on Open Archive Initiative protocols, so that they can be easily interoperable and usable as a single large virtual archive through search engines.

Contrary to the recent great diffusion of open access publications in scientific technological and medical disciplines, the use of open access in Human Sciences is growing more slowly, because of both the shortage of funding and because the necessity and the urgency to make the results more widely accessible is considered less important in Humanities, more orientated towards the realization of monographs which confer greater prestige and provide titles that can be used in the academic career.

All this adds up to a certain degree of inertia and resistance to change in the researchers community especially in the context of human sciences and classical studies that, at least in part, still have doubts about the credibility and reliability of the digital format and prefer to move towards the traditional channels of the paper output.

At the moment it is possible to distinguish among the archaeological journals:

– those that have been recently established in electronic format;
– those which, while also printed on paper, have put all the texts in their entirety available on-line;
– those, the most numerous, that are generally journals of long tradition, which have transmitted on-line only part of their content;
– and the others only available by paying a subscription fee or adopting a pay per view access to single articles.

For the fully open access journals an important benchmark is the web site Directory of open access journals (DOAJ: http://www.doaj.org/), the repertoire of peer-reviewed electronic journals freely accessible on-line without embargos or any other restrictions, created by the University of Lund. This currently includes more than 9700 magazines, working in different disciplines, including academic journals or others that are subject to quality control by a scientific committee. The purpose of this initiative is to increase the visibility and promote the use of the quality open access resources. In 2012 a survey was published in 'Archeologia e Calcolatori' where it was possible to measure the success and the important growth of open access editorial sector (Caravale & Piergrossi, 2012). In about a year the number of the periodicals increased by 1700 titles, with regard to archeology the increase was from 39 to 54 journals, 8 of which Italian.

'Archeologia e Calcolatori' (http://soi.cnr.it/archcalc/) (Fig. 2) is the oldest of the Italian journals in DOAJ. Founded in 1990 by Mauro Cristofani (director of the Institute for Etruscan-Italic Archaeology, now Institute for Ancient Mediterranean Studies of the National Research Council) and Riccardo Francovich and directed by Paola Moscati, it was conceived as an international observatory of theoretical and methodological aspects of computing

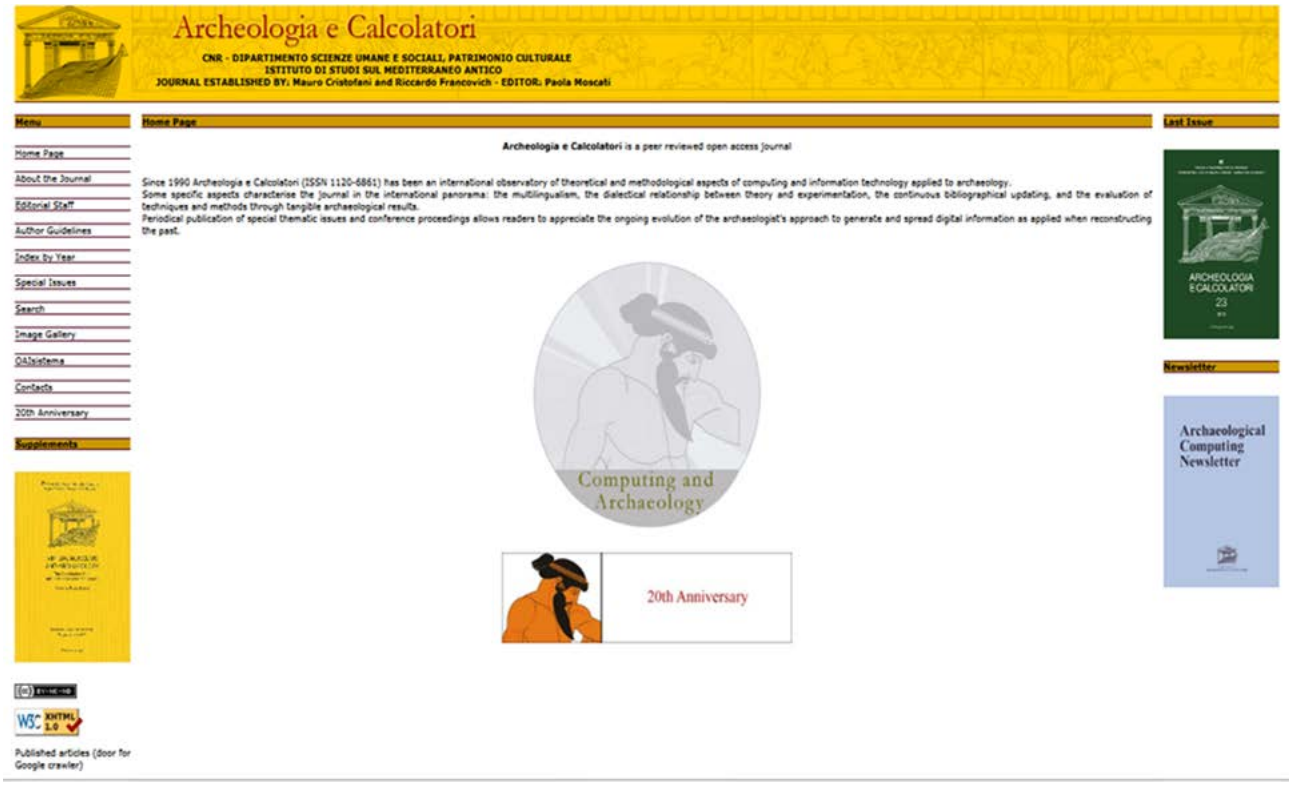


Figure 2: 'Archeologia e Calcolatori' homepage.

and information technology applied to archaeology. The Institute directed by Mauro Cristofani since the Seventies paid specific attention to the so-called 'ancillary sciences' of archaeology and to the methodological renewal that was investing archaeology, which was beginning to apply new tools for research coming from natural and social sciences. With regard to the application of computer techniques for the study and classification of archaeological artifacts, at the beginning of the Eighties our Institute gave rise to a dedicated line of research, which found in the international journal 'Archeologia e Calcolatori' a stable editorial point of reference and a place of convergence for the theoretical and methodological debate.

As it is clearly illustrated in the journal's website, some specific aspects characterise the journal in the international panorama: the multilingualism, the dialectical relationship between theory and experimentation, the continuous and steady bibliographical updating, and the evaluation of techniques and methods through tangible archaeological results.

'Archeologia e Calcolatori' is at the moment a peer-reviewed journal controlled by an international Scientific Committee and a qualified Editorial board which ensures a continuous flow of information and a profitable exchange of data.

The journal covers three distinct parts:

1. The first looks at methodological approaches: it collects articles concerning theoretical aspects of archaeological computing as well as reports on programs conducted by dedicated international institutions.
2. The main section contains articles on various computer applications, such as databases, Geographical Information Systems, quantitative methods, expert systems, computer graphics, image processing, multimedia and web tools. A section is also dedicated to the automatic processing of documentary sources.
3. The third aspect of the journal is characterized by book reviews and bibliographic news, with the aim to provide readers with an up-to-date source of documentation.

Looking at the statistical analysis (Figs. 3-5) compiled on the occasion of the 20th anniversary on the content of the journal, it is possible to retrace the development of archaeological computing.

The application on statistical analysis techniques appears to be one of the more important issues that has a fairly constant rate of recurrence. Contributions discussing databases were more frequent in the 1990's and focused upon data storage and retrieval. Only in recent years an interest in web-based database applications has become apparent. The same can be observed for simulation and Artificial Intelligence applications, which seem to be more frequent in the 1990's issues, and less popular in the following decades.

On the other hand, articles dealing with applications such as Remote Sensing and GIS were less frequent in the early 1990's and more popular in the following years, as

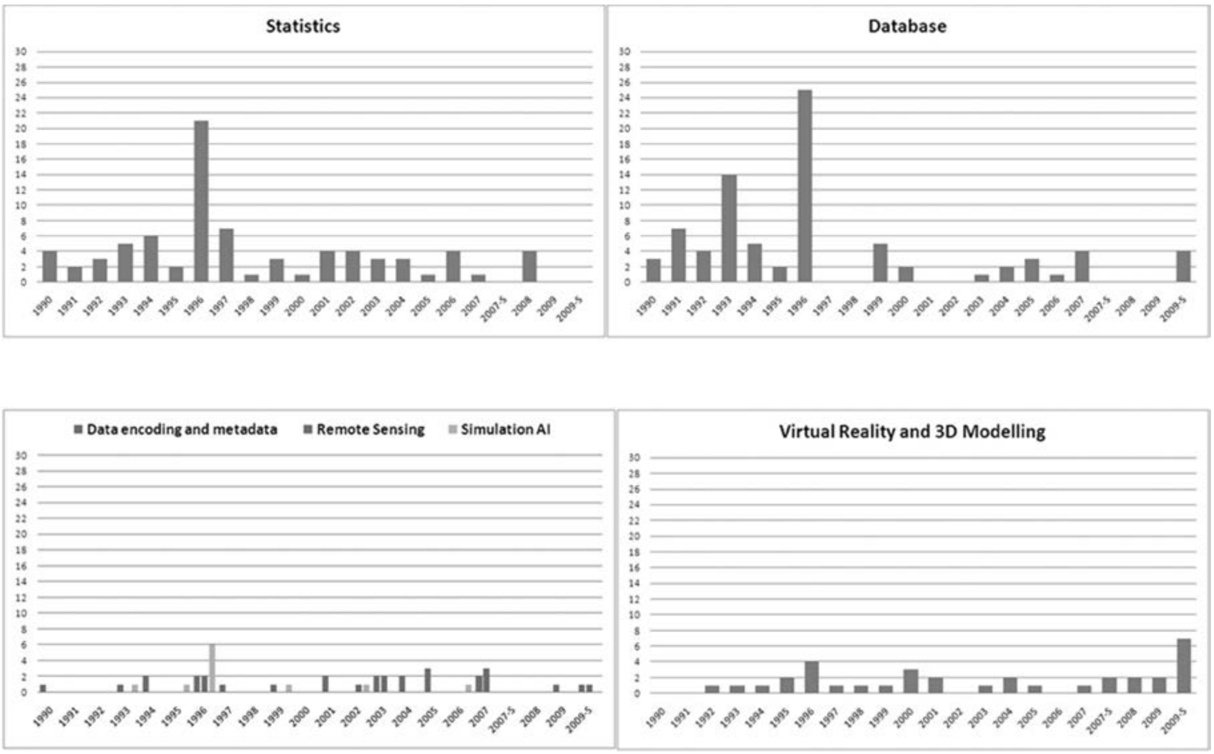


Figure 3: 'Archeologia e Calcolatori': frequency analysis of the archaeological subjects.

a reflection of technical developments. The same can be observed for contributions about Virtual Reality that has dramatically increased in the last years.

The journal's focus on methodological problems is mainly represented by articles on the history of applications, which had its peak in the 20th issue (Moscati, 2009), dedicated to the launch of archaeological computing. The frequency analysis of the archaeological subjects discussed in the journal reveals that there has been a fairly constant number of articles on theoretical and methodological problems, which is one of the main goals of 'Archeologia e Calcolatori'. The same can be said for the two topics of Cultural Resource Management and data dissemination and education. Other subjects which display a higher proportion of contributions throughout the years mainly involve research work conducted in the field (survey and excavations) – one of the most prevalent areas of application – or in the laboratory (classification of archaeological finds).

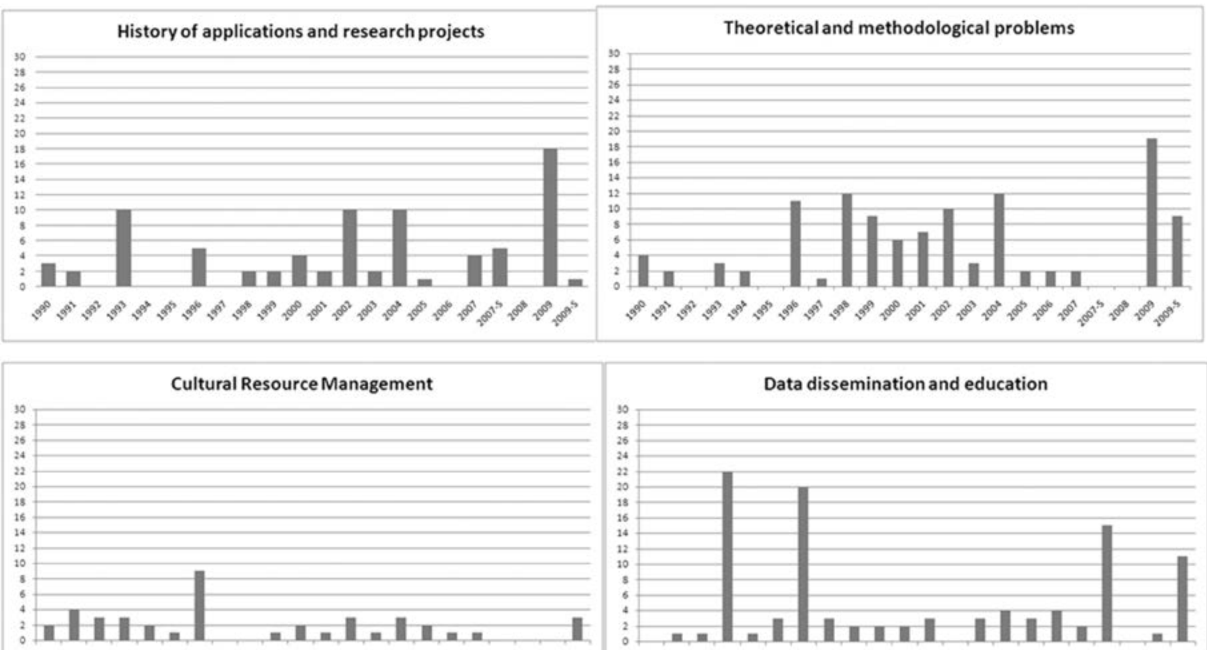


Figure 4: 'Archeologia e Calcolatori': frequency analysis of the archaeological subjects.

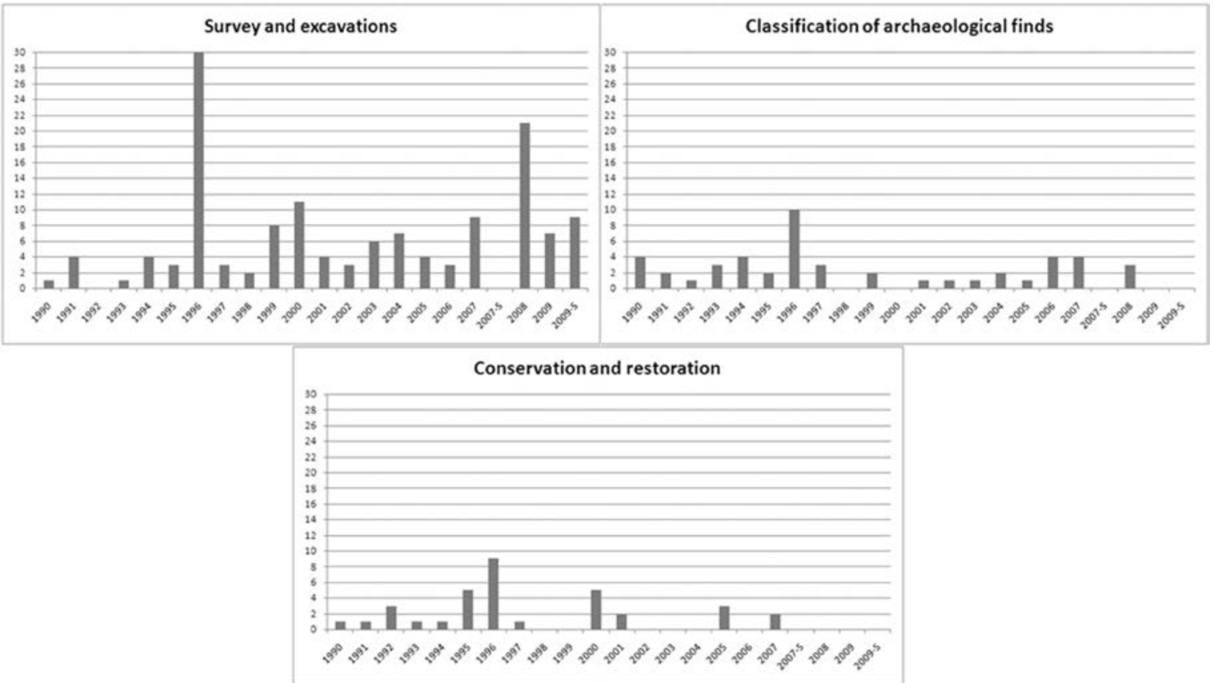


Figure 5: 'Archeologia e Calcolatori': frequency analysis of the archaeological subjects.

With its open access policy, 'Archeologia e Calcolatori' can be considered not only an editorial point of reference, but also an investigative tool for scholars and students. A tangible proof is offered by the implementation of the 'Virtual Museum archaeological computing' (http://archaeologicalcomputing.lincei.it/) – a research project conducted by ISMA together with the Accademia Nazionale dei Lincei – in which the articles published in the journal constitute a resource immediately available for the documentation of all different aspects of the history of this discipline. The journal acts as a tool for research and updating and its pages have seen the evolution of digital cartography, the birth and development of geographic information systems, the use of virtual reality techniques to document and represent territories, cities and monuments covering a wide chronological and geographical perspective, from Italy, to Europe and the other continents. Some articles have also explored issues related to urban and land planning and preventive archeology.

As a result of some experiences gained within the Institute for Etruscan Italic Archaeology related to text encoding, web sites and knowledge sharing, since 2005 in addition to the paper format, 'Archeologia e Calcolatori' is also on-line, providing all articles published since 1998 in pdf downloadable and printable.

To this aim the journal is provided with the Creative Commons license Attribution-NonCommercial-NoDerivs (https://creativecommons.org/). This license is the most restrictive, and allows only to download and redistribute works in any medium or format, as long as license terms are followed and appropriate credts are given to authors; nothing can be changed in any way and articles cannot be used for commercial purposes. The adhesion to the Open access philosophy is closely connected to the policy of the journal, which has always been attentive to the formalization of languages and multimedia communication, aspects that have certainly provoked a technological and at the same time cultural change in archaeological research.

The strength and uniqueness of the journal is in the duplicity between tradition and innovation: innovation linked to its direct experimental commitment aimed at the adoption of cutting-edge technologies for the diffusion of scientific content on the network. The uploading to the web of the journal, also thanks to the far-sighted vision of the publisher All'Insegna del Giglio, has been designed through the creation of a repository where resources are described by Dublin Core metadata complying the OAI-PMH scheme (Open Archives Initiative-Protocol for Metadata Harvesting). The OAI system is an application that allows one to easily administer the collection of 'Archeologia e Calcolatori', to share them in the circuit of the Open Archives and to allow the harvesting from the OAI service provider. The 'Archeologia e Calcolatori''s repository was signalled in 2005 in the OAI mailing list as 'the first Static Repository data-provider registration (…) which uses Patrick Hochstenbach's gateway at Ghent University' and already in 2006 a study of the Los Alamos National Laboratory, examining 10 OAI repositories, found that the journal was in first place regarding the indexing easiness in search engines on the web.

Items are also deposited in the database of CNR SOLAR (Scientific Open-access Literature Archive and Repository, http://eprints.bice.rm.cnr.it/) and can also be retrieved from the PLEIADI portal, the national platform for centralized access to scientific open access literature (http://www.openarchives.it/pleiadi/).

In addition to the OAI support, the application developed for the journal comprises two other web services. The first

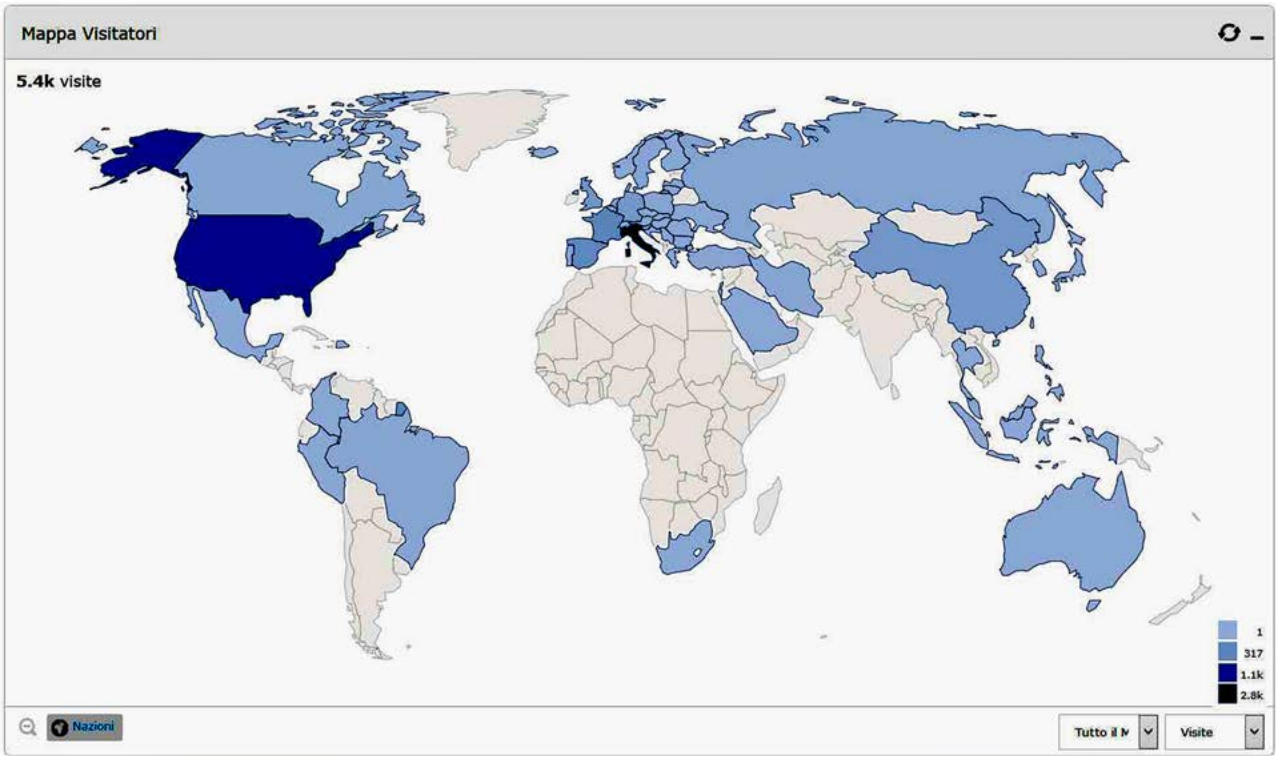


Figure 6: 'Archeologia e Calcolatori' website: accesses' areas of origin.

is made especially for the Google spider: each record in the repository has a static URL associated, reachable through only three hyperlinks from the homepage of the journal (Year, Volume, Article), so the repository can be easily indexed by Google. The second service is a local search engine accessible from the homepage of the website, which allows users to search by author, subject, year, and free text in the metadata for each article. In conclusion, every electronic resource in the repository can be traced through three different paths: OAI, Google and journal website.
From April to June 2014 the counter has detected between 30 and 50 hits per day, with peaks (up to 167) at the time of the on-line publication of the new issue 24. By analyzing the accesses' areas of origin, we observe that in addition to a constant interest from Europe and United States, an increasing attention appears from Asian countries such as China (Fig. 6).

Almost every five years special issues are published that highlight some fundamental stages in the evolution of the archaeological computing, as a basis for a coherent and consistent development of the discipline. The titles are certainly indicative: Choice, Representation and Structuring of Archaeological Information appeared in 1994 (Moscati, 1994), Methodological Trends and Future Perspectives in the Application of GIS in Archaeology in 1998 (Moscati, 1998), New Frontiers of Archaeological Research. Languages, Communication, Information Technology of 2004 (Moscati, 2004). Since 2007 we can also add the publication of supplements such as Virtual Museums and Archaeology (Moscati, 2007), the first and second edition of the meeting of ARCHEOFOSS. Open Source, Free Software e Open Format in the processes of the archaeological research, published respectively in 2009 and 2013 (Cignoni, Palombini & Pescarin, 2009; Serlorenzi, 2013) and the Proceedings of the second and the third meeting d'Informatique et Archeologie de Paris held in 2010 and 2012 focused on standards and rules in archaeology, GIS, and Virtual Reality (Giligny *et al.*, 2012; Costa, Djindjian & Giligny, 2014). The classification by standards and Dublin Core metadata adopted for the journal has also been applied to the articles of the supplements, to which a specific page in the website is now devoted.

A useful tool for archaeological computing, as an autonomous discipline, is the bibliography published in the course of the first ten years of the journal's life, designed to overcome the difficulties of updating and finding sources of information. This bibliography has been poured into a database that will be soon available on-line on the site dedicated to the Virtual Museum of the Archaeological Computing mentioned above.

Data can be queried through different search keys. In addition to the traditional fields such as author, title, date and place of publication, etc., the database has been enriched with the field 'computer applications' that describes the tools of information technology used in research and in data treatment, and a 'subject' field that identifies the area of the archaeological research discussed in the contribution. The bibliographic records, which have reached approximately 2500, will be accompanied by links to the texts available on web, allowing a fast direct access to the resource.

In conclusion the choice to adhere to the open access initiative, also strongly wished by the European Commission Recommendations in July 17, 2012 (http://ec.europa.eu/research/science-society/index.cfm?fuseaction=public.topic&id=1294&lang=1/), at the moment still encounters many difficulties and in the case of 'Archeologia e Calcolatori' the decision to maintain both formats, on-line and on paper in the formula print-on-demand, clashes constantly with the retrieval of the necessary resources and to persevere in the project is a struggling effort that increases year after year.

References

CARAVALE, A. & PIERGROSSI, A. (2012). Archeologia in rete. Le riviste open access: risorse e prospettive. *Archeologia e Calcolatori*. 23, p. 187-207. Available from: http://soi.cnr.it/archcalc/indice/PDF23/11_Caravale_Piergrossi.pdf. [Accessed: 22nd July 2014]

CIGNONI, P., PALOMBINI, A. & PESCARIN, S. (eds.) (2009). Archeofoss. Open Source, Free Software e Open Format nei processi di ricerca archeologica. *Archeologia e Calcolatori*. Supplemento 2.

COSTA, L., DJINDJIAN, F. & GILIGNY, F. (eds.) (2014). *Actes des 3èmes Journées d'Informatique et Archéologie de Paris – JIAP 2012, Paris, 1-2 juin 2012. Archeologia e Calcolatori*. Supplemento 5.

FILIPPI, F. (ed.) (2004). *Manuale per la qualità dei siti web pubblici culturali*. Roma: Ministero per i Beni Culturali, Progetto Minerva.

GILIGNY, F., COSTA, L., DJINDJIAN, F., CIEZAR, P., DESACHY, B. (eds.) (2012). *Actes des 2émes Journées d'Informatique et Archeologia de Paris – JIAP 2010. Archeologia e Calcolatori*. Supplemento 3.

MOSCATI, P. (ed.) (1994). Choice, Representation and Structuring of Archaeological Information. *Archeologia e Calcolatori*. 5.

MOSCATI, P. (ed.) (1998). Methodological Problems and Future Perspectives in the Application of GIS in Archaeology. *Archeologia e Calcolatori*. 9.

MOSCATI, P. (ed.) (2004). New Frontiers of Archaeological Research. Languages, Communication, Information Technology. *Archeologia e Calcolatori*. 15.

MOSCATI, P. (ed.) (2007). Virtual Museums and Archaeology. The Contribution of the Italian National Research Council. *Archeologia e Calcolatori.* Supplemento 1.

MOSCATI, P. (ed.) (2009). La nascita dell'informatica archeologica. *Atti del Convegno Internazional, Roma, Accademia Nazionale dei Lincei, 24 ottobre 2008. Archeologia e Calcolatori*. 20.

SERLORENZI, M. (ed.) (2013). ArcheoFOSS. Free, Libre and Open Source Software e Open Format nei processi di ricerca archeologica. *Atti del VII Workshop, Roma, 11-13 giugno 2012. Archeologia e Calcolatori*. Supplemento 4.

Massive Open Online Opportunity: MOOCs and Internet–Based Communities of Archaeological Practice

Jeffrey P. EMANUEL

Harvard University, USA
jemanuel@fas.harvard.edu

Abstract
While the World Wide Web has provided the public with heretofore-unimagined access to information, the democratization of online content creation has also provided an unprecedented opportunity for the spread of misinformation and misinterpretation. Archaeology is no exception, as developments like the exposing of museum collections, the ability to conduct armchair 'surveys,' and unfettered access to uncontextualized images via simple Web search have combined to confront a new generation of avocational and aspiring archaeologists with myriad explanations and interpretations of artifacts, archaeological data, and history writ large. The rise of MOOCs (both as 'massive open online courses' and as repositories for massively-accessible online content) may help combat this by providing a structured mechanism for practitioners to reach, interact with, educate, and learn from an ever-growing online audience. This is of particular importance for archaeology, a field in which standards of conduct and interpretation are keys to sound and ethical practice.

Keywords: Massive, MOOC, Online Applications, Pedagogy, Communities of Practice

1. Information Flow in the Social Age

In 2013, 2.77 billion people worldwide were classified as 'Internet Users,' as defined by the International Telecommunications Union (ITU 2010; 2013; Figure 1). The rise of the Internet age has provided an ever–increasing worldwide population with heretofore–unimagined access to information, putting the combined knowledge of the world literally at the fingertips of over one third of the planet's population. This, in turn, has allowed for questions to be answered, knowledge to be transmitted, and social endeavors to be engaged in at lightning speeds in some of the most remote corners of the globe. While the World Wide Web is an unparalleled tool of information dissemination, though, perhaps the most noteworthy aspect of the explosion of Internet access and web–based technologies that has taken place in the 21st century is the near–eradication of the one–to–many model that once monopolized Internet design and usage, particularly in the overlapping academic, research, and education spheres. Instead, people both shape the information available on the Internet, making its contents a conversation rather than a tool for broadcast, and engage in direct communication with their fellow users via various tools available to them (*inter alia* Matzat, 2004; Burgess and Green, 2009; Bohn *et al.,*, 2014).

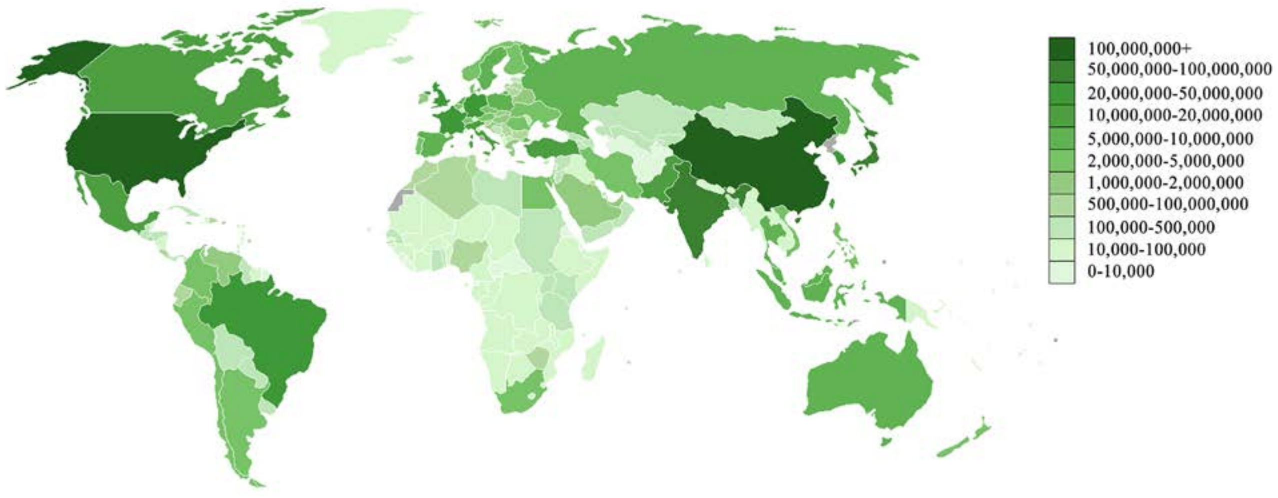


Figure 1: Density map of Internet users worldwide, based on data from the International Telecommunications Union.

Along with the spread of internet–based information has come the opportunity for internet–based content creation, both via well–known collaborative enterprises like Wikipedia, and via lesser–known, but also potentially influential, platforms like social media accounts and personal blogs – all of which are available for free to anybody with an Internet connection and a compatible device (Daugherty, Easton and Bright, 2008; Correa, 2010; Blank, 2013). Unfortunately, the egalitarian – and frequently anonymous – nature of online content creation can lend itself to the spread of misinformation and misinterpretation as much as it can to the creation of reliable content and commentary (e.g. Huffington Post, 2013).

Archaeology is far from immune to this side effect of the 21st century web. So–called 'Adventures in Bad Archaeology' abound on the Internet and beyond (see badarchaeology.com, specialtyinterests.net, biblebabel.net, or watch a few episodes of the History Channel's 'Ancient Aliens' series for just a few of many examples!). In light of this, any sound–minded practitioner could be forgiven for reacting to such a twisting of our profession with a combination of horror and a renewed commitment to preventing archaeological information from reaching the general public prior to being surrounded by comprehensive, carefully–researched and –crafted contextualization and explanation. However, like any other subject or pursuit, archaeology in the 21st century cannot exist in isolation, remaining a tightly guarded secret until each new data point has been mummified inside an impenetrable wrap of contextualization and explanation that would provide non–specialists with one part raw data and ninety–nine parts direction on how to view, think about, and use those data. Fortunately, as this session's very existence demonstrates, many within our field have no intention of holding up archaeology, and the data points archaeologists recover, as the property of specialists alone. This is fortuitous for many reasons, not least of which is the simple fact that the nature of modern communication and the platforms available for content creation combine to make such a move all but impossible.

The nature of the 21st century Web, and of human communication in the Internet age, is one of openness and personal contributions to the discussion surrounding whatever subject an individual may be interested in. This, combined with the opening of museum collections, the ability to conduct armchair 'surveys' via tools like Google Earth, unfettered access to uncontextualized images via simple Web search, and similar developments, has ensured that the web will continue to abound in perpetuity with myriad explanations and interpretations of artifacts, archaeological data, and history writ large. Given this, the last thing we should do is pull back from the public arena of ideas and discourse, as any attempt by specialists in archaeology or any other field to withhold information – even if it is done with the purest of intentions – will in effect cede the battlefield of information and ideas to everybody else.

2. Embracing the Open Internet

Many One partial solution to this conundrum may be found in a location that might seem, at first blush, to be an unlikely place to look for a savior. I speak here of the MOOC, or the 'Massive Open Online Course' – an acronym and term which is increasingly disinclined toward that which it purports to describe. As an offering to the public in the free and open space, the MOOC can serve as a structured mechanism for professionals across the academy to reach, interact with, educate, and learn from an ever–growing online audience (cf. Ho *et al.,* 2014; for an advocacy perspective, see Zapatero and Morales 2014). This is of particular importance for archaeology, a field in which standards of conduct and interpretation are keys to sound and ethical practice.

Though no single MOOC will likely ever compete with the overall traffic of a site like Wikipedia, the presence of these learning experiences in the free and open space is a resounding positive for open access to information, and sends a strong signal that specialists are not prepared to withdraw from the discussions surrounding their areas of expertise and cede the arena of ideas, once again, to everybody else.

The open, inclusive nature of MOOC–based learning experiences can allow them to coexist and compete with similarly free and open sources of information about archaeological topics that are broadly accessible on the public Internet. Further, in the MOOC environment, experts leading open online learning experiences can draw in new participants, while simultaneously ensuring that the facts, techniques, and practices conveyed in their particular learning experience represent accurate scholarly interpretation and understanding, as well as the most up–to–date professional standards and methods.

The multidirectional communication that is characteristic of today's Internet is also a critical role–player in the success of MOOCs as developers and perpetuators of communities of practice. Successful participants, for example, may go on to serve as international and intercultural hubs from which accurate, professionally–conveyed information can flow outward to various peripheries. In addition, the multicultural nature of MOOC audiences may also serve as a mechanism for improving the professional practice of archaeology, in part by creating a feedback loop via which practitioners can be exposed to viewpoints and cultural interpretations that might not be commonly considered.

Now, of course, we come to the most often voiced criticism of MOOCs: the abysmal 'retention' and 'completion' rates they boast, and the correspondingly low percentage of total registrants which goes on to earn a certificate at the end of each learning experience (Parr, 2013; Perna *et al.,*, 2013; Ho *et al.,*, 2014; Hollands and Tirthali, 2014; Lauerman, 2014; Levy, 2014). The outsized attention paid to these data points, though, only reinforces how inapplicable they are to the present space (Ho and Reich, 2014).

In many ways, higher education has remained static for centuries. Our understanding of it is fueled by a shared vocabulary, and by a common understanding of expectations. College courses are formal engagements, which are quarters, semesters, or years in length. They are offered synchronously, with a beginning date, an end date, and a cutoff for registration and honorable withdrawal. They are undertaken by students who have completed an official registration (and payment) process – a process that includes contractually agreeing to participate that course from start to finish, and that lays out the rewards for success and the consequences of failure. In the case of the former, a good grade, college credit, and eventually a certificate or diploma awaits, while the latter offers a negative mark which will remain on the student's transcripts for the rest of his or her life, affecting every other academic undertaking he or she pursues (cf. Ho and Reich, 2014; Emanuel, forthcoming).

Viewed through this lens, a combined withdrawal and failure rate of twenty percent would be highly suspect to say the least, while a rate of nine in ten meeting this description would rightly be viewed as an unmitigated disaster – a breach of contract by both the professor and the institution offering this course. However, in the open online world, where the only barrier to entry is access to an Internet connection and the willingness to provide an email address, this same rule simply does not – and cannot – apply. Terms and concepts like 'registration,' 'retention,' 'completion,' and even 'course' lose their traditional meaning, and their application here makes little more sense than using maritime terminology to describe modern interstate travel.

A more relevant comparison would be other open online resources that offer engagement, track participation and contribution, and pose similarly low barriers to entry and contractual obligations for the user. For a frame of reference, we may consider some basic metrics from the Massive Open Online Course 'The Ancient Greek Hero' (CB 22x), a learning experience produced by Harvard University (HarvardX) that focuses on the close reading of ancient Greek texts and vase paintings. 43,563 people registered for this learning experience (which was referred to internally as a 'Project' rather than as a 'Course'). Registration, which was possible up to three months prior to the Project's official start date, consisted of to a user providing edX with their email address, viewing the CB 22x information page, and clicking 'register for this course.' Of the 43,563 registrants, 25,686 (59 percent) were active participants, meaning that they actually logged in to the Courseware and accessed the available materials once the Project was live (Reich *et al.,* 2014). 3.2 percent of total registrants (and 5.5 percent of active participants) completed the necessary requirements to earn certificates of completion (ibid; Figure 2).

Now let us compare these numbers to an open online community that has certainly affected the way that we as scholars and researchers, and the way our students, do business: Wikipedia. Using September 2013 as sample,[1] we see that a little over 116 million people visited the site during this period (Wikipedia, 2014). Of those, the number of 'active participants' – people who made five or more edits to Wikipedia's vast array of articles – was 7,789, or a little under seven thousandths of one percent. The number of active contributors – people who made 100 or more edits – was 778, or six ten–thousandths of one percent (ibid). This leaves 99.993% of visitors to be classified as non–participatory and non–contributory (Figure 3) – a statistic which starts to make participation in 'MOOCs' sounds a little better!

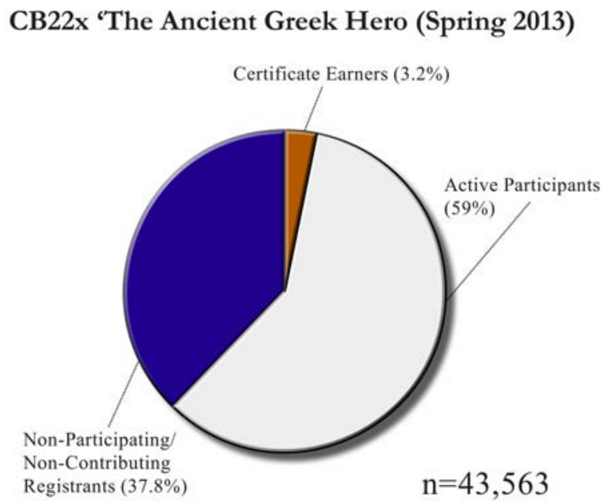


Figure 2: Certificate earners, active participants, and non-participatory and non-contributory enrollees as percentages of overall CB22x enrollment (illustration by the author, based on data from Reich *et al.,* 2014).

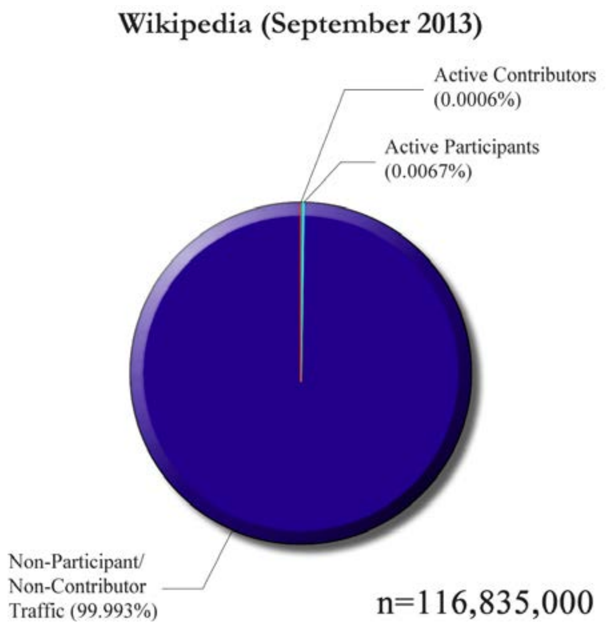


Figure 3: Active participants and contributors versus non-participatory and non-contributory web traffic to Wikipedia.org, 1-30 Sept. 2013 (illustration by the author, based on data from Wikipedia 2014).

[1] At the time of this writing, September 2013 is the most recent month broken down in detail on Wikipedia's statistics page.

As several studies and visualizations have shown, the seven–thousandths of a percent which actively participates – and especially that seven–ten–thousandths of a percent which actively contributes – can be very active in engaging the content and in engaging each other, bringing a remarkable level of interactivity that active contributors bring to a subject matter about which they are passionate (Forte and Bruckman, 2006; Viégas *et al.,* 2007; Viégas and Wattenberg, 2014; Wagner and Prasarnphanich, 2007; Yang and Lai, 2010).

This is by and large the nature of social interactions and interest–based groupings in the online world, and we can see it borne out in MOOCs, as well. The graph below shows the level of successful engagement with assessment exercises charted against the number of chapters of content viewed within CB22x (Reich *et al.,* 2014: fig. 5). As you can see, user types – and, we can infer, user intent – covers the entirety of the spectrum, from non–participatory, non–contributing registrant through the learner who viewed every chapter of content provided and aced every assessment. An excellent case in point on the different modalities of engagement available to participants is what we call 'the listener,' an identification which at its extreme refers to a participant who engaged with every chapter of content but none of the assessments – in other words, a person who truly came to learn and to interact, but without the desire to 'earn a certificate' (compare to readers vs. active participants in Wikipedia; Antin and Cheshire, 2010).

This should provide the impetus for our looking at MOOCs in a different light. Rather than 'courses' whose 'retention' and 'completion' rates we count like beans, the significant differences in user intent and modality of participation should be taken into account when creating, running, and evaluating MOOCs.

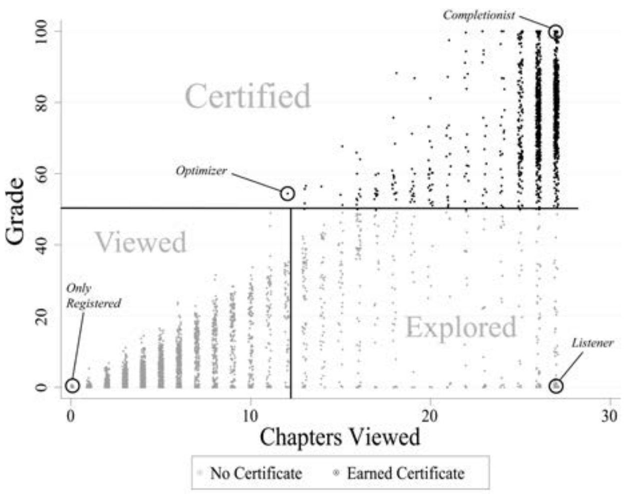


Figure 4: Scatter plot showing the level of successful engagement with assessment exercises charted against the number of chapters of content viewed by enrollees in CB22x (after Reich *et al.,* 2014: fig. 5).

This is also consistent with the move among some to change C in the acronym 'MOOC' from 'Course' to 'Content,' and to consider these objects for what they might be more accurately described as: massive open multimedia textbooks, with – in many cases – a significant social component to them, which can be leveraged to create both feedback loops within, and hubs for accurate information transmission beyond, these learning experiences (cf. Fowler *et al.,* 2014; Emanuel, forthcoming).

Finally, the rise of the MOOC is helping to provoke serious thought about just how our field and many others can be properly communicated to a broad, international, and truly diverse population of learners, enthusiasts, and future professionals (cf. Dufton, Durusu–Tanrıöver and Alcock, 2014). This will hopefully grow into acceptance and consideration of the multiple modalities of participation that consumers of MOOC content desire, but it has already served as an opportunity to invest in better ways for we as archaeologists to communicate our story. In other words, it has spurred investment in better tools and approaches for the conveyance of archaeological understanding and practice, and has provided a means for interested persons to congregate around the topics that make our field so fascinating, while also allowing them to disseminate that information themselves, within their own circles both online and off.

These tools include, for example, media–rich annotation of all modes of content presentation, including text, image, and video (Desenne, 2014; Desenne and Reis–Dennis, 2014; Mondenero Moya *et al.,* 2014), as well as expanded 3D imaging of objects (e.g., inter alia, Zhou *et al.,* 2012: 13–140; Sanders, 2013; Emanuel, 2014; Manuelian, 2014) and vastly improved interactive viewers that can be used to engage with objects of all types within our museum and library collections and beyond (Harward *et al.,* 2014) – along with robust analytics that enable and perpetuate the feedback loops that make MOOCs special, by allowing for measurement of participant engagement, and thus providing an opportunity to recognize where we need to recalibrate our message.

Further development is underway to shrink not participation in these experiences, but the experiences themselves, so as to create overlap between the multiple modalities of participation that learners wish to engage in. An example of this is Humanities (HUM) 1x 'The Book: Histories Across Time and Space,' a hyper–modular learning experience currently under development at HarvardX, which combines archaeology, history, art history, law, music, and several other schools and specialties into a user–directed, interdisciplinary learning experience made up not of a lengthy, 'in–or–out' type 'course,' but of numerous discrete learning experiences focused loosely on one overall topic – in this case, the History of The Book – which can be arranged into learning paths or engaged with on their own. Each of these discrete learning experiences can, in turn, be engaged with via user–directed paths, in terms of both order and depth.

3. Conclusion

The fact that we are willingly entering the open online world is both encouraging and worthwhile, and it is made vastly moreso by a willingness to engage in multidirectional dialogue, and to welcome multiple modalities of participation, trusting that those who participate in even a small portion of our learning experiences may take a portion of the knowledge they gained from us and use it positively, both to educate their peers and to combat the false information that abounds on the 21st century web. Further, though no single MOOC will suddenly challenge Wikipedia or any other multi–million–user resource on the web for Internet supremacy, the simple fact is that only by refusing to cede the field and leave the arena will we be able to truly educate individuals – and, through them, the world – about our profession, and about the critical knowledge, context, and information that is necessary to truly understand the past.

Acknowledgements

I would like to extend special thanks to the organizers of the session 'Archaeology at Large: Embracing Massive Audiences for Online Applications,' J. Andrew Dufton, Müge Durusu–Tanrıöver, and Susan E. Alcock (Brown University), as well as HarvardX, whose funding made attendance at CAA2014 possible.

References

ANTIN, J. and CHESHIRE, C. (2010). Readers Are Not Free–Riders: Reading as a Form of Participation on Wikipedia. Paper presented at the *2010 ACM Conference on Computer Supported Cooperative Work*, Savannah, GA, 6–10 Feb. Available at: http://dl.acm.org/citation.cfm?id=1718918.1718942.

BLANK, G. (2013). Who Creates Content? *Information, Communication & Society* 16(4), pp. 590–612.

BOHN, A., BAUCHTA, C., HORNIK, K. and MAIR, P. (2014). Making Friends and Communicating on Facebook: Implications for the Access to Social Capital. *Social Networks* 37, pp. 29–41.

BURGESS, J. and GREEN, J. (2009). *YouTube: Online Video and Participatory Culture*. Cambridge, MA.

CORREA, T. (2010). The Participation Divide Among 'Online Experts': Experience, Skills and Psychological Factors as Predictors of College Students' Web Content Creation. *Journal of Computer–Mediated Communication* 16(1), pp. 71–92.

DAUGHERTY, T., EASTIN, M. S. and BRIGHT, L. (2008). Exploring Consumer Motivations for Creating User–Generated Content. *Journal of Interactive Advertising* 8(2), pp. 16–25.

DESENNE, P. (2014). Annotations in the edX MOOC Platform. Paper presented at the *New Media Consortium 2014 Summer Conference*, Portland, OR, 17–19 June.

DESENNE, P. and REIS–DENNIS, L. (2014) Open edX Annotation Tools: Breaking the Unidirectionality of Online Course Content. Paper presented at the conference *Learning With MOOCs: A Practitioner's Workshop*, Cambridge, MA, 12–13 Aug.

DUFTON, J. A., DURUSU–TANRIÖVER, M. and ALCOCK, S. E. (2014). Archaeology at Large: Embracing Massive Audiences for Online Applications. Paper presented at the conference *Computer Applications and Quantitative Methods in Archaeology 2014*, Paris, France, 22–25 April.

EMANUEL, J. P. (2014). *Prospectus Omni ab Angulo*: 3D Imaging in Archaeological Reconstruction and Analysis. Paper presented at the Symposium *Thinking With Your Eyes: Visualizing the Arts, Humanities, and Sciences*, Cambridge, MA, 27–28 Feb.

EMANUEL, J. P. (forthcoming). MOOC as Multimedia Textbook: Rethinking the Open Online 'Course' in Light of Learner Intent. In J. P. Emanuel and C. Filos (eds.), *Digital Humanities in the Age of the MOOC*. Classics@ 12. Washington, DC.

FORTE, A. and BRUCKMAN, A. (2006). From Wikipedia to the Classroom: Exploring Online Publication and Learning. Paper presented at the *7th International Conference on Learning Sciences*, Bloomington, IN, 27 June–1 July 2006. Available at: http://www.cc.gatech.edu/~asb/papers/conference/forte–bruckman–icls06.pdf.

FOWLER, R., MEINKING, K. A., MORRELL, K., *et al.,* (2014). *Adapting Content from a Massive Open Online Course to a Liberal Arts Setting*. Transformations: National Institute for Technology in Liberal Education. http://www.academiccommons.org/collaboration-and-blended-learning/case-studies-on-digital-collaboration-and-blendedhybrid-learning/

HARWARD, J., HAMBURGER, J., STERN, R., SINGHAL, R. and EMANUEL, J. P. (2014). *Oculus: Using Open APIs to Share Harvard's Digitized Books and Manuscripts*. Paper presented at the *4th Annual Harvard University Information Technology Summit*, Cambridge, MA, 5 June. Available at http://www.academia.edu/7266427/Oculus_Using_Open_APIs_to_Share_Harvards_Digitized_Books_and_Manuscripts.

HO, A. D. and REICH, B. J. F. (2014). The Tricky Task of Figuring Out What Makes a MOOC Successful. *The Atlantic*, 23 Jan. Available at http://www.theatlantic.com/education/archive/2014/01/the–tricky–task–of–figuring–out–what–makes–a–mooc–successful/283274/.

HO, A. D., Reich, B. J. F., Nesterko, S. *et al.,* (2014). *HarvardX and MITx: The First Year of Open Online Courses*. HarvardX Working Papers 1, Cambridge.

HOLLANDS, F. M. and TIRTHALI, D. (2014) *MOOCs: Expectations and Reality*. Center for Benefit–Cost

Studies of Education, Teachers College, Columbia University. Available at: http://cbcse.org/wordpress/wp–content/uploads/2014/05/MOOCs_Expectations_and_Reality.pdf.

Huffington Post (2013) Celebrity Death Hoaxes Through the Years. *Huffington Post*, 14 Nov. Available at http://www.huffingtonpost.com/2013/11/14/celebrity–death–hoaxes_n_4268875.html.

International Telecommunications Union. (2010). *Definitions of World Telecommunication/ICT Indicators*. Available at: http://www.itu.int/ITU–D/ict/material/TelecomICT_Indicators_Definition_March2010_for_web.pdf.

International Telecommunications Union. (2013). *Percentage of Individuals Using the Internet*. Available at: http://www.itu.int/en/ITU–D/Statistics/Documents/statistics/2013/Individuals_Internet_2000–2012.xls.

LAUERMAN, J. (2014). Harvard, MIT Online Courses Dropped by 95% of Registrants. *Bloomberg*, 21 Jan. Available at: http://www.bloomberg.com/news/2014–01–21/harvard–online–courses–dropped–by–95–of–registered–study–says.html.

LEVY, A. (2014). Massive Online Education Gets Less Massive. *Bloomberg*, 14 Jan. Available at: http://www.bloomberg.com/news/2014–01–14/massive–online–education–gets–less–massive.html.

MANUELIAN, P. D. (2013). Giza 3D: Digital Archaeology and Scholarly Access to the Giza Pyramids. Paper presented at the conference *Digital Heritage 2013*, Marseilles, France, 28 Oct–1 Nov 2013. Available at: https://www.academia.edu/7572994/_Giza_3D_Digital_Archaeology_and_Scholarly_Access_to_the_Giza_Pyramids._The_Giza_Project_at_Harvard_University_.

MATAZAT, U. (2004). Academic Communication and Internet Discussion Groups: Transfer of Information or Creation of Social Contacts? *Social Networks* 26(3), 221–255.

MONDENERO MOYA, J. J., CEBRIÁN ROBLES, D. and DESENNE, P. (2014). Usabilidad y Satisfacción en Herramientas de Anotaciones Multimedia Para MOOC. *Comunicar 44*.

PARR, C. (2013). Not Staying the Course. *Times Higher Education*, 10 May. Available at: http://www.insidehighered.com/news/2013/05/10/new–study–low–mooc–completion–rates.

PERNA, L., RUBY, A., BORUCH, R. *et al.,* (2013) *The Lifecycle of a Million MOOC Users*. Paper presented at the *MOOC Research Initiative Conference*, Austin, TX, 5 Dec 2013. Available at: http://www.gse.upenn.edu/pdf/ahead/perna_ruby_boruch_moocs_dec2013.pdf.

REICH, B. J. F., EMANUEL, J. P., NESTERKO, S. *et al.,* (2014). *HeroesX: The Ancient Greek Hero: Spring 2013 Course Report*. HarvardX Working Papers 3, Cambridge. Available at: http://www.academia.edu/5788891/CB22x_The_Ancient_Greek_Hero_Spring_2013_Course_Report

SANDERS, D. (2013). The Gurob Ship–Cart Model in Virtual Reality. In S. Wachsmann, *The Gurob Ship–Cart Model and Its Mediterranean Context*, College Station, pp. 209–218.

VIÉGAS, F. B., WATTENBERG M., KRISS, J. and VAN HAM, F. (2007). Talk Before You Type: Coordination in Wikipedia. In *Proceedings of the 40th Annual Hawaii International Conference on System Sciences*, New York. Available at: http://hint.fm/papers/wikipedia_coordination_final.pdf

VIÉGAS, F. B. and WATTENBERG, M. (2014). Keynote Address. Presented at the Symposium *Thinking With Your Eyes: Visualizing the Arts, Humanities, and Sciences*, Cambridge, MA, 27–28 Feb.

WAGNER, C. and PRASARNPHANICH, P. (2007) .Innovating Collaborative Content Creation: The Role of Altruism and Wiki Technology. In *Proceedings of the 40th Annual Hawaii International Conference on System Sciences*, New York. Available at: http://ieeexplore.ieee.org/stamp/stamp.jsp?tp=&arnumber=4076406

Wikipedia. (2014). *Wikipedia: Statistics*. Available at: https://en.wikipedia.org/wiki/Wikipedia:Statistics.

YANG, H–L. and LAI, C–Y. (2010). Motivations of Wikipedia Content Contributors. *Computers in Human Behavior* 26(6), pp. 1377–1383.

ZAPATERO, L. A. and MORALES, M. M. de (2014). MOOCs as a Way to Disseminate Scientific Knowledge and Expertise Against the Death Penalty. Paper presented at the *8th International Technology, Education and Development Conference*, Valencia, Spain, 10–12 March. Abstract available at: http://library.iated.org/view/ARROYOZAPATERO2014MOO.

ZHOU, M., ROMANOWSKA, I., WU, Z., XU, P. and VERHAGEN, P., eds. (2012). *Revive the Past: Proceedings of the 39th Conference on Computer Applications and Quantitative Methods in Archaeology*. Amsterdam.

Moving Instruction Beyond the Museum's Walls: Priorities in Online Public Education at the Oriental Institute

Kathryn GROSSMAN,[1] Catherine Kenyon,[2] Megaera LORENZ,[3] Brittany HAYDEN[3]

[1] Massachusetts Institute of Technology, USA
[2] The Oriental Institute of the University of Chicago, USA
[3] University of Chicago, USA
kmgrossm@mit.edu , ckenyon@uchicago.edu, megaera@uchicago.edu, blhayden@uchicago.edu

Abstract
This paper presents the results of the first three years of a new program of online museum education at the University of Chicago's Oriental Institute, a research organization and museum devoted to the study of the ancient Near East. For this new initiative in online education, the Oriental Institute has offered three completely web-based courses for adult learners. Using participant demographics, institute membership rates, satisfaction surveys, and course completion rates as metrics, we argue that our experiences with fee-based, closed access, online courses can contribute to the development of more successful open access courses (e.g. Massive Open Online Courses or 'MOOCs'). We also explore the ways that online courses (both fee-based courses and MOOCs) can help to increase our public engagement and build our membership. Such a study has implications both for the Oriental Institute and for museum education programs more broadly.

Keywords: MOOCs, Online Learning, Public Education, Museums, Near East

1. Introduction

This paper presents the results of a new program of online museum education at the Oriental Institute of the University of Chicago. The Oriental Institute is a research center and museum within the University, dedicated to the study of the ancient Near East. Within the Oriental Institute, the Public Education and Outreach Department brings our scholarly research to a general audience; crucially, this audience is often highly educated and well informed about the ancient world. In this paper, we focus specifically on the public education department's adult programming, but it is worth noting that we also serve families, teachers, and school children. Our adult education programs have traditionally been offered in a classroom setting and are offered to both the general public and our membership.

The Oriental Institute has always tried to reach ever-larger audiences, bringing our research to more of the general public. That is part of the mission of being a museum and participating in an engaged archaeology more broadly. In 1996, we began offering online courses on the Near East as part of an effort to broaden the geographic scope of our audience. We recently looked into the benefits of offering a MOOC ('Massive Open Online Course') as a way to further broaden that scope. Although ultimately, the Institute will not be pursuing a MOOC in the near future, this paper is one result of our research into the issues surrounding online public education, both fee-based as well as massive and open. In it, we present some of the goals of our program of online education, the methods we have used to develop and offer these courses, the results of our education program, and some thoughts about how our experience in online education so far can inform MOOC programs at institutions like ours.

MOOCs have become a lightning rod in debates over the future of higher education in the United States (for a recent review of MOOC research, see e.g., Reich, 2015). The most undeniably positive aspect of these courses is the open access to education (both content and instruction) that they offer to virtually unlimited audiences. MOOCs also offer unprecedented opportunities to grow the brand of universities and research institutes alike. For instructors, MOOCs offer a chance to experiment with new technologies and reach global audiences (for a survey of MOOC instructors, see Kolowich and Newman, 2013).

Despite these important benefits, however, recent research has also demonstrated a number of critical issues regarding MOOCs. First, several studies have shown that MOOCs struggle to retain their enrollment over the life of the course (see, e.g. Breslow *et al.*, 2013) and that those who do enroll are often younger males with extensive previous educational experience (Emanuel, 2013; Ho *et al.*, 2014). This suggests that the educational benefit is, in reality, limited to a smaller, less 'massive' population, and particularly one that would also benefit from traditional courses. Second, the financial benefit of MOOCs for the institutions that host them has yet to be clearly demonstrated. Typically, these open courses have been offered by wealthy colleges and universities that can afford to experiment with the new educational medium (Hollands and Tirthali, 2014; Hoxby, 2014). Third, the production of a MOOC is highly labor-intensive for the professors who teach them. This can have dramatic consequences for the research productivity of those professors (see Hollands and Tirthali, 2014: 148-151).

The debate surrounding the benefits and drawbacks of MOOCs has been, so far, largely limited to colleges and universities; we argue, however, that such discussions are also relevant for research institutes like ours. We are interested in reaching a vast and international audience; we are interested in exploring how MOOCs can provide both free access to education and a financial benefit for our institute; and we are interested in exploring how MOOCs can directly benefit the researchers who invest time and effort in teaching them. In this paper, we show how our experiences with fee-based, closed access online courses can contribute to the development of successful open access courses that will continue to provide a firm financial foundation for our institute. Such a study has implications both for the Oriental Institute and for museum education programs more broadly.

2. Goals

In addition to the academics from around the world who come to the Institute to do research, the Oriental Institute has a vast international public audience. Guest logs show that a high proportion of our Museum visitors is international, as is a small percentage (2.3%) of our membership. One of our goals in offering our current program of fee-based online courses is to reach out to this audience and to allow members outside the Chicago area to take advantage of one of the perks of membership - discounts on education courses.

Our nonlocal audience is eager to engage with the Oriental Institute through social media platforms such as Facebook and Twitter, so we know that there is a desire for increased engagement with the institute through online avenues. However, we have been grappling with ways to reach these large numbers of offsite people. Over the last few decades, we offered a series of postal correspondence courses on ancient Near Eastern languages, but these courses tended to target a highly specialized and self-selecting audience. Web-based public education courses, therefore, offered an excellent opportunity for us to engage with a more diverse cross section of nonlocal students. When the current program of online teaching was conceptualized in 2009, we saw a number of benefits for both the department and the institute.

The benefits of this program work along two lines: audience building and revenue. In opening up the Oriental Institute to a nonlocal audience through online teaching, we hoped to attract new members, whose membership would support our research goals, and whose engagement with the institute would grow over time. These courses would also offer a deeper level of interaction with our current, nonlocal membership, making it possible for those who could not attend regular face-to-face classes to participate. Online courses have been offered to members and nonmembers for a fee ranging from 175 to 365 USD. The fee-based structure of these courses is intended to provide graduate students with paid teaching opportunities. With the increased attention to web-based teaching in higher education, our students also gain a distinct advantage in the academic job market by learning to teach online while still in graduate school. The fees from these web-based courses also go to support the programming budget for the Public Education Department.

3. Methods

Web-based adult education has a long history at the Oriental Institute. We offered our first online course, An Introduction to Ancient Egypt, in 1996, as higher education was first experimenting with the Internet. Peter Piccione, now at the University of Charleston, taught this first course. He offered the course twice, and both times it filled to capacity. This early version of an online course had a very similar structure to the organization of our more recent offerings; it was asynchronous, hosted entirely online, and was built around a similar structure of weekly readings and activities.

The recent online education initiative began in full force, however, in 2011. The University of Chicago had, in 2010, upgraded to a more user-friendly version of Blackboard, a common course management software system in the United States. The University supported the Oriental Institute's plans for offering online courses by providing access and tech support for building our classes on this platform.

We offered our first course, Dawn of History: Society and Culture in Ancient Mesopotamia, in the winter of 2012. That class has since been taught three more times and is now scheduled to be offered on an annual basis every autumn. We offered a second course, the Art and Architecture of Ancient Egypt, for the first time in the winter of 2013. That course was offered again in 2014 and is now being offered during the winter quarter of each year. Our third online adult course offering, Deciphering the Past: Beginning Egyptian Hieroglyphs, was offered for the first time in the spring of 2014.

4. Structure and Components of the Courses

The structure and components of our online courses have allowed us to create a highly successful program. While not all of these structures and practices would translate directly into a MOOC offering, they have given us a good sense for what works and what does not work in our public education offerings. Here we present some of the unique features of our courses and some of the reasons that we developed them.

4.1. Course Staff

The staff for each course includes Public Education Department employees, a graduate student instructor, a faculty advisory board, and a part-time Tech Support specialist. The instructors provide the content for the online courses; this content includes texts that will be used as course readings. Instructors also select and edit images and videos, create assignments, build maps in ArcGIS, and create additional materials such as glossaries, exams, and timelines. Because our courses are fee-based, we have had to be careful to build them using only our own institutional materials or copyright-free content.

In order to be sure that content was of the highest academic caliber and that our students would get the same rigorous educational experience that they had come to expect from both the Oriental Institute and the University of Chicago, all course material is vetted by a University of Chicago faculty committee. The committee then gives comments and suggestions to the instructor. The graduate student who builds the course content serves as the course instructor during its first iteration. Our goal, however, is to build a library of courses for the Public Education Department that graduating students can hand off to younger students, who will continue to teach the course using the same material. Our first experiment with that took place this fall, when a new student took on the teaching responsibilities for the ancient Mesopotamia course for the first time. The transition was smooth, with the incoming instructor using much of the same content, modified slightly for his particular approach to teaching.

During the development of the courses, the staff of the Public Education office designed the framework on which the content would hang; provided guidance and structure for the students developing it; and funded the initial work. The initial investment was enormous, in terms of both time and resources. The Mesopotamia and Egypt courses each took about seven months to build, and the hieroglyphs course took about eight months. Once the courses began, the Education staff took an even more active role, especially in administering the course and helping instructors tailor it to an adult audience. The education staff organized registration and fees, made sure that students were enrolled in the Blackboard system, supervised the online instruction, and managed the course evaluation process.

A Technical Support specialist was also assigned to each course. Before each course began, the tech specialist checked to see that all of the content was working properly as it was loaded online. Once each course was up and running, the tech specialist worked directly with the course participants to deal with technical issues. This freed the instructor from dealing with those issues, significantly reducing their workload.

4.2. Student Demographics in Our Web-Based Courses

The demographics of students in our web-based courses tend to mirror the demographics of our in-person adult education offerings. Although there is a great deal of variability, our students tend to be older (often retired) (Fig. 1) and well educated (often with some post-graduate education) (Fig. 2), with a strong prior interest in the Near East. For most this is their first online course, and many are not familiar with Internet terminology or functionality. In fact, our audience's lack of familiarity with technology and the internet has been one of our biggest challenges.

4.3. Instructional Design

Because the first two courses, devoted to Mesopotamia and Egypt, were structured similarly, we can describe those together. The Hieroglyphs course was structured somewhat differently, and we will present the design of that course separately.

4.3.1 Instructional Design for the Mesopotamia and Egypt Courses

In terms of instructional design, a separate course content module is generally released each week and remains active throughout the remainder of the course. Each module (Fig. 3) includes a weekly reading, illustrated with web links, as well as audio, video, and image files. Modules also contained a glossary of terms, a map of sites mentioned in the reading, a short quiz, a private journaling task, and a discussion assignment, all organized around a common theme. There are, therefore, no 'class meeting times' as there are for traditional face-to-face courses. This asynchronous setup permits students who live in different time zones or are juggling busy schedules to participate fully in the course. The downside to such asynchronism, however, is that much of the community built in a classroom setting is lost.

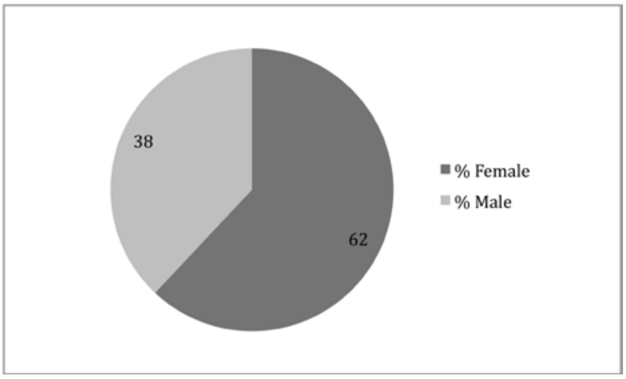


FIGURE 1: GENDER DEMOGRAPHICS IN WEB-BASED COURSES AT THE ORIENTAL INSTITUTE.

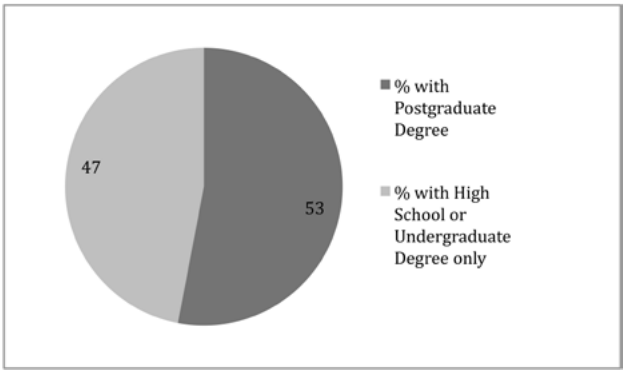


FIGURE 2: EDUCATION LEVELS IN WEB-BASED CLASSES AT THE ORIENTAL INSTITUTE.

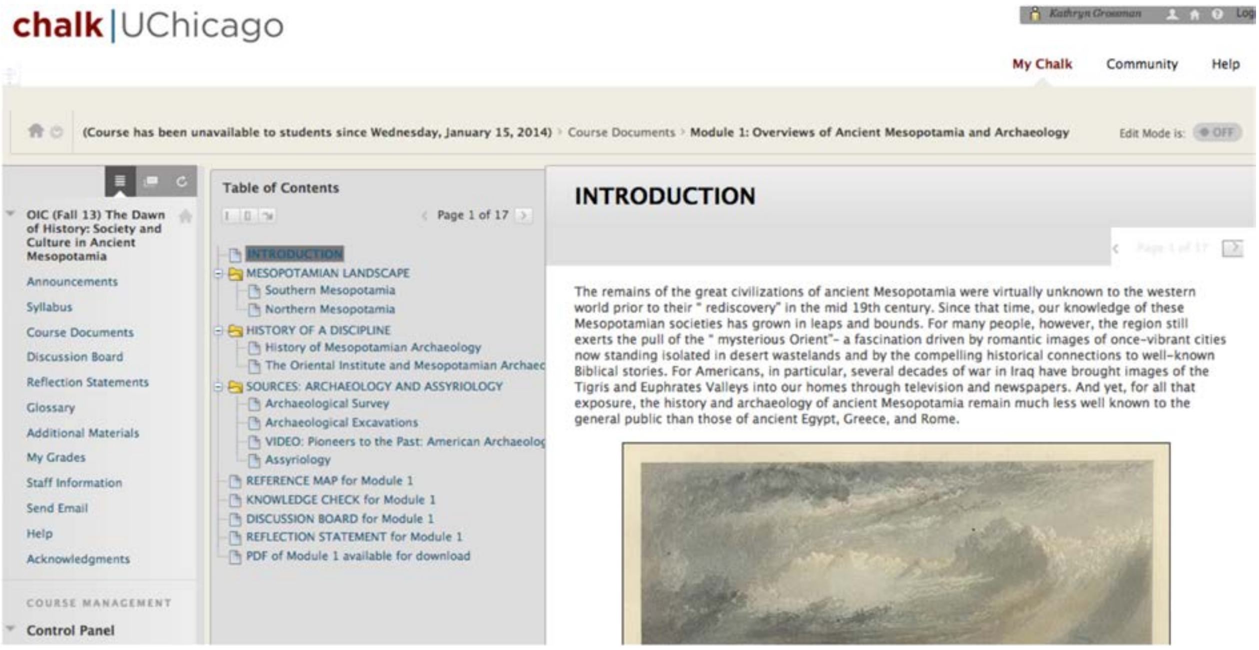


Figure 3: Screenshot of a Weekly Module for the Mesopotamia Course.
Module components shown in the Table of Contents (center left in image).

We have rebuilt that community by creating multiple avenues for student engagement with the course, with the instructor, and with one another. The most important is the discussion board (Fig. 4). The discussion board was designed to encourage students to interact with one another and with the instructor. Our discussion board activities are different each week, and for each class, but the theory is always the same. There is a weekly activity or set of questions that students are asked to complete. These might ask them, for example, to compare two academic articles or to analyze maps and plans. The students post their response and are then required to respond to two other students' posts. This gets the students talking to one another and working together on the material. The whole class can see the discussion, even if they do not participate.

While many students, especially older students, were initially distrustful of it, the discussion board has been the most successful part of many of our online courses. Responses to the assignment often begin appearing immediately after it is posted and in a typical week, the class as a whole might write upwards of 150 long and

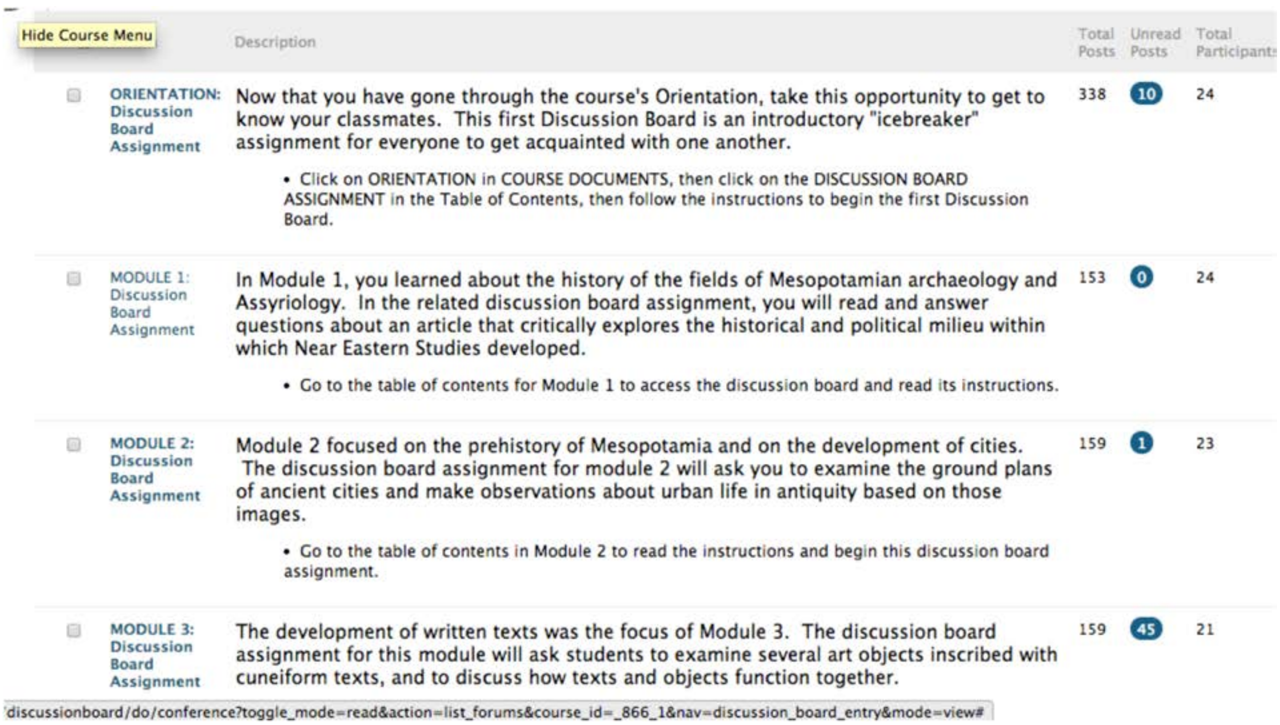


Figure 4: Screenshot of the Discussion Board for the Mesopotamia course.

thoughtful answers to the questions posed. In most weeks, the discussion expanded far beyond the initial prompt, and students used the space to answer one another's questions about the course material more broadly. The instructor's role was to keep the points relevant to the module, to make sure everyone's voice was heard, to correct mistakes, to answer the questions that inevitably arose, and, at the end of the week, to tie it all together with a long wrap-up post. Despite the lack of a physical classroom setting, the interaction with the students was, through the back-and-forth of the discussion board, far more intense than in traditional lectures courses.

In addition to the academic discussion board, each course also hosts a 'Tech Talk' board, which operates using the same system and is monitored by the Tech Support Specialist, rather than the instructor. Here, students can asynchronously discuss their technical troubles with the tech specialist and with one another. In so doing, they are very often able to solve technical issues themselves.

We have found, however, that some students just want to read the modules and avoid the discussion board. Some are turned off by the constant and intense nature of the discussions, and others simply do not have the time to devote to it. For these students, a journal feature has proven a good way to keep them active in the course. Journals, in our system, are a private venue for the student to communicate with the instructor. Students can ask questions, give feedback on the course itself, and reflect on the week's module.

We also post weekly reading quizzes, usually ten questions in multiple choice or true-false format. These are a way for students to check the knowledge that they have gained from the reading. Even though our courses are noncredit, students take these quizzes tremendously seriously. Grades are assigned for the courses, not for academic credit, but as a motivation.

4.3.2 Instructional Design for the Hieroglyphs Course

Our most recent new course offering, a course on Egyptian hieroglyphs, has been run somewhat differently. The course grew out of our old mail based correspondence courses, and so retains much of the structure from that system. There is, for example, a great deal of homework each week, asking students to work independently on their translations and learning the hieroglyphic signs. After the homework is submitted, the instructor records and posts a video to offer students both additional assistance and answers to common questions and mistakes made on each assignment. Each module also begins and ends with videos taken in the Oriental Institute's museum, where the instructor uses objects in the galleries to provide cultural context for the weekly grammar lesson. The discussion board in this course serves more as a location for dealing with technical issues and questions, and less as the central meeting place for the class.

The hieroglyphs course also makes use of a wider range of technological features, including digital flashcards (Fig. 5) that have an audio component and freeware hieroglyphic fonts used by professional Egyptologists. Students download the software for these fonts so that they can write in hieroglyphs online. Getting students to correctly access the fonts has been, in terms of technology, the trickiest part of the class.

5. Results

We carefully monitor the results of our courses through a series of Survey Monkey evaluations that we ask students to fill out before and after each course. These evaluations include questions about how comfortable the students are with the technology, how much they have learned, and how satisfied they have been with specific aspects of the courses, among others. We have found that our courses are well received, and we have been able to successfully

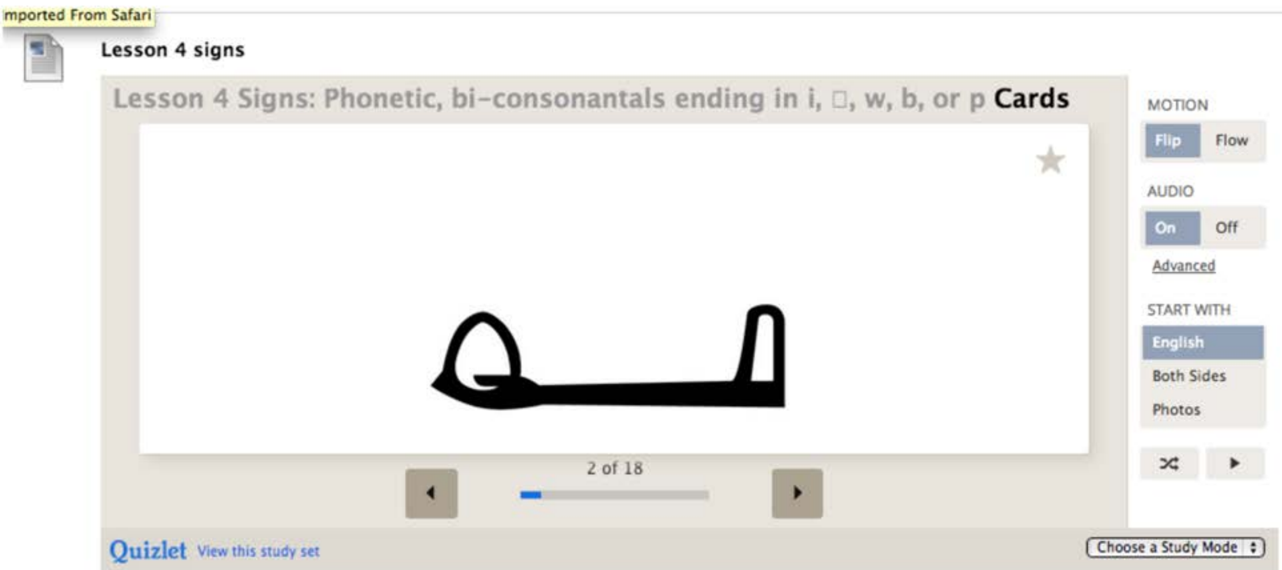


FIGURE 5: SCREENSHOT OF A DIGITAL FLASHCARD FOR THE HIEROGLYPHS COURSE.

modify future iterations of the courses based on the feedback from these surveys.

Perhaps the most astonishing result of our online courses is their completion rate. Over 100 students have enrolled in these classes so far, and we have achieved a roughly 60-70% completion rate. 'Completion,' in the case of these courses means that the student has remained engaged in class activities throughout the entirety of the course. We can monitor these rates because our courses have built-in tools that allow us to see when students log in to the course and which course components they access, even if an assignment is not submitted. Our completion rate is much better than the completion rates for most MOOCs, and far better than even our old postal correspondence courses. The Hieroglyphs course, however, achieved a lower completion rate (19%). Based on the survey results, this seems to be due to the more complicated content; some students who enrolled, for example, did not realize at the onset that learning hieroglyphs was far more complicated than learning a romance language. We are revaluating the course now, but it seems that we may need to rethink its target audience. Rather than being aimed at an interested, but non-academic audience, it might be a better choice for students in bachelors degree and masters degree programs who want to prepare for a doctoral program in Egyptology, but do not have access to a hieroglyphs course.

Our most hoped for result, that these online courses would lead to membership growth, has also come to pass. We found that many nonmembers joined the Institute when they enrolled in these courses in order to take advantage of the reduced rate for members. We hope (and believe) that many of these new members will continue their relationship with the Institute into the future. It is gratifying to know that outreach through online courses can, in fact, be effective at growing our membership. The growth, while measurable, is small; our advertising budget, for example, is limited, and our outreach has typically been aimed at a public audience that is already interested in the ancient world.

The students have been overwhelmingly satisfied with the courses so far, with over 90% saying that they learned a great deal from the course, 98% reporting that they would recommend courses at the Oriental Institute to others, and 100% reporting that they were interested in taking more courses with us. Initially we had worried that an online course platform would turn our older students way, but we found that many students were so pleased with the format that they did in fact continue to enroll in other online course after successfully completing the first one.

In financial terms, we have seen our first course, Dawn of History, turn an overall profit after its third offering. We are confident that the Art and Architecture course will turn that same corner this winter when it is offered a third time, and we assume the same timeframe of profitability will be true for the Hieroglyphs course.

6. Building on our Successes: Toward the Integration of MOOCs in Public Education Programs

Moving forward, we plan to teach each of these fee-based online courses once a year – we do not have enough enrollment to support continuous offerings of each course year round. Looking ahead, however, our next projects include developing an online Cuneiform course and, perhaps in the future, expanding to offer a MOOC. We had initially hoped to offer a MOOC during 2015, but the initial time investment was too great. Nevertheless, it is still something that we hope to pursue in the future. The goal in offering a MOOC is to reach a vastly larger number of people than we could ever reach with a paid course. As one example, such an approach has worked for the American Museum of Natural History (AMNH); in 2013, the Museum offered a MOOC for educators through the course-hosting platform Coursera. Surveys of the enrollees showed that the majority (62%) had no prior knowledge of the Museum (Hollands and Tirthali 2014: 71). This level of engagement with new audiences is one of the main benefits we see in offering a MOOC through our own museum. The promotion of the AMNH course through the well-known Coursera platform seems to have been a key factor in reaching such a large audience. As advertising costs have been been a major hurdle to increased enrollment in our fee-based courses, the option to host a MOOC through one of these major platforms is attractive. It would allow us to reach a far greater audience than much of our previous outreach.

In general, we have found that our fee-based online courses reach an audience very similar to that of our face-to-face instruction. Archaeology, however, is a discipline that is of great interest to the broader public; we want to parlay that interest into more direct engagement with the Institute, using it to build our membership and create buzz about our research projects. We hope to spark interest in the ancient Near East in people who would not otherwise have known about the Oriental Institute. In addition to their sheer size, MOOCs tend to enroll a slightly different demographic than our courses have reached. For example, MOOCs typically reach a much younger audience than our course offerings have (Fig. 6). The median age of MOOC enrollees is 23-30, and only 6% are older than 50 (Ho *et al.*, 2014). We are particularly excited about the possibility of engaging with such a different demographic group. The one demographic feature that our current courses and MOOCs have in common is prior education. MOOCs, like our courses, tend to enroll students with substantial prior educational experience (Emanuel, 2013).

It is important to point out that we (and indeed, higher education in general) do not know yet what the benefits will be for institutions that offer MOOCs. Certainly their widely publicized, low completion rates are troubling. However, studies have shown that the people who fully participate in (and complete) MOOCs are generally students like those who already take part in our other museum education courses: well educated and already

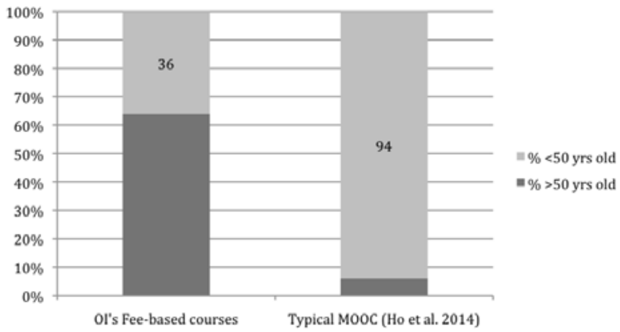


FIGURE 6: AGE STRUCTURE COMPARISONS, ORIENTAL INSTITUTE COURSES AND TYPICAL MOOCS.

engaged academically. This suggests to us that typical MOOC completers might be interested in continuing to engage with the Oriental Institute. As we move forward, however, we need to carefully monitor enrollment in our regular online classes to judge the degree to which offering a MOOC will impact interest in our fee-based courses. This is especially important because our online program does not offer degrees, certificates, or other incentives for people to continue their enrollment. We do not see MOOCs taking the place of our fee-based courses; instead, we anticipate that there will always be an audience for both.

Despite these lingering questions, there can be no doubt that online education has already increased our exposure. The courses have also provided an avenue for nonlocal Institute members to retain their connection with us and they have provided many graduate students with invaluable exposure to web-based teaching. Our online education program stands poised to continue helping the Institute further its goal of introducing the public to the history and archaeology of the ancient Near East. As we see it now, it can be nothing but a positive step forward for our institution and others like us.

Acknowledgements

We would like to gratefully acknowledge the help, advice, and support of the following individuals: Gil Stein, Steve Camp, Nadine Moeller, Peter Piccione, Emily Teeter, Jack Green, Wendy Ennes, Carol Krucoff, John Sanders, Paul Ruffin, Jamie Cunningham, Paul Bergen, and Emily Bembeneck. Thanks also to the anonymous reviewers who provided helpful feedback. Any mistakes are our own

References

BRESLOW L. , PRITCHARD D., DEBOER J., STUMP G., HO A., SEATON D. (2013). Studying Learning int he Worldwide Classroom: Research into edX's First MOOC, Reasearch and Practice in Assesment, 8. pp. 13-25.

EMANUEL, E. (2013). Online education: MOOCs taken by educated few, *Nature* 503, p. 324.

HO, A. D., REICH, J., NESTERKO, S., SEATON, D. T., MULLANEY, T., WALDO, J., & CHUANG, I. (2014). HarvardX and MITx: The first year of open online courses. *HarvardX and MITx Working Paper No. 1*.

HOLLANDS, F. and D. TIRTHALI. 2014. *MOOCs: Expectations and Reality*. New York: Columbia University Teachers' College, Center for Benefit-Cost Studies of Education.

HOXBY, C. (2014). The Economics of Online Postsecondary Education: MOOCs, Nonselective Education, and Highly Selective Education, *NBER Working Paper No. 19816*.

KOLOWICH, S. and J. NEWMAN. (2013). The Minds Behind the MOOCs: Survey Results, *Chronicle of Higher Education* 21 March 2013.

REICH, J. (2015). Rebooting MOOC Research, *Science* 347(6217). pp.34-35.

Crowd- and Community-Fuelled Archaeology. Early Results from the MicroPasts Project

Chiara BONACCHI,[1] Andrew BEVAN,[1] Daniel PETT,[2] Adi KEINAN-SCHOONBAERT[1]

[1] Institute of Archaeology, University College London, UK
[2] British Museum, UK
c.bonacchi@ucl.ac.uk, a.bevan@ucl.ac.uk, DPETT@britishmuseum.org, adi.keinan.09@ucl.ac.uk

Abstract

The MicroPasts project is a novel experiment in the use of crowd-based methodologies to enable participatory archaeological research. Building on a long tradition of offline community archaeology in the UK, this initiative aims to integrate crowd-sourcing, crowd-funding and forum-based discussion to encourage groups of academics and volunteers to collaborate on the web. This paper will introduce MicroPasts, its aims, methods and initial results, with a particular emphasis on project evaluation. The evaluative work conducted over the first few months of the project already demonstrates the potential for crowd-sourced archaeological 3D modelling, especially amongst younger audiences, next to more traditional kinds of crowd-sourcing such as transcription. It has also allowed a comparative assessment of different methods for sustaining contributor participation through time and a discussion of their implications for the sustainability of the MicroPasts project and (potentially) other archaeological crowd-sourcing endeavours.

Keywords: Crowd-sourcing, crowd-funding, Public Archaeology, online communities

1. Introduction

MicroPasts is a web-enabled crowd-sourcing and crowd-funding project whose overall goal is to promote the collection and re-use of high quality research data via institutional and community collaborations, both online and offline. The MicroPasts model for participatory research should be seen in the context of recent developments in the field of Public Archaeology and in the practice of archaeological and heritage crowd-sourcing. Below, the aims and methods chosen to assess this model are discussed, together with the first results produced in the three months since the public launch of the MicroPasts crowd-sourcing website and forum. The final section of this paper then reflects more widely on some continuing challenges of archaeological crowd-sourcing and how these can be addressed.

1.1. Archaeology, Digital Technologies and 'the Public'

In the last few years, researchers in the field of Public Archaeology have urged a thorough review and consolidation of the theoretical and methodological approaches that can be applied to examine the multiple forms of the interaction between archaeology and society (see the discussion in Bonacchi, 2014: 379, and a forthcoming issue of World Archaeology dedicated to Public Archaeology). A large part of these reflections have considered the impact of the rapidly changing media and communication landscape on the public's uses of archaeological resources for research and other purposes, as well as on their participation in heritage policing. Increasing attention has been dedicated, in particular, to understand the potential of digital technologies for facilitating new, more collaborative and creative forms of public engagement with the human past (e.g. Bonacchi, 2012; Richardson, 2013). In parallel, in the area of archaeological science, researchers specialising in computational approaches to archaeology have addressed issues concerning open access and open data (e.g. Kansa *et al.,* 2012; Lake, 2012), sometimes also considering the wider implications of adopting these relatively novel practices for society at large, or specific online and offline communities (e.g. Beale, 2012; Bevan, 2012; Hole, 2012). While investigating different aspects of the relationship between archaeology, new digital media (McQuail, 2005: 38; Lister *et al.*,, 2009: 13) and citizens, these threads of research have ultimately had to confront (more or less explicitly) the question of the public value(s) of archaeology. By doing so, they have supported (and partly reflected) efforts made at national policy level to encourage the sustainability of heritage organisations and a more proactive social role for research institutions in a time of crisis. Expressions of these concerns by policy makers in Great Britain are, for example, the calls opened in 2013 by the UK Arts and Humanities Research Council (AHRC) for pilot projects and reviews intending to explore fresh ways of understanding the value of arts and culture, or to collaborate with galleries, libraries, archives and museums for developing and critically evaluating public engagement agendas.

1.2. Crowd-sourcing in Archaeology

In this broader context, crowd-sourcing, as a way of collecting information, services or funds in small amounts from large groups of people over the internet, has received growing attention from archaeologists as well as other cultural heritage professionals (Dunn and Hedges, 2012). This method emerged less than a decade ago in the commercial sector, where companies had been looking for ways of 'out-sourcing' labour to potentially interested 'crowds' of workers around the globe (Howe, 2006). Today, it is also being explored for not-for-profit purposes, for example as a means of conducting both science and humanities research, curating museum collections and managing heritage resources in collaboration with the public (see, for example, Oomen and Aroyo, 2011; Dunn and Hedges, 2012; Carletti *et al.,* 2013; Ridge, 2013). In archaeology more specifically, crowd-sourcing endeavours have spanned a wide array of activities, including the transcription of textual records (e.g. Ancient Lives project), the inspection of imagery for archaeological features (e.g. Field Expedition Mongolia), the interrogation of built architecture (e.g. Bodiam Castle Project) and the public recording of metal artefacts (e.g. Portable Antiquities Scheme), to name just a few.

Although increasingly popular, however, crowd-sourcing initiatives in archaeology have been primarily of a 'contributory' nature (Simon, 2010; Oomen and Aroyo, 2011: 139), inviting volunteers to offer their time, skills and experience to help with projects that had been designed entirely by 'professionals'. Until now, the latter have largely sought help for transcription and geo-referencing processes requiring the completion of mechanical and editorial tasks (Dunn and Hedges, 2012: 21, 36-37). In addition, a substantial number of projects have aimed at micro-financing (crowd-funding) archaeology. Being considered as a kind of crowd-sourcing by some commentators (Oomen and Aroyo, 2011: 140) and rejected as such by others (Dunn and Hedges, 2012: 6), crowd-funding is increasingly pursued to support a wide range of activities in the archaeological domain, spanning excavations, conservation, student dissertations or the production of documentary films. While the majority of these endeavours rely on 'generalist' commercial crowd-funding platforms like Kickstarter or Indiegogo, a minority use heritage-themed crowd-funding websites (e.g. CommonSites, DigVentures).

In summary, it seems that, until now, crowd-based methods have not been leveraged to encourage public involvement in the creation and use of archaeological information from the outset - in the sense of the 'co-creative' projects described by Simon (2010). Furthermore, very little evaluative work has been conducted so far in order to investigate the motivations leading different people to engage with archaeological crowd-sourcing and crowd-funding, the ways in which volunteers participate, the value placed on these exercises by contributors and partner organisations, and their longer-term sustainability, scalability and applicability beyond Anglophone countries. Amongst other goals, MicroPasts aims to address the two key concerns outlined above, by developing and evaluating a novel model for participatory archaeological research that utilises web technologies and crowd-sourcing.

2. The MicroPasts Model

2.1. Overview of the Project

MicroPasts is a collaboration between the Institute of Archaeology, University College London (UCL), and the British Museum, supported by the AHRC 'Digital Transformations in Community Research Co-Production in the Arts and Humanities' fund. This scheme was designed to facilitate research which proposes to draw on original digital methods for creating resources of enduring value to both academics and relevant communities. As part of the MicroPasts project, a website (micropasts.org) was created where communities that are already established offline (e.g. archaeological and historical societies, groups of metal detectorists, etc.) as well as more ubiquitous online 'crowds' can participate in one or more of three activities in collaboration with 'traditional' academics. We hope that, while doing this, participants also progressively form new online communities of (general or particular) archaeological interest. Here we characterise a 'crowd' as largely anonymous and fleetingly involved, as opposed to a 'community' that is repeatedly involved, with clearer group consciousness and interconnectedness (Haythornthwaite, 2009). The three activities that can be undertaken on the MicroPasts website consist of co-producing archaeological and historical open data via crowd-sourcing; designing new research agendas involving both volunteers and traditional academics; and crowd-funding some of these new collaborations that have been dreamt up collectively.

The first six months of the project were dedicated to developing a series of core open source web components through shared coding practices and version control using GitHub (https://github.com/micropasts). The forum (community.micropasts.org) has been built using the Discourse discussion software,[1] which draws on Ruby on Rails, Postgres and Redis, while the crowd-sourcing site (crowdsourced.micropasts.org) relies on the PyBossa framework (code base: Python, Postgres, Bootstrap theme).[2] Finally, the crowd-funding platform is constructed by modifying and extending the Neighbor.ly and Catarse[3] open source crowd-funding frameworks, which are built on Ruby on Rails, Postgres and Redis, and will (at least initially) use PayPal to process payments. Our aim is to encourage members of the public to first participate in crowd-sourcing and subsequently to get involved in the design of new projects and in crowd-funding.

We launched both the crowd-sourcing and forum components of the MicroPasts website on 16 April 2014, before the crowd-funding platform was completed. Our outreach strategy for the launch was articulated into two separate strands. The first consisted of tailored communications via email, talks and social media to reach our target audiences: archaeological and historical societies based in the UK, groups of metal detectorists and

[1] See Discourse at https://github.com/discourse/discourse.
[2] See Pybossa at https://github.com/PyBossa.
[3] See https://github.com/neighborly/neighborly and https://github.com/catarse/catarse.

other 'communities of interest' connected to the partnering institutions (UCL and the British Museum). In order to attract an as yet unknown online 'crowd' potentially interested in archaeology, history, or (even) the digital methods that are used in the project, we drew instead on the joint action of the UCL and British Museum press offices, who contacted a number of local and national media outlets including newspapers, magazines, radio and TV broadcasters. As a result, news about MicroPasts circulated via the social media accounts of (amongst others) UCL, the UCL Institute of Archaeology, the British Museum and the Portable Antiquities Scheme, through mailing lists (e.g. the Museum and Computer Group, Heritage, the UCL Centre for Audio-Visual Study and Practice in Archaeology), numerous popular blogs (such as io9.com) and relevant websites (e.g. ICCROM, the Megalithic Portal, etc.). However, so far, the project has featured in very few magazines (e.g. Heritage Daily), and has received virtually no coverage from general media outlets with national or international reach.

2.2. Three Components: Crowd-sourcing, Crowd-funding and Co-design

The MicroPasts crowd-sourcing site was launched with just two initial types of applications, both focusing on British Prehistory and British Museum collections. The first type involves archival transcription and geo-referencing tasks, while the second focuses on photo-masking to support 3D object modelling. By participating in archival transcription, volunteers can help digitise over 30,000 object cards that document Bronze Age metal artefacts found mostly in Britain from the nineteenth century onwards. These cards are part of the National Bronze Implements Index (NBII), an archive that was first developed around 1914 and has been housed at the British Museum since the 1960s. The NBII forms the first extensive catalogue of Bronze Age objects in Britain and Europe, and represents an untapped source of information about later prehistory. The cards are organised in numbered drawers by object type (e.g. spearhead, axe, sword, etc.) and find-spot (generally, county, town, and/or museum/private collection). Via crowd-sourcing, MicroPasts users transcribe object cards online and geo-reference the sites of discovery on a map powered by OpenLayers3 (Figure 1). By doing this, volunteers facilitate further research into the history of NBII (for example, by gaining information on the full geographic and chronological scope of its collections), as well as advancing existing knowledge of curatorial practices in Britain over the last century.[4] More importantly perhaps, this newly digitised resource will be invaluable for the study and comparison of Bronze Age objects, enriching the extensive Portable Antiquities Scheme's (PAS) spatial database, which records metal artefacts discovered in England and Wales from 2003 to the present day. By combining these two databases, the MicroPasts project will complement the public-facing nature of the PAS as well as form potentially one of the largest digital archives on prehistoric metal objects anywhere in the world.

The second type of crowd-sourcing application involves 'photo-masking'. Volunteers are invited to click around the outline of an artefact shown to them in a photograph. For each artefact, a set of at least 50 photographs that cover the object's entire external surface (Figure 2) is captured. Via an increasingly popular method known as Structure-from-Motion, common features can be identified in overlapping

[4] See http://finds.org.uk/info/advice/aboutus.

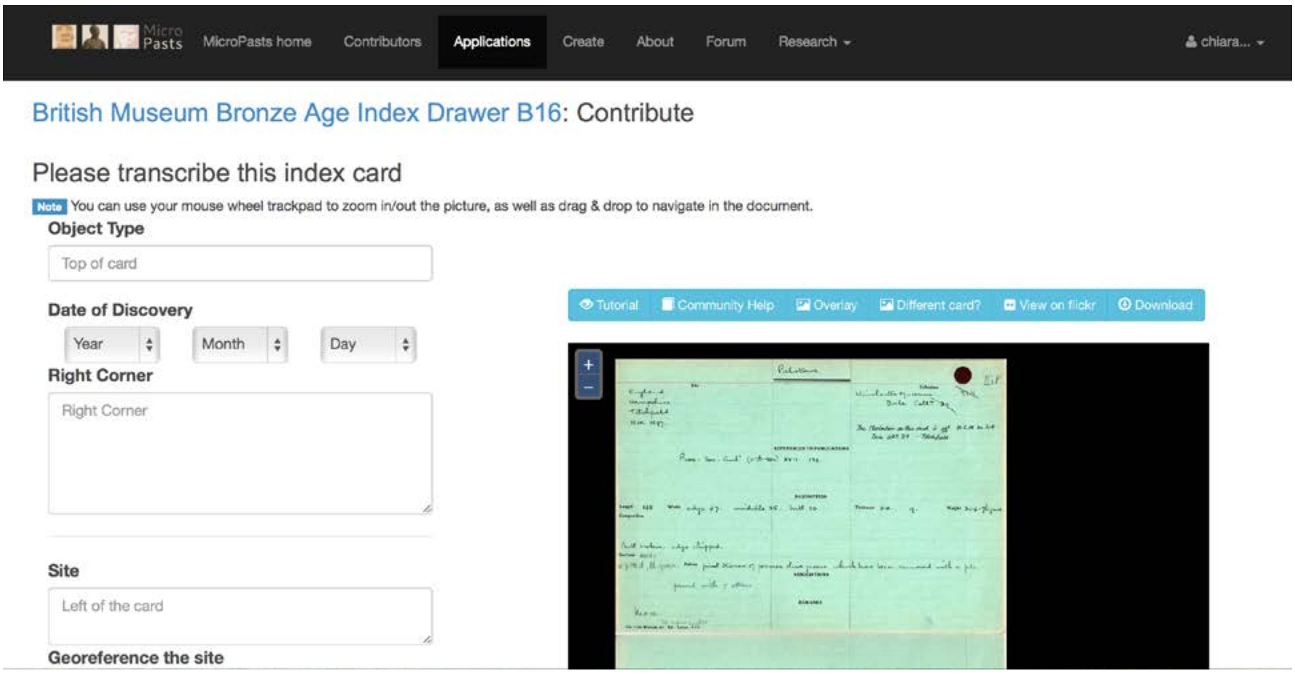


Figure 1: Interface of one of the MicroPasts transcription applications, showing the upper part of the form where data are entered by contributors (on the left), and the card that is being transcribed (on the right).

photographs of the same object and these can then be used to build a high quality 3D model of the object (e.g. in archaeology, Ducke *et al.,*, 2011; Remondino *et al.,*, 2012; Verhoeven *et al.,*, 2012; Green *et al.,*, 2014). By drawing the outline of the object in each photograph, MicroPasts contributors allow us to 'mask' out the background and focus our model-building on the object only (Figure 3). This simple but important task considerably improves the quality of 3D models we can generate (Figure 4), and getting two contributors to mask each photograph makes it possible to check the quality of the results.[5] Currently, on the MicroPasts crowd-sourcing website, users can mask images for the creation of 3D models of Bronze Age palstaves and other metal artefacts that are recorded in the NBII archive. Once ready, large samples of models for similar objects will allow statistical shape analysis that is useful for rethinking artefact typologies (e.g. Bevan *et al.,* 2014). On 2 July 2014 we also launched a new application, developed in collaboration with the Petrie Museum, to enable the photo-masking of an Egyptian funerary figurine, a shabti. Small-scale 'pop-up' applications of this kind are easy to implement, and while their role in a specific research agenda is often less clear-cut, they are usually proposed with aims of quality checking in mind, and allow diversifying the sub-types of crowd-sourcing applications and the range of participating institutions on the site (see the section 3.2.).

To offer opportunities for volunteers to learn more and develop further practical skills if they wish, we have developed a 'Learning' page on micropasts.org with a number of resources. These include step-by-step working notes explaining how 3D models can be created offline with different kinds of software, other aspects of method (for example concerning crowd-sourcing itself), and background information on the collections and on the geographic and chronological contexts with which MicroPasts is concerned. These resources can be discussed and enriched by volunteers via the forum, which is a space where contributors can debate how they would like the MicroPasts platforms and project to develop (Figure 5). At a more advanced stage in the production of crowd-sourced data, community.micropasts.org will also be useful to host open conversations on how the derived data produced by volunteers could support new research agendas. These newly co-designed archaeological projects will be able to seek funding via the MicroPasts crowd-funding website.

Through the MicroPasts crowd-funding platform it is possible to raise funding for archaeological research projects not focusing on excavation and which have been developed jointly by mixed groups of academics and community partners, either on the MicroPasts forum or elsewhere. Teams of this kind are able to submit proposals indicating a minimum and a maximum funding goal (initially up to £5,000), aims and context of the collaboration as well as outcomes and digital outputs. The research findings

Figure 2: Photo-taking of Bronze Age implements at the British Museum. The images were then uploaded to Flickr and made available for photo-masking.

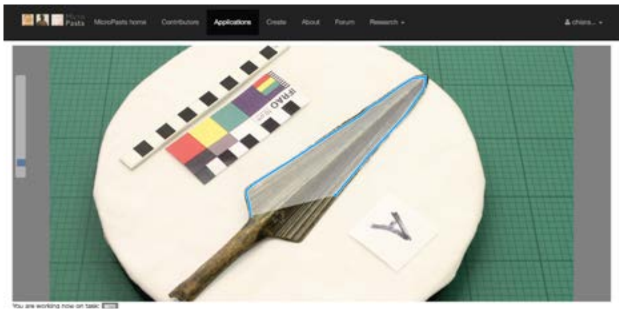


Figure 3: Outline drawing on one of the photo-masking applications.

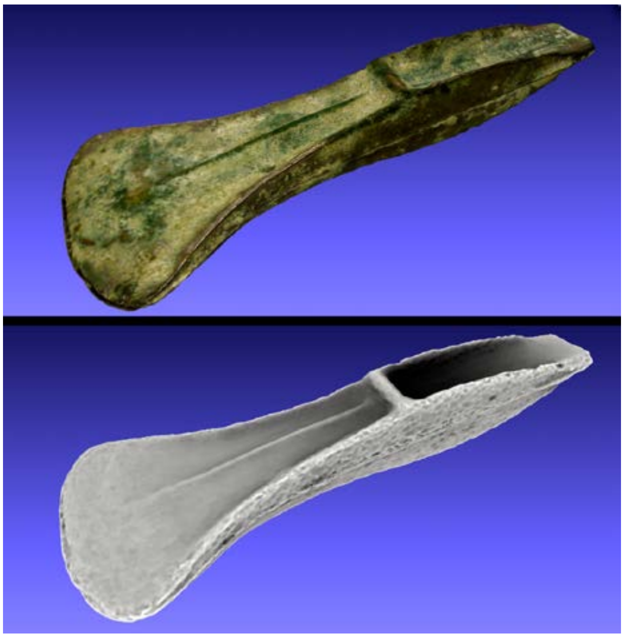


Figure 4: A 3D model of a Bronze Age palstave, shown both with a photographic texture and with an 'ambient occlusion' surface (for an online version visible in most browsers, see micropasts.org/3D/).

[5] The algorithm developed for quality validation purposes is stored on GitHub and is available at https://github.com/findsorguk/MicroPasts-Scripts/blob/master/photoMasking.py

from these projects will be made available online under open license. We are looking then to explore what rewards, beyond access to the final data, contributors really might wish to have from an archaeological project in exchange for financial support (e.g. direct participation, better access on special open days, participation in discussion with the project team, or something more traditional such as a T-shirt, a badge, or a book). Facilitating the crowd-funding of collaborations between academics and communities is also, ultimately, a way of opening up institutional doors to the public in a more structured and systematic manner, and of investigating the extent to which open access and open data can expand the already considerable impact of offline and local community archaeology, at least in certain countries such as the UK (Beale, 2012).

3. Evaluation

3.1. Aims and Methodology

The lead author of this paper has a particular research focus on evaluating: (a) the processes via which online communities of interest in the human past develop through the MicroPasts websites; (b) how these sites are used by different contributors to participate in archaeological and historical research and the value of this participation; (c) the likely longer-term sustainability of the MicroPasts model. To address these aspects and contribute to a wider discussion on the role of crowd-sourcing as a method for initiating and sustaining the co-production of science and humanities research, we draw on both quantitative and qualitative approaches, combining more traditional ‘digitised’ methods (e.g. online surveys) with ‘natively’ digital methods (i.e. methods that have not just been ‘transferred’ from the offline to the online world, but that exist exclusively online; see Rogers, 2013), and offline ones (e.g. interviews). Throughout the project, information (anonymised where appropriate) will be collected on user motivations, behaviour, and socio-demographic characteristics. Not all of this information, however, has started to be collected yet or can be analysed at this early stage (as of 25 July 2014).

The discussion below moves on to consider how the MicroPasts volunteer society has formed so far, in what ways people have participated and, in the light of this, how sustainable the project might be (in this admittedly initial phase of its lifecycle). It draws upon 313 responses to an online survey that was coded in to pop up after the completion of the first crowd-sourcing task. The survey enquires about only three things via closed questions: how contributors have heard of MicroPasts,[6] whether they ‘work with archaeology or history as part of their main job’,[7] and their age.[8]

Other data that will be discussed are acquired via Google Analytics or extracted directly from the MicroPasts platforms. Finally, we will comment on the outcomes of an ad hoc survey emailed to three archaeological and historical societies who were introduced to the MicroPasts project through a talk (the Wiltshire Archaeological and Natural History Society and the Chess Valley Archaeological and Historical Society), or thanks to an email sent by British Museum curator Wilkin (to the Later Prehistoric Finds Group).

[6] How did you find out about MicroPasts? Tick ALL the options that apply. Options: 1. Via the Portable Antiquities Scheme; 2. Via British Museum people/websites/social media; 3. Via University College London people/websites/social media; Via people/websites of another university (NOT University College London); Via an archaeological/historical society; From my school; I was told by someone who does NOT belong to any of the categories listed above; From an online newspaper/magazine; Casually, browsing the web.
[7] Do you work with archaeology or history as part of your MAIN job? Please choose only ONE option from the list below. Options: No; Yes.
[8] Your age: Please choose only ONE option from the list below. Options: 8-11; 12-17; 18-24; 25-34; 35-44; 45-54; 55-64; 65-74; 75+.

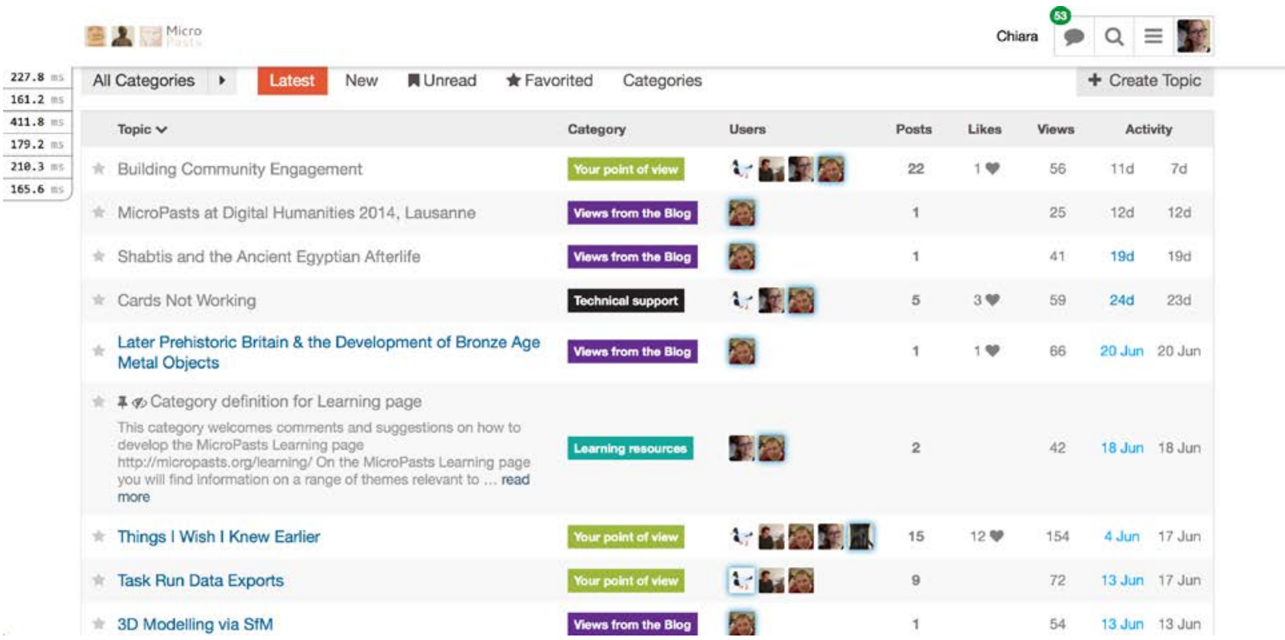


Figure 5: The MicroPasts community forum.

3.2. Initial Findings

A number of published articles and reviews have already commented on the kinds of citizen science groups that emerge from heritage crowd-sourcing, discussing their size, the level of interaction between participants and the nature of individual contributions (e.g. Dunn and Hedges, 2012; Owens, 2013; Proctor, 2013; Ridge, 2013, to name a few). These studies emphasise that, more often than not, 'crowd'-sourcing endeavours actually involve 'small' rather than 'very large' groups, and that most of the volunteers are already connected in some way to the institutions proposing the activities (e.g. Owens, 2013; Proctor, 2013). Despite being limited in quantity and mainly anecdotal, existing evidence suggests that archaeological crowd-sourcing has also tended to involve low numbers of (albeit often enthusiastic) contributors. This trend is apparent in the Ancient Lives and Bodiam Castle projects, as well as some web-based research collaborations engaging a handful of committed participants (e.g. the Durham Deanery project [Masinton 2014] - and the 'crowd-sourced' interpretation of LiDAR data presented in Duckers 2013).

As previously mentioned (section 2.1.), MicroPasts aims to further benefit, and benefit from, organised groups who are already pursuing archaeological or historical research offline, as well as a more dispersed and international 'crowd'. As of July 2014, after three months of operation, two main phases can be identified in the process of building a group of volunteers on crowdsourced.micropasts.org. The first three weeks saw a large number of people exploring the MicroPasts websites and registering as members of the crowd-sourcing site, in step with a flurry of news and publicity (Figure 6). By the end of the fourth week, however, as the initial publicity drive dropped off, sessions across all of the MicroPasts sites dropped as well, from 3,733 (sessions in the third week) to 1,381, and new membership rates tailed off to an average of just 7 new contributors per week. By this time, a previously large group of users based primarily in the UK and US, but also worldwide (as evidenced by Google Analytics) had transformed into a much smaller group of more active, regular volunteers. However, it is fair to say, that this latter group has mainly focused on completing the proposed tasks rather than using social media or the forum for wider discussion: for example, in the few cases where the forum was utilised by contributors, this was typically to raise technical issues. Hence, it is certainly not possible to claim that a distinctively MicroPasts community (that talks amongst itself) exists yet, but that impression is not unlike the one noted by other heritage crowd-sourcing projects such as Transcribe Bentham (Causer *et al.*,, 2012: 14).

More positively, across both phases described above, MicroPasts has already managed to involve people who do not work with archaeology or history as part of their main job, with 71% of survey respondents falling into this category (see section 3.1.). However, only a very small proportion of these contributors belong to those already established communities of interest that we have been specifically targeting. No more than eight respondents out of 313 claimed to have heard of MicroPasts from an archaeological or historical society, and only 13 via the UK Portable Antiquities Scheme. In addition, only six members of the three archaeological societies who were invited to try MicroPasts and questioned about their experience via an ad hoc online survey actually submitted a completed questionnaire. Perhaps the limited take-up of MicroPasts amongst archaeological groups could be explained either in terms of a mismatch between the generally younger age of savvy users of digital technologies and that of societies' members, or due to these society members already getting their required access to archaeological activity offline. Regardless, the issue clearly requires further and closer scrutiny over a longer time period.

For most volunteers, both anonymous and registered, participation in MicroPasts crowd-sourcing seems to start by trying out the photo-masking task (Figure 7). However, despite its initial appeal and the 'Learning' page and online 3D model viewer created to provide context and purpose to this kind of application, masking is soon abandoned by the majority of users and, overall, a greater number of transcription tasks than masking tasks are submitted (Figure 8). As shown in Figure 9, a handful of super-transcribers account for most of the transcription work, as frequently happens in crowd-sourcing projects (Holley, 2010; Causer and Wallace, 2012). Most probably, photo-masking holds people's attention for less time, because its purpose is less obvious, the end result (a 3D model) is not immediately produced, and the task is exclusively 'mechanical' (Dunn and Hedges, 2012: 36-37). Transcription also has a longer historical pedigree as a task for skilled volunteers and therefore may seem like a more valuable kind of research to pursue. Notwithstanding these results, we should be careful before dismissing photo-masking too quickly, for example because there is a statistically significant difference between people under 35, who prefer these masking tasks, and those aged 35 and above, who are instead more active on transcription (Figure 10). The difference could be explained in the light of the stronger appeal of 3D modelling amongst younger volunteers and/or the greater popularity of transcription amongst older ones.

Providing different kinds of applications probably makes the site look more diverse and interesting, but for now there is no evidence that greater variety in kind would lead to a significant increase in the number of tasks submitted by the same individual across more than one application type. For example, although most of the top 20 contributors in the leader board have explored both transcription and masking, they have in fact soon opted for one or the other (whilst still moving between applications of the same type). Hence, in the present state of our knowledge, it seems better (i.e. more effective in terms of responding to participants' interests) to increase the number of examples of the same kinds of application rather than provide lots

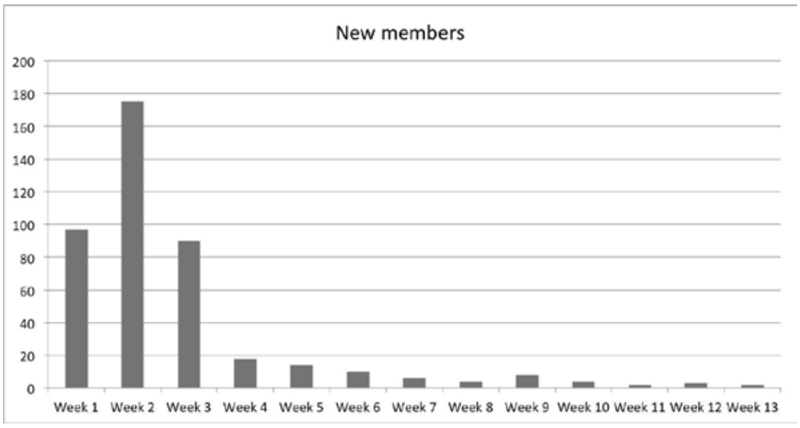


Figure 6: Number of contributors registering themselves as 'members' of the MicroPasts crowd-sourcing website every week. Week 1: 14-20 April 2014; Week 13: 7-13 July 2014.

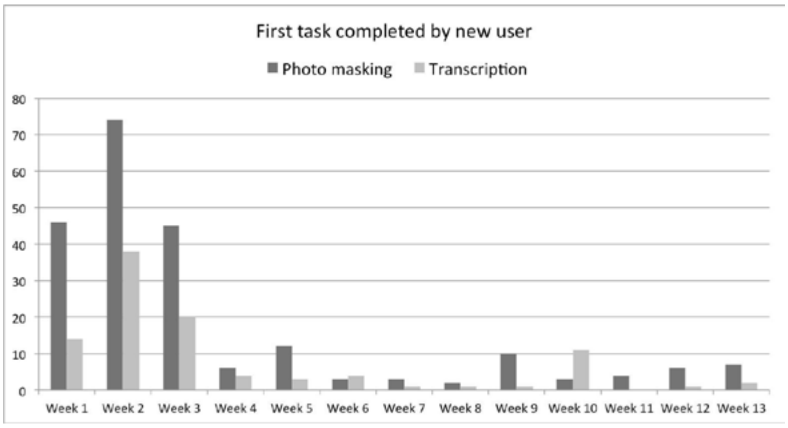


Figure 7: Number and type of 'first tasks' completed by a new user every week. Week 1: 14-20 April 2014; Week 13: 7-13 July 2014. The calculation is based on the number of responses to the survey appearing after the completion of a first task on crowdsourced.micropasts.org.

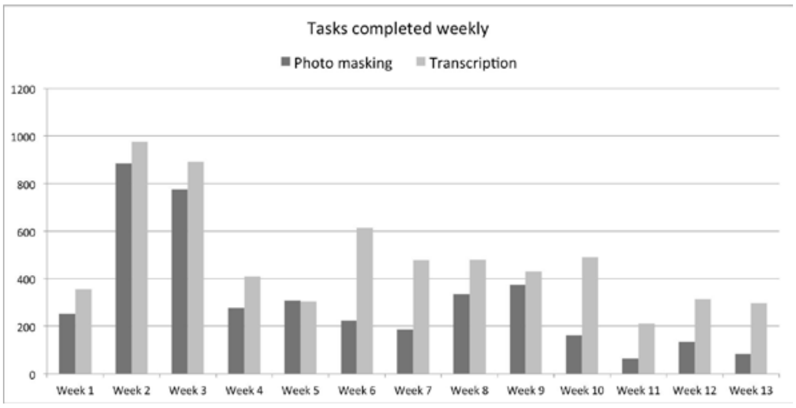


Figure 8: Overall number of tasks completed weekly by anonymous and authenticated users.
Week 1: 14-20 April 2014; Week 13: 7-13 July 2014.

Application	Total No. & kind of tasks	Min	1st Qu.	Median	3rd Qu.	Max
British Museum Bronze Age Index Drawer B16	968 tasks Transcription	1	1	2	5	564
British Museum Bronze Age Index Drawer B8	905 tasks Transcription	1	2	4	13.5	882
British Museum Bronze Age Index Drawer A9	1019 tasks Transcription	1	1	4	11	992
Photo Masking of British Museum Objects	1740 tasks Photo masking	1	1	3	9	1132
Photo Masking the Arreton Hoard	589 tasks Photo masking	1	1	2	8	55
Photo Masking Petrie Museum Shabti	110 tasks Photo masking	1	1	4	6	67
British Museum Bronze Age Index: Arreton Hoard	15 tasks Transcription	1	1	5	11	15

Figure 9: Summary statistics for all the transcription and photo-masking applications that have been available on the MicroPasts website until now.
The 'Min' and 'Max' columns indicate, respectively, the minimum and maximum number of tasks submitted by the same user for each crowd-sourcing application.

Cross-tabulation		Age		Total
Application type		12-34	35+	
Photo masking	Count	128	90	219
	Expected Count	115.7	102.3	219
	% within Application	58.7%	41.3%	100
	% within Age	78.0%	62.1%	70.9%
	% of Total	41.4%	29.1%	70.9%
Transcription	Count	36	55	90
	Expected Count	48.3	42.7	90
	% within Application	39.6%	60.4%	100
	% within Age	22.0%	37.9%	29.1%
	% of Total	11.7%	17.8%	29.1%

[x^2 = 9.458 with 1 df; P = 0.002].

Figure 10: Cross-tabulation between 'Age of contributors' and 'kind of first crowd-sourcing task completed' on MicroPasts.

Photo Masking Petrie Museum Shabti	13 days for 110 tasks	8.5 average daily tasks
British Museum Bronze Age Index Drawer B8	92 days for 878	9.5 average daily tasks
British Museum Bronze Age Index Drawer A9	92 days for 815	8.9 average daily tasks
Photo Masking of British Museum Objects	92 days for 922	10 average daily tasks
British Museum Bronze Age Index Drawer B16	92 days for 687	7.5 average daily tasks

Figure 11: Completion rates for the four British Museum applications launched on the 16th of April 2014, compared to the smaller-scale Petrie Museum application launched on the 2nd of July 2014.

of different types of applications, in order to sustain contributor enthusiasm through time and increase the completion rates of whole research projects.

A final comparison can be made between the completion rates of the four applications that were launched on 16 April 2014 and that of the Petrie Museum Shabti, which was made available to the public on 2 July 2014. In the case of the Shabti app, we tested a different crowd building strategy, consisting of launching an application with a low number of tasks, in order to create a 3D model of one artefact housed in a museum, whose help was also invoked to reach and involve new volunteers. The 110 tasks necessary to photo-mask the Shabti were completed in 13 days, meaning that an average of 8.5 tasks per day were submitted. This number is very similar to other applications (see Figure 11), suggesting that, even if applied systematically, this strategy of small, novel projects may not generate greater commitment from existing users. The result is in line with what was highlighted for example by Holley (2010): the greater the final goal of crowd-sourcing is, the more likely users are to perceive and pursue this as a common goal.

4. Conclusions

The MicroPasts project is the first to experiment with the use of crowd-based methodologies to enable a joint 'contributory', 'co-creative' and 'hosted' approach (Simon 2010) to participatory archaeological research. Building on the long tradition of offline community archaeology in the UK, this initiative is aiming to integrate crowd-sourcing, crowd-funding and forum discussions to facilitate the formation of a cohesive group (no matter how large or small) of academics and volunteers collaborating on the web. The current phase in the project lifecycle is not advanced enough to allow a measurement of the impact of the model on archaeological practice and heritage funding policy (not least because the crowd-funding platform has not been completed yet). However, the evaluation conducted so far already provides insights that are useful not only to inform the future development of MicroPasts, but also (beyond the specific case study) to support others who might consider crowd-sourcing as a means of fostering 'citizen archaeology'.

The discussion of initial findings has revealed how the emergence of an interconnected online community is proving to be a challenge for MicroPasts, as for other crowd-sourcing projects in the cultural heritage and science domains in the past. It is to be expected, however, that time and the activation of the crowd-funding website will help overcome this issue, as volunteers become progressively more confident in the use of the platform and able to see (and comment on) the research outputs produced thanks to their commitment (cfr. the Old Weather project). In addition, successful crowd-funding bids, we hope, will lead to the development of new crowd-sourcing applications and it is our hypothesis that the latter will receive the attention of their donors, a (most probably large) part of whom will already have links with one another offline (e.g. being members of the same archaeological group).

An early assessment of the MicroPasts project also made clear the potential for archaeological 3D models based on crowd-sourced data, especially amongst younger audiences. Novel applications supporting the creation of 3D models of artefacts seem to be worth pursuing, next to those that have a more established tradition (e.g. transcription), as long as the notion of these tasks supporting a clear research agenda can be upheld. However, it will be necessary to study ways of enhancing the current photo-masking exercises proposed by MicroPasts in order to increase the extent to which this application is perceived as worthwhile by volunteers and (as a result) the overall duration of contributors' engagement with masking. More generally, a method for sustaining volunteer participation through time, particularly for crowd-sourcing undertakings with limited time and financial resources secured from the outset (such as MicroPasts), is to focus on diversifying the content and examples of a same kind of application. Analysis of MicroPasts volunteer behaviour so far is instead suggesting that offering a variety of crowd-sourcing types on the same platform is not an effective strategy to respond to contributors' interests and increase completion rates of whole projects. Sustainability, together with the quality of the research data generated and the value assigned by volunteers to their participation, will be measures of the future success of the MicroPasts platform, but also of the adoption of a Public Archaeology approach for a more open and active role of archaeological research institutions in contemporary society.

Acknowledgments

We are thankful to Jennifer Wexler, Neil Wilkin and Rachael Sparks, with whom we have been developing MicroPasts in the past months. We are also grateful to Roger Bland, Ian Carroll, Tim Causer, Nathalie Cohen, Stuart Dunn, Susie Green, Lorna Richardson, Mia Ridge, Stuart Robson, Peter Schauer, Melissa Terras, Lisa Westcott Wilkins and Brendon Wilkins, who have offered invaluable advice on the project. Finally, thank you to all MicroPasts' collaborators. The MicroPasts project has been granted initial funding by the UK Arts and Humanities Research Council.

References

BEALE, N. (2012). How community archaeology can make use of open data to achieve further its objectives. *World Archaeology 44* (4). pp. 612-633. doi: 10.1080/00438243.2012.743252.

BEVAN, A. (2012). Value, authority and the open society. Some implications for digital and online archaeology. In: Bonacchi, C. (ed.). *Archaeology and Digital Communication: Towards Strategies of Public Engagement*. London: Archetype Publications. p. 1-14.

BEVAN, A., LI, X.J., MARTINON-TORRES, M., GREEN, S., XIA, Y., ZHAO, K., ZHAO, Z., MA, S., CAO, W. and REHREN, T. (2014). Computer vision, archaeological classification and China's terracotta warriors. *Journal of Archaeological Science* 49 (2014). pp. 249–254 doi: 10.1016/j.jas.2014.05.014.

BONACCHI, C. (ed.) (2012). *Archaeology and Digital Communication. Towards Strategies of Public Engagement*. London: Archetype Publications.

BONACCHI, C. (2014). Understanding the public experience of archaeology in the UK and Italy: a call for a sociological movement in Public Archaeology. *European Journal of Post-Classical Archaeologies* 4 (2014). pp. 377-400.

CARLETTI, L., GIANNACHI, G., PRICE, D. and MCAULEY, D. (2013). Digital Humanities and Crowdsourcing: An Exploration. In: Proctor, N. and Cheery, R. (eds). *MW2013: Museums and the Web. April 17-20, 2013, Portland, OR*. Silver Spring, MD: Museums and the Web. Available from: http://mw2013.museumsandtheweb.com/paper/digital-humanities-and-crowdsourcing-an-exploration-4/. [Accessed: 24th July 2014].

CAUSER, T., TONRA, J. and WALLACE V. (2012). Transcription maximised; expense minimised? Crowdsourcing and editing *The Collected Works of Jeremy Bentham. Literary and Linguistic Computing*. 27 (2). p. 119-137.

CAUSER, T. and WALLACE, V. (2012). Building A Volunteer Community: Results and Findings from *Transcribe Bentham. DHQ: Digital Humanities Quarterly. 6* (2). Available from: http://www.digitalhumanities.org/dhq/vol/6/2/000125/000125.html. [Accessed: 24th July 2014].

DUCKE, B., SCORE, D., and REEVES, J. (2011). Multiview 3D reconstruction of the archaeological site at Weymouth from image series. *Computers and Graphics*. 35 (2). p. 375-382.

DUCKERS, G. L. (2013). Bridging the 'Geospatial Divide' in Archaeology: Community Based Interpretation of LIDAR Data. *Internet Archaeology* (35). doi:10.11141/ia.35.10

DUNN, S. and HEDGES, M. (2012). *Crowd-sourcing Scoping Study. Engaging the Crowd with Humanities Research*. Report for the UK Arts and Humanities Research Council Connected Communities Scheme. Available from: http://crowds.cerch.kcl.ac.uk/wp-content/uploads/2012/12/Crowdsourcing-connected-communities.pdf. [Accessed: 24th July 2014].

GREEN, S., BEVAN, A. and SHAPLAND, M. (2014). A comparative assessment of structure from motion methods for archaeological research. *Journal of Archaeological Science*. 46. pp. 173-181.

HAYTHORNTHWAITE, C. (2009). Crowds and communities: light and heavywork of peer production. In: Haythornthwaite, C. and Gruzd, A. (eds) *Proceedings of the 42nd Hawaii International Conference on System Sciences. Los Alamitos, CA: IEEE Computer Society*. Available from: https://www.ideals.uiuc.edu/handle/2142/9457. [Accessed: 24th July 2014].

HOLE, B. (2012). A Call for Open Scholarship in Archaeology. In: Bonacchi, C. (ed.) *Archaeology and Digital Communication. Towards Strategies of Public Engagement*. London: Archetype Publications. pp. 114-126.

HOLLEY, R. (2010). Crowdsourcing: How and Why Should Libraries Do It? *D-Lib Magazine*. 16 (3/4). doi:10.1045/march2010-holley.

HOWE, J., (2006). The Rise of Crowdsourcing. *Wired*. 14 June 2014. Available from: http://archive.wired.com/wired/archive/14.06/crowds.html. [Accessed: 24th July 2014].

JAMES, M. R. and ROBSON, S. (2012). Straightforward reconstruction of 3D surfaces and topography with a camera: Accuracy and geoscience application. *Journal of Geophysical Research* 117. F03017. doi: 10.1029/2011JF002289.

KANSA, E., KANSA, S. and WATRALL, E. (eds) (2011). *Archaeology 2.0: New Approaches to Communication and Collaboration*. Los Angeles: Cotsen Institute of Archaeology.

LAKE, M. (2012). Open archaeology. *World Archaeology*. 44.4. p. 471-8. http://dx.doi.org/10.1080/00438243.2012.748521.

LISTER, M., DOVEY, J., JIDDINGS, S., GRANT, I. and Kelly, K. (2009) *New media: a critical introduction*. New York: Routledge.

MASINTON, A. (2014) Crowd-sourcing projects. Personal communication [email], 12 March 2014.

MCQUAIL, D. (2005) *McQuail's Mass Communication Theory*. London: SAGE.

OOMEN, J. and AROYO, L. (2011). Crowdsourcing in the Cultural Heritage Domain: Opportunities and Challenges. *5th International Conference on Communities and Technologies, 29 June-2 July 2011, Brisbane, Australia*. p. 138-149. Available from: http://www.iisi.de/fileadmin/IISI/upload/2011/p138_oomen.pdf [Accessed: 24th July 2014].

OWENS, T. (2013). Digital Cultural Heritage and the Crowd. Curator: *The Museum Journal*. 56 (1). pp. 105-106.

PROCTOR, N. (2013). Crowdsourcing - an Introduction: From Public Goods to Public Good. Curator: *The Museum Journal* 56 (1). pp. 121-130.

REMONDINO, F., DEL PIZZO, S., KERSTEN, T. P. and TROISI, T. (2012). Low-cost and open-source solutions for automated image orientation - a critical overview. In: Ioannides, M., Fritsch, D., Leissner, J., Davies, R., Remondino, F., Caffo, R. (eds). *Progress in Cultural Heritage Preservation.* P*roceedings of the 4th International Conference, EuroMed 2012. Euromed2012, Lemessos, Cyprus. October 29-November 3, 2012*. Berlin, Heidelberg: Springer. pp. 40-54.

RICHARDSON, L. (2013). A Digital Public Archaeology? *Papers from the Institute of Archaeology* 23(1). p.1-12. doi: http://dx.doi. org/10.5334/pia.431.

RIDGE, M. (2013). From Tagging to Theorizing: Deepening Engagement with Cultural Heritage through Crowdsourcing. Curator: *The Museum Journal* 56 (4). pp. 435-450.

ROGERS, R. (2013). *Digital Methods*. Cambridge, Massachusetts: MIT Press.

SIMON, N. (2010). *The Participatory Museum* (Museum 2.0). Available from: http://www.participatorymuseum.org/. [Accessed: 24th July 2014].

VERHOEVEN, G., DONEUS, M., BRIESEC, C. and VERMEULEN, F. (2012). Mapping by matching: a computer vision-based approach to fast and accurate georeferencing of archaeological aerial photographs. *Journal of Archaeological Science* 39. pp. 2060-2070.

The ACCORD Project: Archaeological Community Co-Production of Research Resources

Stuart JEFFREY,[1] Alex HALE,[2] Cara JONES,[3] Sian JONES,[4] Mhairi MAXWELL[1]

[1] Digital Design Studio, Glasgow School of Art, UK
[2] Royal Commission on the Ancient and Historical Monuments of Scotland (RCAHMS), UK
[3] Archaeology Scotland, UK
[4] University of Manchester, UK
s.jeffrey@gsa.ac.uk, alex.hale@rcahms.gov.uk, c.jones@archaeologyscotland.org.uk,
sian.jones@manchester.ac.uk, m.maxwell@gsa.ac.uk

Abstract

This paper introduces the AHRC funded ACCORD project, a partnership between the Digital Design Studio at the Glasgow School of Art, Archaeology Scotland, the University of Manchester and the RCAHMS. The ACCORD project examines the opportunities and implications of digital visualisation technologies for community engagement and research through the co-creation of 3D models of historic monuments and places. Despite their increasing accessibility, techniques such as laser scanning, 3D modelling and 3D printing have remained firmly in the domain of heritage specialists. Expert forms of knowledge and/or professional priorities frame the use of digital visualisation technologies and forms of community-based social value are rarely addressed. Consequently, the resulting digital objects fail to engage communities as a means of researching and representing their heritage. The first part of this paper presents how the ACCORD project seeks to address this gap through the co-design and co-production of an integrated research asset that encompasses social value and engages communities with transformative digital technologies. The second half of this paper (section 4) presents a case study of an ACCORD project based in Argyll which highlights the nature of community relations with expert groups, issues of archaeological authority and the transformative power of co-production using digital recording techniques.

Keywords: Community Co-production, 3D, Visualisation, RTI, Authenticity, Social Value

1. Introduction

There have now been nearly three decades of research and development of digital visualisation technologies in archaeology and heritage. Approaches that utilise photogrammetry, laser scanning, 3D modelling, and real time virtual reality have become standard practice in the academic archaeological sphere, commercial archaeological ventures and cultural heritage management. However, there is little community engagement with digital visualisation technologies, despite interest in the technologies themselves (Banks, 2011). Expert forms of knowledge and/or professional priorities, rather than community ones, invariably inform digital visualisations. Furthermore, digital visualisations commissioned in these traditional sectors rarely integrate forms of community-based social value relating to the historic environment into the recording exercise. Consequently, the resulting digital objects often fail to engage communities as a means of researching and representing their heritage. The low levels of community use and re-use, let alone co-production, of these resources also highlights concerns relating to perceptions of authenticity and value. Recent work by Latour and Low (2011) suggests that replicas and reproductions of original historic objects can indeed acquire authenticity, but that this is dependent on the nature of their production, as well as the quality of the resulting outputs. Other research suggests that relationships between people, places and things are important in the experience of authenticity in relation to historic monuments, buildings and objects (e.g. Jones, 2009; Macdonald, 1997). It is not surprising then, that 3D models produced by 'experts' can seem disconnected, clinical, and irrelevant to the broader communities of interest that accrue around heritage places.

Community engagement has been increasingly prominent in archaeology and heritage over the last two decades (Newman and McLean, 1998; Smith and Waterton, 2009). Community archaeology projects are widespread, usually defined by community participation in the design, management, implementation and/or dissemination of archaeological field research (see Marshall, 2008; Moser *et al.*, 2002). Furthermore, in heritage management, social value has become an important aspect of international conservation charters and national policies. Defined as a collective attachment to place that embodies meanings and values important to communities (Johnston, 1994:10), social value is something that heritage organisations attempt to integrate into the conservation, curation and management of heritage assets. Sometimes this involves the active participation of community groups.

As 3D digital visualisation technologies are increasingly employed in research and representation in the heritage sector, it is also important that forms of social value and community engagement are incorporated into the production of digital models. Issues around access to equipment, cost of recording and lack of technical expertise are no longer significant barriers to community participation. The on-going fall in hardware and software costs and the simplification of hitherto highly complex recording processes have brought at least some techniques easily within the reach of community groups. The development of freeware with associated support communities and online forums further enhances accessibility. The time is therefore ripe for the development of new approaches that bring academic researchers, heritage organisations and community partners together to explore the multiple opportunities created by the co-design and co-production of 3D digital visualisations of historic monuments and places. In particular, it will be important to investigate how co-design and co-production impact on the forms of social value and authenticity attributed to the resulting 3D models. What kinds of values are produced through co-design and co-production? Do the resulting models acquire different forms of authenticity and significance than those that are produced entirely within professional/ academic domains? What is the potential of such models for community re-use and research?

2. The ACCORD Project: aims and objectives

The ACCORD project, funded by the UK Arts and Humanities Research Council's Connected Communities and Digital Transformations programmes, is a partnership between the Digital Design Studio at the Glasgow School of Art, Archaeology Scotland, the University of Manchester and the RCAHMS. At the time of writing we are mid-way through the project, which started in late 2013 and will complete in spring 2015.

The primary aim of ACCORD is to examine the opportunities and implications of digital visualisation technologies for community engagement and research through the co-design and co-production of 3D models of historic monuments and places. The project actively engages existing community groups in the process of designing and producing 3D records and models of heritage places, many of which they have ongoing relationships with. A range of recording and modelling techniques is being made available for selection during the co-design process with each community heritage group. These include: consumer level photogrammetry; Reflectance Transformation Imaging (RTI); and 'white light scanning'. Depending on the outcomes of the co-design process, time of flight laser scanning has also been deployed for a sub-set of groups. 3D printing technology has also been used in select cases to create physical models from the captured data. The datasets resulting from the project will contain full contextual and technical metadata generated in collaboration with the community groups involved.

The ACCORD Project also aims to reflect on the nature of the relationships between community groups, digital heritage professionals, and the outputs they have created, particularly in comparison to similar outputs produced in more traditional professional domains. The participation of interested communities in the design process will allow contemporary social values associated with heritage places to be explored and embedded in the resulting digital records and 3D objects, as well as the associated contextual metadata. Groups are encouraged to integrate existing research and community-generated material relating to their chosen heritage place(s) such as survey work, independent research, oral histories, folklore, and/or photographs. Throughout each ACCORD sub-project, the team also conduct qualitative reviews of the transformative aspects of the process, investigating changes in attitude to 3D recording technologies during the life of each project, as well as the forms of significance, authenticity, and value acquired by the resulting 3D objects.

Finally, the ACCORD Project aims to broaden capacity for the creation and reuse of digital visualisation technologies in community heritage activities and research and to increase awareness of the potential of 3D technologies for community heritage practice. This will be facilitated through the production of an open-access dataset. In addition to direct in-person collaborative work with the ACCORD groups, social media including the ACCORD blog, Twitter account and Facebook page, plays an important role by increasing access to the work. Beyond this we hope to create a community of communities by facilitating interaction between the ACCORD community groups that will extend beyond the lifetime of the project. The long-term legacy of the project is a research asset prepared for permanent archive by the community groups with the support of the ACCORD project team. This will be available via the UK's Archaeology Data Service (ADS) in a 'special programme' format i.e. an integrated archive specifically designed to be capable of regular update during the life time of the project. An interactive website facilitating promotion, communication, comment and social media integration exists separately. The technical team is actively guiding the co-production of technical metadata to ensure that each data asset has a well-formed OAIS Submission Information Package (ISO 14721:2012, version 2, published 2012-08-21) enabling rapid archival ingest and dissemination of the datasets. These OAIS Archival Information Packages will also be deposited with the RCAHMS.

3. Project methodology

3.1. Group selection

Since ACCORD only runs for 15-months, groups who already have well-established working relationships with partners on the ACCORD team were approached in the first instance. Ultimately we will work with 10-12 community groups, each of which will form a sub-project within the overall ACCORD Project. A number of community groups have been drawn from Archaeology Scotland's highly successful Adopt-a-Monument Scheme (out of 55 groups) and one from Glasgow Life's community scheme. In the first instance, a process of facilitated self-selection was used to identify a cohort of groups with an interest in 3D recording, printing or digital visualisation for artistic reconstruction and heritage representation. Within this process of self-selection, the ACCORD project team also endeavoured to include a variety of different kinds of community groups and a wide geographical distribution across different regions of Scotland and between urban and rural environments. While a range of different historic monuments, buildings and objects is desirable, it is not an a priori consideration since the collaborative process at the heart of ACCORD means that the subject matter is prescribed by community interests and attachments. The community groups fall into a number of broad, sometimes overlapping, categories ranging from long established archaeological/heritage societies to local interest groups, action groups and loosely formed communities of interest/practice.

3.2. Timetable

Each ACCORD sub-project is scheduled to take place over a 2-3day period. The process is split into a number of phases. Phase 1 consists of focus group designed to explore baseline data on the nature of the group and the heritage places they engage with. Phase 2 is the co-design phase where the community groups and ACCORD project team together select targets and technologies. Phase 3 marks the start of co-production with field recording. Phase 4 focuses on data processing and initial model/RTI production. This is followed by the final Phase of engagement, which involves a further focus group reflecting on the process, issues of significance, authenticity and value, and the impact of the technologies.

3.3. Co-design and co-production

Collaboration in the design and production of digital 3D records and models sits at the heart of the ACCORD project. In the first instance, the historic monuments, buildings and objects selected for recording are based on the specific interests and attachments of the community groups concerned. These are ascertained through the initial focus group. Co-design must also take into account timescales and technologies. The recording and modelling must be feasible within the timetable described above, and amendable to the technologies available. Consumer level photogrammetry and Reflectance Transformation Imaging (RTI) are favoured because they are more accessible and enable full community participation. Depending on the outcomes of the co-design process, time of flight laser scanning is sometimes also deployed. 3D printing technology is also used in select cases to create physical models from the captured data. Technical staff from Glasgow School of Art's Digital Design Studio facilitate community recording and modelling, but the ACCORD Project does not undertake formal training. In the case of laser scanning DDS technical staff usually lead the recording and modelling process, but they do so according to the collaborative design and they encourage as much community participation as possible. In the interests of co-production, we aim to always include photogrammetry and/or RTI work alongside laser scanning.

3.4. Social Value Capture

During the co-design process, two semi-structured focus group sessions are held, one at the beginning of our work with each community group and one at the end. The first focus group explores the nature of the group and the historic monuments, buildings and/or objects they are interested in. There is a discussion of the group's make-up and history. There is also an in-depth exploration of the meanings and values associated with specific monuments, buildings and/or objects. Feelings of attachment and ownership are examined. The second focus group deals with the group's experience with 3D visualisation, including the recording and modelling process. It also explores the group's response to the models themselves and the forms of ownership and authenticity associated with them, if any. Finally, the ways in which participation in 3D visualisation has impacted on the interests and attachments of the group is considered.

In the intervening period between the focus groups, ACCORD project members actively participate in co-design and co-production. During this they observe and record community participants' responses, particularly changing attitudes to 3D technologies and the heritage places being recorded/modelled. A copy of social-media interactions, email and written communication is also being kept (including the project's own blog, facebook and twitter accounts). The extent of this social media will vary between groups and individuals, and is affected by practicalities (such as access and familiarity with the use of the internet in this way).

3.5. Ethics and intellectual property

Due attention has been given to issues of inclusion and power relationships, both within and between community groups, as well as between community groups and project investigators/partners. The project builds on the relationships developed by its partners who are experts in the community engagement field. Throughout the community engagement process, careful attention has been given to the ethical dimensions of this research, particularly issues of informed consent. Potential tensions surrounding open-data and community feelings of ownership and intellectual property are also the subject of overt discussion. We draw heavily on the Ethical Guidelines created through the AHRC Connected Communities Programme (CSCJA, NCCPE 2012) (http://connected-communities.org/). The seven ethical principles of mutual respect, equality and inclusion, democratic participation, active learning, making a difference, collective action, and personal integrity underpin all aspects of our engagement. More specific legal issues relating to ownership of the resulting assets (intellectual property and copyright) are openly negotiated with the community groups. Groups are strongly encouraged to adopt as open a regime as possible (i.e. Creative Commons CC0) for the ACCORD outputs that they have co-created. Attribution and non-commercial re-use of data can be guaranteed by the CC-NC-BY license where appropriate.

4. Case Study: Colintraive and Glendaruel Development Trust Archaeology and History Group

At the time of writing three sub-projects have been undertaken, each with a particular character and set of issues that help to highlight the nature of community relations with expert groups, archaeological authority and digital recording techniques. The case study below focuses on work at Colintraive and Glendaruel in the west of Scotland. The other two completed projects took place at Dumbarton Rock on the river Clyde working with rock climbers specifically interested in the sporting heritage of the site (as opposed to the renowned castle atop the rock), and with the Friends of Glasgow Necropolis in the centre of Glasgow city who have specific conservation and recording objectives which were ideally suited for the application of RTI.

Colintraive and Glendaruel is situated in Argyll, in the Parish of Kilmoden, on the West coast of Scotland. The ACCORD team worked with the Colintraive Glendaruel Development Trust (CGDT) Archaeology and History Group on 6th April and 21st – 22nd June 2014. The CGDT Archaeology and History Group is a recently formed sub-committee of the CGDT (http://cgdt.org/about/). The CGDT is governed by a Board of Directors, elected from and by the community. It currently has c.40 official members; membership is open to all those who are permanently resident in Colintraive and Glendaruel and on the electoral roll. The Trust is established as a Company Limited by Guarantee (SC350010) and is registered as a Charity (SC040002). The CGDT Archaeology and History Group was founded following an Archaeology Scotland Adopt-a-Monument workshop held on the 23rd and 24th of November 2013. The group is currently made up of around 12 passionate and driven members of the local community, who see their work in the spirit of the CGDT as 'Making Colintraive and Glendaruel a better place to live and work'.

Together we recorded and modelled 3 monuments located in the Kilmoden Parish: the Lephinkill Chambered Cairn; a cup and ring marked stone; and a World War One memorial to two men who fell at Gallipoli. For this sub-project photogrammetry, RTI and laser scanning were deployed. As will become clear, the value of the monuments discussed here is in part a product of their relation to the Stronafian Forest which the CGDT purchased in 2013 and the ability of these technologies to make the invisible visible.

The CGDT Archaeology and History Group describe themselves as 'incomers', holiday-homers, part-time inhabitants and 'locals' who have family ties going back generations. The group's recent interest in the archaeology and history of the area is framed within the CGDT's purchase of the Stronafian Forest and the sense of ownership and belonging it entails. The Group express a strong sense of continuity with the past. This motivates some to learn more about their landscape; indeed for some this is a duty or responsibility. For some, 'getting the facts right' about the dates and national significance of the archaeological remains is a very important issue. For others it is more about 'the story' and a passion to get more 'in touch'. The CGDT purchased Stronafian Forest (which covers 600 ha) as a social enterprise project in February 2013. This desire for development is motivated by a desire to enhance the area's social, economic, and historic value, while it is also seen to represent the ambition and resilience of the ColGlen community. There is a strong sense of pride in this vibrant, exciting and special place. It is not just about legal ownership; working together as a community is seen as a key part of taking ownership of 'our forest'.

Against this backdrop of ambition and resilience, there is also a sense that the ColGlen community is a 'community at risk'. It is hoped that the Stronafian forest will form a social hub and focal point, for the dispersed inhabitants. Purchase of the Forest is associated with amenity, educational, and wellbeing benefits for the community. The community forest is also recognised as an economic asset for tourism, leisure and local businesses, offering the opportunity to create something that will 'entice' and 'attract' public interest, and encourage those who are passing through the area to stop and explore. Archaeology and history have a prominent place in this vision.

4.1. The ACCORD-CGDT Archaeology and History Group project monuments

The first site of interest identified was Lephinkill Chambered Cairn (NGR NS 0027 8432). This chambered cairn is situated above the valley floor about 600m E of Clachan of Glendaruel in a clearing within a forestry plantation. Described as a 'Clyde-type long cairn', it appears as a vegetation covered irregular mound of stones. The chamber is at the north end and is entered through a concave facade, now blocked by stones. It has been disturbed and robbed for building material. The cairn is not scheduled but is historically significant in terms of understanding the distribution of Neolithic settlement and burial. It is included in the National Monument Record (Canmore ID 40540 http://canmore.rcahms.gov.uk/en/site/40540/details/lephinkill/), volume 6 of the RCAHMS Argyll Inventory and Henshall's *Chambered Cairns of Scotland* (Henshall, 1972: 329-30). In terms of social value, current local interest in this cairn has recently accrued since the purchase of this forest and is framed by its location here. Despite the focus on its ruinous state in monuments records, it is seen as one of the most pristine monuments in the forest and is perceived to be aesthetically pleasing. The age of the cairn is also important, being seen as the oldest monument and the anchor for a narrative of human inhabitation of the area. Other factors which underpin the interest of the CGDT History and Archaeology Group in this monument are its visibility and accessibility, which make it a good focal point for a footpath positioned at a good viewpoint above the valley. Therefore, the group see it as an attraction and 'hub' from which a sense of connection to the past emerges, and from which people can further explore the landscape.

For the ACCORD recording exercise, the CGDT Archaeology and History Group expressed a strong desire

Figure 1: The Lephinkill cairn during laser scanning using a Leica C10 scanner.

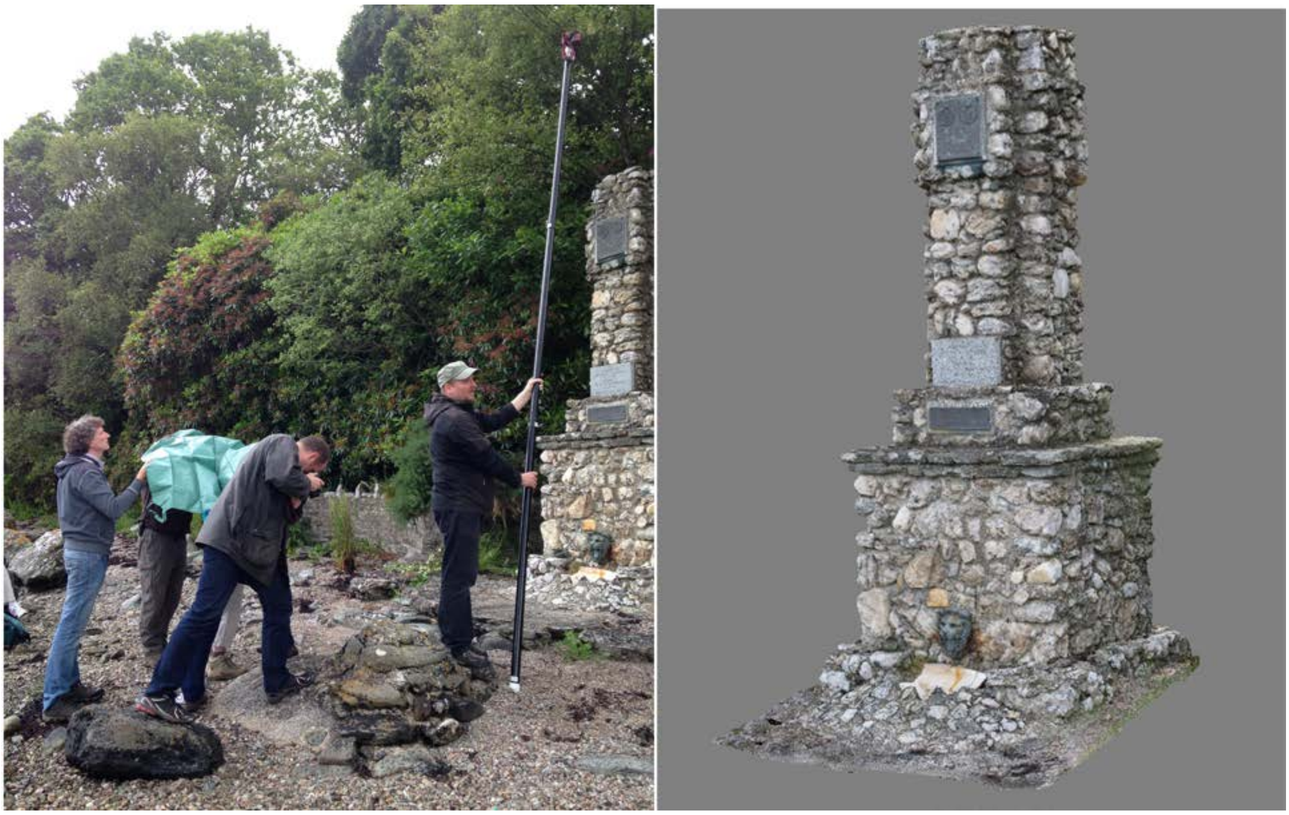


Figure 2: Left – The group capturing images for photogrammetric modelling. Right – A screenshot of the completed model.

to focus on the cairn for the reasons outlined above; as a visible, ancient and aesthetically pleasing monument that people were increasingly attached to. However, the monument was not conducive to photogrammetry, so over the weekend 21st and 22nd of June it was topographically recorded by low resolution laser scan.

The second identified site was a Cup and Ring Marked Stone (NGR NR 9985 8408). This decorated boulder is situated in a recent deciduous plantation 100m SE of the road bridge (A886) over the Clachan Burn. The Canmore entry describes it as a sunk schist boulder. Its upper surface ‘bears ten cupmarks, one of which is surrounded by a single

Figure 3: Left -The stone being prepared for RTI recording, Right – a screen shot of the resulting RTI.

ring and another by a keyhole-shaped ring' (Canmore ID 39960). This Neolithic rock art is described as both hard to find and difficult to discern except 'when wet and in low sun'. The boulder is not a Ancient Monument, but it is included the national monuments record, and Morris's The prehistoric rock art of Argyll (Morris 1977).Although maps in the Colintraive Community Hall show this boulder located in the forest, to the group this monument was shrouded in mystery and none of the members present had visited it. Nevertheless its chronological and cultural connection to the cairn instigated a desire to locate it. When it was eventually found, it was, fortuitously, near to one of the 'desire lines' (proposed path route) leading up to cairn. With the ACCORD project team the group 3D modelled and recorded it using photogrammetry and RTI (RTIBuilder and PTMFitter).

Finally the group selected a WWI War Memorial ('The Gallipoli memorial') (NGR NS 028 750): This memorial, located on a rocky foreshore is dedicated to two brothers, R.F. McKirdy and P.M. McKirdy, who fell at the Battle of Gallipoli. It is not included in the National Monument Record but there is an entry in the Imperial War Museum online War Memorial Archive where it is described as an ornamental drinking fountain and unworked stone cairn. There are bronze lions on the fountainhead and the Badges of RNVR and Argyll and Sutherland Highlanders. The date of erection is unknown and it is noted that coastal erosion has affected the memorial. Only one member of the group, having only moved to Colintraive in September from Australia, knew about the existence of 'the Gallipolli memorial'. On discussion, another member of the group remembered that they had a personal connection to the son of one of the individuals commemorated on the monument. The whole group felt it was important to record this monument in order to 'raise awareness' of this 'overlooked' site. It has also been since noted that next year is the 100 year anniversary of the deaths of the McKirdy brothers. The memorial was modelled using photogrammetry (Photoscan) by the Group in collaboration with the ACCORD team on the weekend of 21st and 22nd June.

4.2. The impact of 3D visualisation in Colintraive and Glendaruel

It was clear that the process of recording and modelling the monuments had an impact on how people perceive and engage with them. Only the chambered cairn was a strong focus of interest before the ACCORD weekend, but the opportunity to explore their local heritage in a new engaging way resulted in new relationships with the cup and ring marked boulder and the War Memorial. As one group member noted: 'It's the things that are on your doorstop which you don't explore.' In the case of 'the Gallipoli memorial', a member of the group who had just returned after 20 years living in Australia and New Zealand decided to find out more about the McKirdy brothers since 'Gallipoli, during World War One, is an event etched in the stone memorials and hearts of every ANZAC'. His personal account is now on the ACCORD Blog (http://accordproject.wordpress.com/). In the case of the rock art, the value of 3D recording lay in rendering the invisible visible. As one member of the group said, 'I couldn't see why it was important and exciting before… now I see why!'

The group were extremely enthusiastic about the potential for photogrammetry and RTI. The results were recognised as bringing the monuments to life in a visually stunning way. There was also excitement about using the 3D models in various ways: for communication, analysis (especially RTI), raising the profile of the forest, increasing access to the archaeology for educational and interpretation purposes, encouraging visitors to the sites, and enhancing the monuments which are already there, though importantly never replacing them. Members of the CGDT History and Archaeology Group felt they offered a low cost to failure; constraints of time and money were not an issue, while results did not necessarily need professional or specialist validation

5. Conclusion and future work

The innovative nature of the project derives from its use of digital visualisation technologies to engage communities in the co-creation of 3D models of historic monuments and places. There is a great deal of research demonstrating that community participation both enhances existing significance of the historic environment and in some cases generates new forms of significance (Smith and Waterton 2009; Moser et al 2002; Marshall 2009). The case study above shows clearly that a co-design and co-production process has a transformative affect on the relationship between the group, their heritage, and the digital outputs of the recording process. It is too early to say to what extent that transformation is long lasting or profound. However, these initial results suggest that the approach adopted by ACCORD has the potential to deal with the long-standing problems of lack of engagement with digital outputs. It may also counteract the wariness of some non-specialist groups to metaphorically take ownership of local heritage, which they perceive as the domain of experts, professionals and authoritative bodies. Through a collaborative process – addressing what is to be recorded, how it is recorded, who records it, and how it is modelled – a network of relationships (that is participatory, inclusive and local) is created around the outputs. There is also some evidence emerging that the creation of this network of relationships impacts on the perceived authenticity of the digital outcomes and therefore their relevance and re-use potential.

The ACCORD project will continue to engage with a further 5-8 groups throughout the remaining project period, while we will additionally revisit a selection of communities in order to evaluate the legacy effects of the project. The use of digital technologies to enhance and generate forms of social significance is an important outcome, adding distinctive value to existing heritage assets and our understandings of them. Community groups will be able to draw on the resulting digital datasets for various strategic purposes, including public access and presentation, education, and tourism initiatives. The records and models resulting from the project will also provide important research resources. For instance community groups will be able to continue to integrate them into their research initiatives, such as survey work, interpretation, oral history projects or forms of artistic intervention. The integration of contemporary social value in the design and production of the 3D models and the associated contextual data means that the resulting assets also provide heritage managers with a new source of information about the local significance of heritage places. Finally for academic researchers interested in community heritage initiatives and the generation of social value, the project promises to provide an important exemplar for analysing how digital technologies mediate the complex relationships that people have with heritage places.

References

BANKS, S. (2011). Connected Communities scoping study, http://www.ahrc.ac.uk/Funding-Opportunities/Research-funding/Connected-Communities/Scoping-studies-and-reviews/Documents/Community-based%20Participatory%20Research.pdf

CSCJA, NCCPE (2012). http://connected-communities.org/, accessed 29/07/2014.

HENSHALL, A. S. (1972). *The chambered tombs of Scotland*, vol.2 Edinburgh: Edinburgh University Press. pp. 329-30.

ISO (2012). Space data and information transfer systems Open archival information system (OAIS) Reference model, ISO 14721:2012, version 2, https://www.iso.org/obp/ui/#iso:std:iso:14721:ed-2:v1:en

JONES, S. (2009). Experiencing authenticity at heritage sites: some implications for heritage management and conservation. *Conservation and Management of Archaeological Sites*.12(2). pp. 133-147.

JOHNSTON, C. (1994). *What is Social Value? A discussion paper*. Canberra: Australian Government.

LATOUR, B. & LOWE, A. (2011). The migration of the aura, or how to explore the original through its facsimiles. In: Bartscherer T (ed.). *Switching Codes*, 275-98. Chicago: University of Chicago Press.

MARSHALL, Y. (2009). Community archaeology. In: B. Cunliffe, C. Gosden and R. Joyce (eds). T*he Oxford Handbook of Archaeology*. Oxford: Oxford University Press, pp. 1078-1102

MACDONALD, S. (1997). A people's story: heritage, identity and authenticity, In: C Rojek and J Urry (eds). *Touring Cultures: transformations of travel and theory*. London: Routledge. pp. 155–175.

MORRIS, R. W. (1977). *The prehistoric rock art of Argyll*. Poole: Dolphin Press.

MOSER, S., GLAZIER, D., PHILLIPS, J. E., Lamya NASSER EL NEMR, SALEH MOUSA, M. NASR AIESH, R. RICHARDSON, S., CONNER, A., and SEYMOUR, M. (2002). Transforming archaeology through practice: strategies for collaborative archaeology and the community archaeology project at Quseir, Egypt. *World Archaeology* 34. pp. 220–48.

NEWMAN, A. & McLEAN, F. (1998). Heritage builds communities: the application of heritage resources to the problems of social exclusion. *International Journal of Heritage Studies* 4(3). pp. 143-53.

RCAHMS (1988). *Argyll: an inventory of the monuments volume 6: Mid-Argyll and Cowal, prehistoric and early historic monuments*, Edinburgh: The Royal Commission on the Ancient and Historical Monuments of Scotland. Page(s): 48, no. 18 Plan, 48 (B).

SMITH, L. & E. WATERTON (2009). *Heritage, Communities and Archaeology*. London: Duckworth. The ACCORD blog, http://accordproject.wordpress.com/, accessed 30/07/2014.

The CGDT website, http://cgdt.org/, accessed 29/07/2014

Chapter 5. Archaeological Information Systems

12000 Years of Human Occupation, 3 Meters Deep Stratigraphy, 12 Hectares… A Geographical Information System (GIS) for the Preventive Archaeology Operation at Alizay (Normandie, France)

Sylvain MAZET, Cyril MARCIGN, BrunoAUBRY, Isabelle COMTE, Philippe BOULINGUIEZ

Institut National de Recherches Archéologiques Préventives (INRAP), France
sylvain.mazet@inrap.fr, cyril.marcigny@inrap.fr, bruno.aubry@inrap.fr,
isabelle.comte@inrap.fr, philippe.boulinguiez@inrap.fr

Abstract

An archaeological investigation carried out by INRAP (Institut national de recherches archéologiques préventives) revealed that banks of the river Seine at Alizay (Normandy, France) were occupied during 12000 years. The specifications of the excavation were to study various occupations (Upper Palaeolithic, Neolithic, Iron Age and Middle Age) applying the appropriate methodology for each major period.

A Leica TS15 robotic total station was used to rapidly record archaeological features and store directly in the field pre-formatted GIS data. The aim of the process was to inventory and study, during the excavation, the 120 000 exhumed artefacts using a Microsoft Access database linked to a geodatabase in Esri ArcGIS.

This specific archaeological information system was designed to guide the earth removal, by producing digital elevation model of each stratigraphic context, facilitate the management and inventory of artifacts and spatialize the data produced during the excavation.

Keywords: Preventive archaeology, GIS, Process, Methodology

1. Introduction

An archaeological operation carried out by INRAP (Institut national de recherches archéologiques préventives) revealed that banks of the river Seine at Alizay (Normandy, France) were occupied during 12000 years. In an extensive area of 12 hectares, archaeologists had to excavate a three meters deep stratigraphy covering the Middle Age, the Iron Age, the Neolithic, to finally reach the earliest settlement from the upper Palaeolithic (Marcigny et al., 2013). The specifications of the excavation were to study these various occupations applying the appropriate methodology to each major period.

A Leica TS15 robotic total station was used to rapidly record archaeological features and store directly in the field pre-formatted GIS data such as number ID, stratigraphic unit, nature of the artefacts, etc. The aim of the process was to inventory and study during the excavation the 120 000 exhumed artefacts using a Microsoft Access database linked to a geodatabase in Esri ArcGIS.

This specific archaeological information system was designed to guide the earth removal, by producing digital elevation model of each stratigraphic contexts, facilitate the management and inventory of artifacts and spatialize the data produced by specialists such as palynologist, malacologist, geomorphologist, geophysicians, finds specialists (lithic, pottery), etc.

Beyond technical and methodological aspects, this GIS is aimed to produce new data that could enrich our understanding of human occupations at Alizay like habitat, burials or activity areas, through time.

The present publication will explain the entire archaeological data process insisting on its different steps such as field recording, database design and management as well as GIS use for a large scale archaeological excavation.

2. Site specification

Situated in front of the city of Pont de l'Arche, in Normandy (NW of France), 40 hectares on the right bank of the river Seine, were archaeologically investigated, by B. Aubry, in 2007 and 2009. It revealed the high archaeological interest of this alluvial plain (Aubry *et al.*, 2011). 12 ha of this area located on the territories of the villages of Alizay and Igoville (Eure) were directly concerned by the future exploitation of gravel quarries owned by Lafarge and Cemex (Figure 1).

The presence of gravel under 3 to 4 meters of alluvial deposits, sealing human occupations from the upper Palaeolithic to the Middle Age, has required a preventive archaeological excavation conducted by INRAP in 2011 and 2012 during one year and half. More than 70 people were involved in this project which represents approximately 7000 man-days. Up to six 24t caterpillar excavators were employed at the same time during the digging as well as to remove 275 000 m3 of earth.

The scientific aim of the operation was the detailed study of the pre-and protohistoric occupations applying a multiscalar approach. The four main topics are focusing on the stratigraphy (the different deposits and the study of their evolution), the palaeoenvironment, the human occupation from a cultural and a chronological point of view and the

Figure 1: Aerial view of the excavation (H. Paitier).

palaeoethnology. In order to address these questions, a multi-proxy study has been set up involving molluscan analysis, charcoal analysis, geophysics, zooarchaeology, palynology, geomorphology, sedimentology and micromorphology.

The archaeological investigation revealed the presence of 2 700 archaeological features (hearth, postholes, graves, oven, ditches, etc.) associated to 114 000 artifacts (68 000 lithics, 34 000 ceramics, 800 metallic finds, etc.). 110 different occupations or settlements have been identified and put into a chronological context by 180 C14 dating.

3. Methodology

Regarding the stratigraphy, the extent of the study area, the volume of earth, the schedule and the amount of data, the issue of the data recording process appeared as a cornerstone of this big scale excavation. A specific

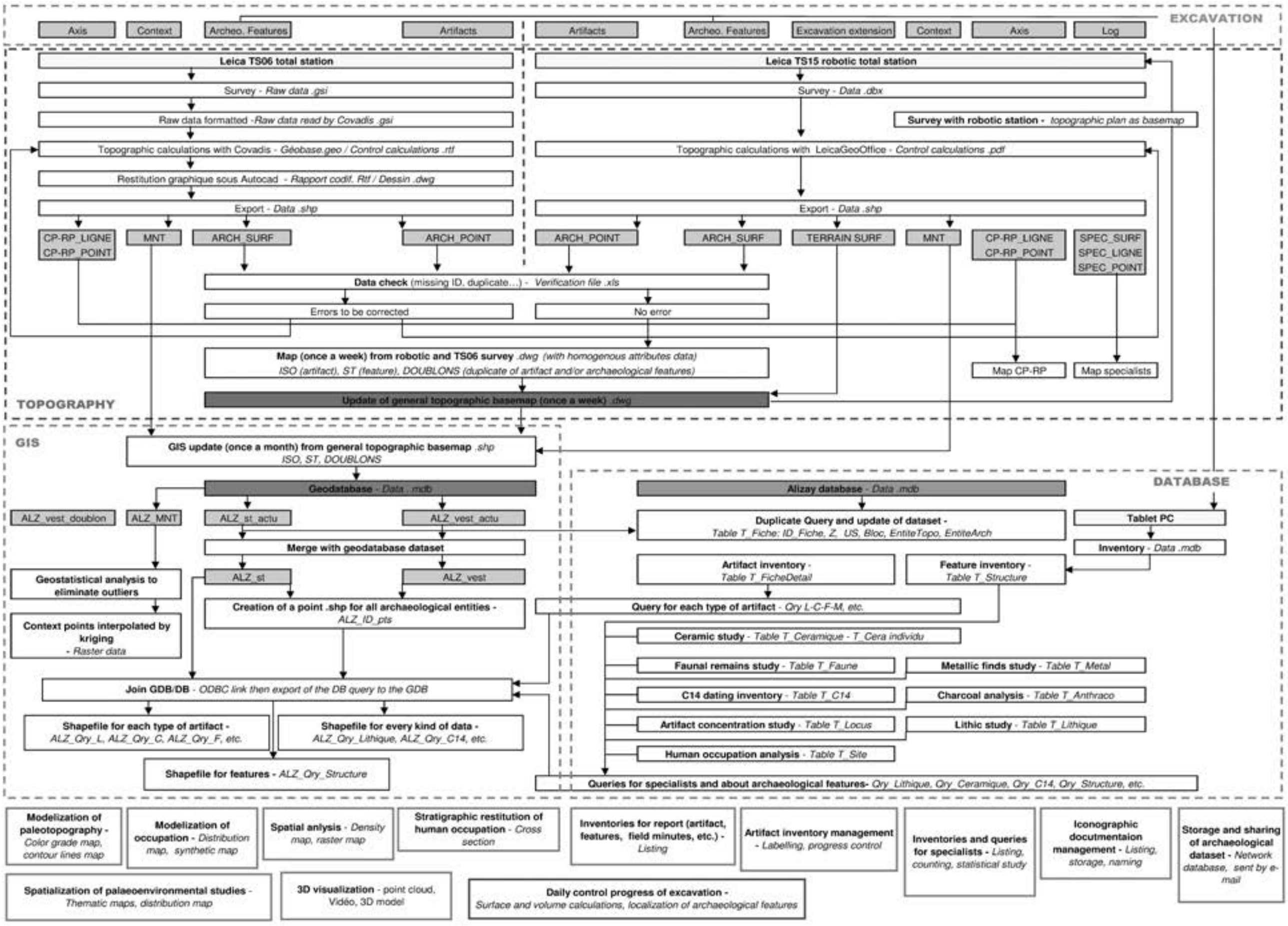


Figure 2: Alizay data recording and analyzing process.

methodology had to be set up to answer the followings questions:

– How to deal with so much data?
– How to record, describe, store and spatialize the data?

In the field, the topographical recording has to be efficient and the archaeological recording exhaustive. In the office, the operation needs to have a sustainable archaeological data archiving and the archaeological and palaeoenvironmental data must be spatialized.

A solution is to use a robotic total station for the topographical recording, a database to store the data and a GIS to spatialize these data. In this robust process of data recording, each archaeological remains (artifact or feature) has to be defined by a unique ID from 1 to n. This number identifies the graphic entity recorded by the topographer, the archaeological feature (using a label), the description of this feature in the database, the entity in the GIS which will be linked with this ID to the database through an attribute join.

The methodology used during the field investigation is shown in figure 2. This chart is composed of five parts showing the five steps of the data recording process.

First of all, we have to identify and conceptualize what has to be recorded in the field. The 3D position of the artifacts and position of stratigraphic units can be represented by points. The shape of archaeological features or the extent of trial trenches or boundaries of the mechanic overburden removal are drawn with polygons. Finally, the axis of cross section is represented as a line.

The topographer is in charge of recording the features in the field. His job is to rapidly record, all day long and accurately the features and allocate an ID to each feature. In the office, he will verify the accuracy of the recording, control errors such as duplicates or missing information and he will produce plans on a weekly basis. Two to three topographers have been employed during 288 days, recording 80 000 different features

Centimetre accuracy in the three dimensions and a control of this precision were required for the topographical equipment; for these reasons, we used a total station instead of a GPS device.

On the other side, it has to be a robotic device (Leica TS 15) which allows to record natively all the features enumerated above as point, line or polygon. The standard total station Leica TS06 used at INRAP can not do the same: a codification has to be used in Geomedia COVADIS, a plugin software in Autodesk AutoCAD which will draw line and polygon from the point cloud recorded in the field. It is not the only benefit of the use of the robotic total station in comparison with the standard equipment; we could cite the autonomy of the topographer during the work, real-time topographic computing, the visualisation of maps in the field, and the possibility to record different descriptive data linked to the graphic entity. As a matter of fact, we were able to record for each archaeological feature their ID, the stratigraphic unit to which it belongs, the nature of the artefact, and additional comments. A data dictionary was designed in Leica Geo Office for this purpose and upload in the Leica CS15 and its software Leica SmartWorx Viva.

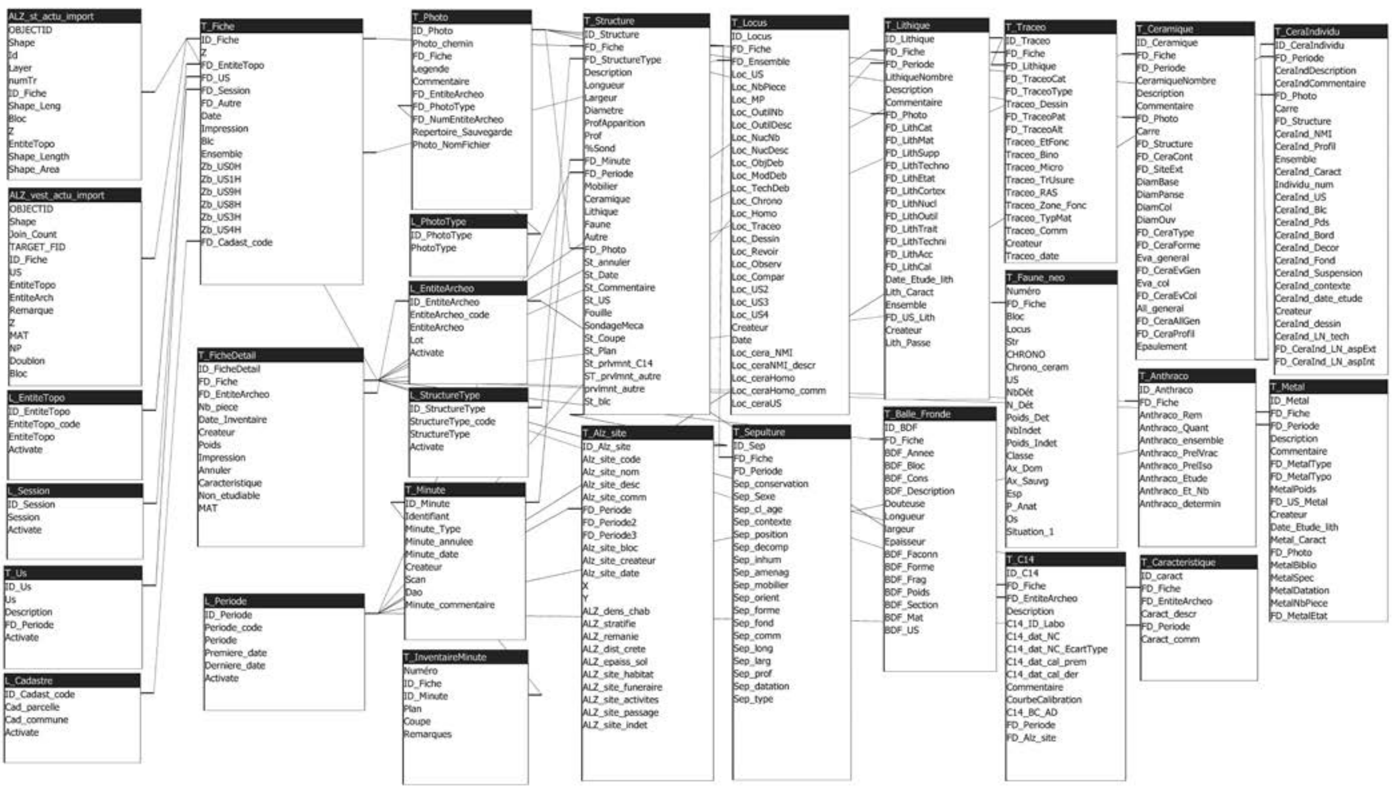


Figure 3: Simplified relation table of Alizay database.

The keypoint of the use of the robotic total station is the normalization of attribute data directly in the field which then allows a GIS export of topographical data as an ESRI shapefile. This format which could be read by Autodesk AutoCAD, Leica Geo Office and of course Esri ArcGIS provides a good interoperability between software used in this process.

The use of two different total stations, a robotic one and a standard one, prompted two different processes as seen in figure 2. The one used with the robotic station is shorter than for the standard station. The latter requiring more steps to produce the same quality of topographical data, which is then used for the GIS update.

Just before this update, the dbf file from the shapefile containing topographical data, is used to check duplication or incoherence in the data set. 'Alz Verif Topo' is a VBA routine developed in an Excel file by P. Boulinguiez, one of the topographer of the operation, which reads one or several dbf files and check duplicates of feature's ID, errors in the z value regarding a fixed range of heights, jumps in the ID numbering or missing information (stratigraphic units or nature of artifacts). A report is then produced, summarizing the different problems which can be displayed or printed in order to proceed with modifications to the data set.

In parallel with the recording of the graphic entity, the archaeological recording has to be done in both the field and in the office. A tool for the whole team is necessary, with the possibility of describing single item or a batch of items and to store interdisciplinary data. The 'Alizay' relational database has been designed with Microsoft Access, ensuring a good interoperability with a geographical information system and a light weight storage file. The database of the project weights 60 MB whereas a zipped version represents only 13 MB: sharing of data between members of the project is facilitated even by e-mail.

Most of the time during the operation, archaeologists were working at the same time on a network database which was stored on a network area storage (NAS): the aim was to avoid redundancy of information. The database was used to inventory artefacts and produce labels for archiving but also for specialists to store studies about

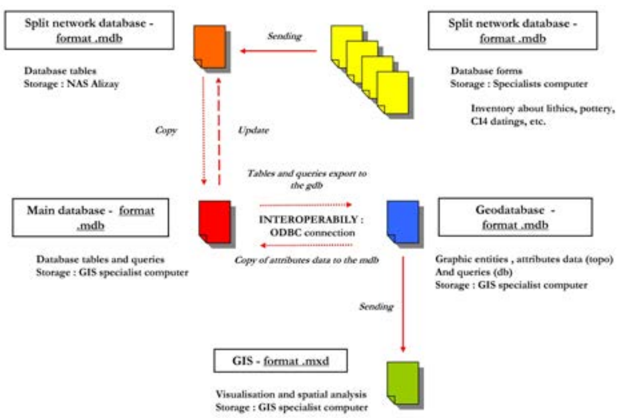


Figure 4: Architecture of Alizay process.

lithics, pottery, charcoals, anthropology, C14 dating, etc. The archaeological features were registered directly in the field through a lighter version of the form designed for a tablet PC.

Figure 3 is a simplified version of the relation's table of Alizay database where the primary key of most of tables is the unique ID used to define archaeological features. New data created from the survey on a monthly basis, is loaded in a generic feature layer ('ALZ_artefact_update') from the vector geodatabase of the project and linked permanently to the Alizay database through an attached table. This is possible because the Esri geodatabase is a mdb. file and appears as a Microsoft database file. It allows to harvest topographical data concerning ID, type of artefact, geographic coordinates (X, Y and Z), stratigraphic unit, cadastral parcel ID, remarks, etc. These data are stored in a specific table called 'T_Fiche' table, the main table of the database. Data derived from artifacts inventory ('T_FicheDetail'), from specialist studies ('T_Lithique', 'T_Ceramique', 'T_Anthraco', 'T_BalleDeFronde', 'T_C14' etc) or from archaeological feature analysis (T_structure) are linked to the main table. Photos or scanned and vectorized drawings can be displayed into forms related to artefacts or archaeological features: a relative path is stored in a specific table but not the picture itself otherwise it will increase the size of the database.

Basic queries have been designed in order to be exported to the GIS. The aim is to present the distribution of the different type of artifacts ('Qry_L' for lithic, 'Qry_C' for pottery, 'Qry_F' for faunal remains, 'Qry_M' for metallic finds, etc.), the density of artifacts or features according to the stratigraphic unit or the chronology ('Qry_St_NEA' for the structures dated from Early Neolithic, 'Qry_St_NEM for the structures dated from Middle Neolithic, etc.)

The GIS of the project is composed of a vector geodatabase and a raster geodatabase, in order to improve the management of data. A personal geodatabase has been set up to store rasters such as topographical and geological maps, Lidar or aerial imagery, DEM produced, georeferenced photos, old maps (Napoleonic cadastre, atlas of Trudaine, etc.).

The vector geodatabase stores all the graphic entities and their attribute data. As explained above, two specific feature classes were created to import new data derived from the topographic survey: one for the update of artifacts (point feature class) and one for archaeological features (polygon feature class). These two feature classes will be added on a monthly basis to the main feature classes containing the whole data set. In case of errors, duplicates or problem of versioning, entities are stored in two specific feature classes according to their nature (point or polygon).

A lot of different feature classes concerning archaeological evaluation, cadastral parcels, extent of the mechanic overburden removal, etc. are also stored in this geodatabase.

Joining attributes table to a large set of graphic entities in Arcmap (>100 000) can be very difficult in terms of display and labels. A direct join with a table stored in a database through an ODBC (Open Database Connectivity) link will crash but if the table is stored with graphic entities in the same geodatabase, then displays and labels will refresh as fast as with a small dataset.

Therefore, an ODBC link created in Esri ArcCatalog with 'Alizay' database allows importing queries designed earlier in Microsoft Access. This import is done automatically through a routine created in the modelbuilder of ArcGIS. Graphic entities and attributes data coming from the database are stored in the same geodatabase and then can be joined and displayed quickly in Arcmap.

As explain above, the whole process set up for Alizay operation requires a specific architecture to deal with a large dataset that has to be shared by different actors (specialists, topographer, GIS operator, etc.). Figure 4 shows how such a process can work with a network database and a GIS. The cornerstone of the process is the storage of the Alizay database and the geodatabase on the GIS operator's personal computer. Copies and updates are made from this database to and from the shared database on the network and to the geodatabase of the GIS. The backup of the most important data of the project can be done from the GIS operator's PC: Cobian Backup from Cobiansoft was employed to copy the Alizay database and

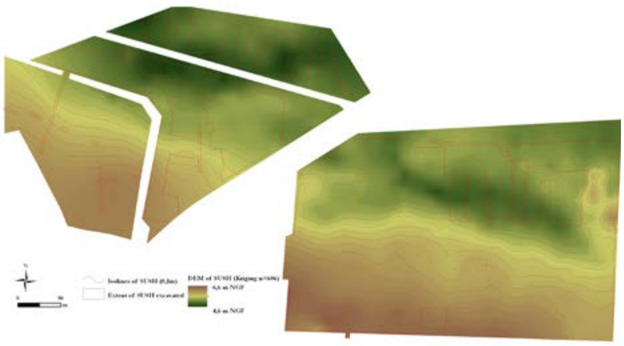


Figure 5: Modelization of a stratigraphic context (8H) of Alizay (Interpolation by kriging).

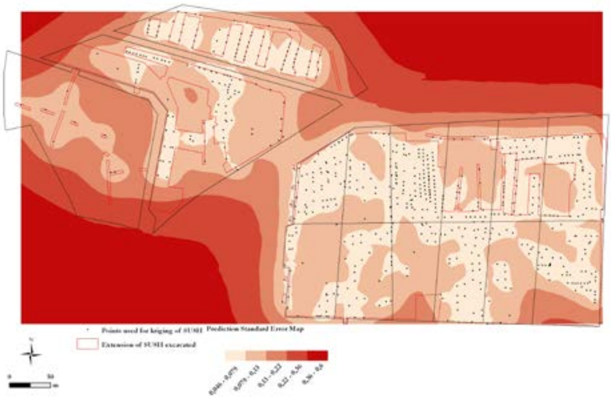


Figure 6: Standard prediction error map of the DEM (figure 5)

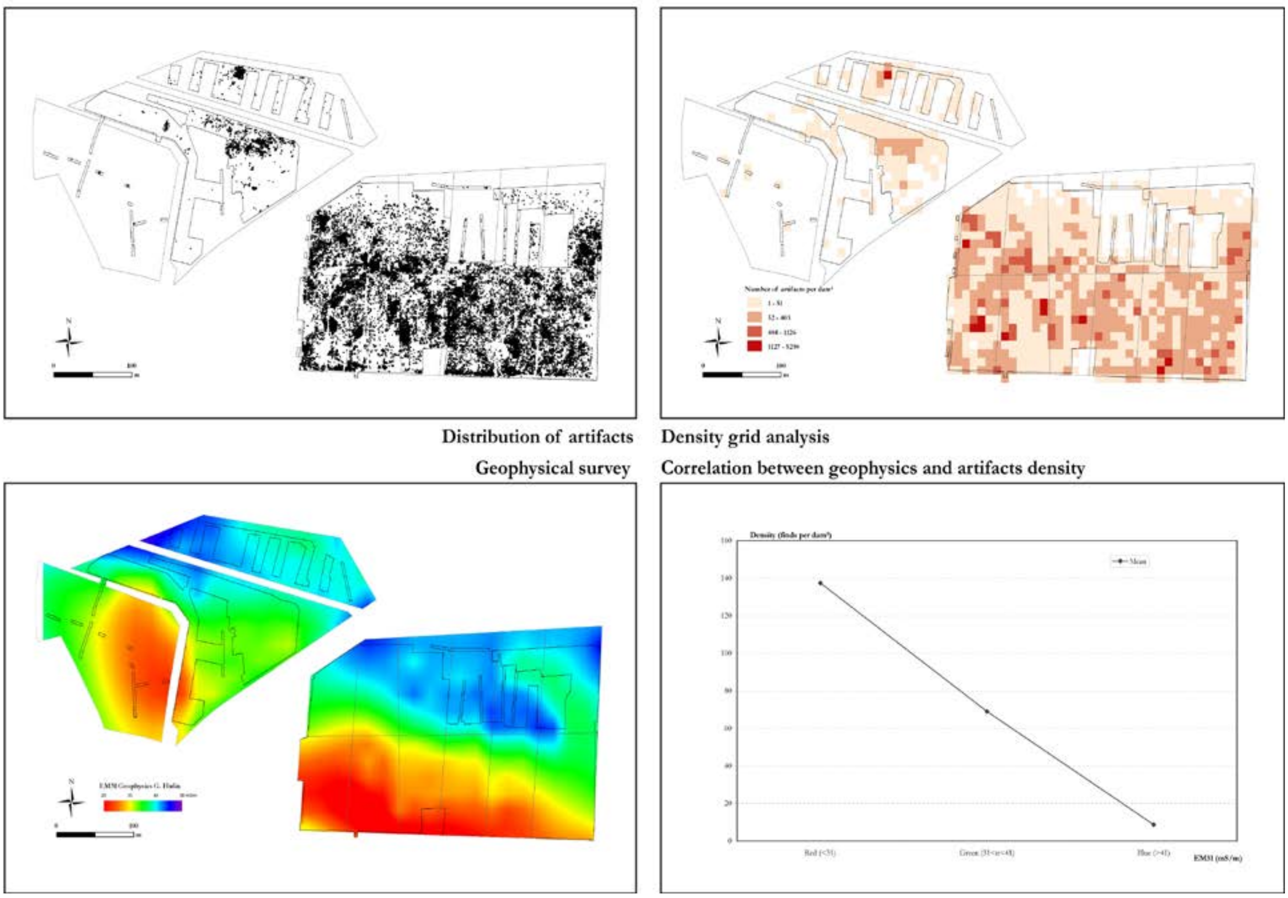


Figure 7: Alizay data recording and analyzing process.

the two geodatabases everyday on different support: NAS, hard disk drives and in the cloud.

The last section of figure 2 concerns applications of the whole process. For the topography, as explained earlier, the aim is to be able to produce maps showing the extent of the overburden removal, archaeological features and artifacts, on a weekly basis. Surface and volume calculation allowed accurate monitoring of the progress of the excavation.

The main interest of the database is to archive data in a safe and sustainable manner. When queries are set up, it provides up to date artifacts or features inventories which have to be present in the official report of the archaeological operation. It also provides statistics for specialist's studies. Moreover, the management of iconographic documentation is ensured by storing, listing and naming pictures and drawings.

Several GIS applications have been done during the fieldwork. One of the most useful was the production of DEM of excavated stratigraphic units exhumed. 2507 points have been recorded for five different archaeological layers. An interpolation of these points by kriging was run in order to reconstruct the paleotopography (figure 5). At the same time, errors have been quantified and spatially located to validate the model (see figure 6). The model was then used as a basemap for the paleoenvironemental reconstruction using multi proxy studies.

Another application was the use of the distribution of artifacts to run a density grid analysis (10x 10 m). This rasterization of artifacts distributions was then compared to the resistivity maps produced by the geophysical survey. The comparison revealed a lower density of finds in the clayish zone which corresponds to palaeochannels. At the opposite, higher density of finds are located where gravels are present, on ancients river banks (see figure 7).

We use also the GIS to apprehend the stratigraphic context through the display of artifacts point cloud in Esri Arcscene. We also produced vertical distributions of artifacts using a plugin called Crossview, a A-Prime Software: it projects finds on a cross-section created by data derived from DEMs. It appears to be a very good tool to understand successive human occupation located on a same spot.

4. Conclusion

As a conclusion, the deep stratigraphy, the presence of prehistoric settlements with significant number of finds, the time constraints and the scale of the project required a complex process to ensure the detailed recording of a significant amount of archaeological data.

This process made possible the excavation and the study of several occupations and sites like an Upper Palaeolithic settlement (9800-9300 cal BC) associated to long flint blades production and the butchering of six aurochs (Bos primigenus), several Middle Mesolithic sites (8th millennium) with geometric microliths (lunates), an Early Neolithic occupation represented by few pits containing a pottery of La Hoguette culture dated by C14 on bone temper (5370 – 5222 cal BC). One hundred hearths of the Middle Neolithic with associated flint and ceramic concentrations have been excavated; the river bank of the Seine appears at this time as an area of significant activity. A high concentration of artifacts and two series of postholes indicate the presence of two different habitations, close to each other and dated to the Recent/Late Neolithic. Several loci dated between 2400 and 2050 BC were found in the south-western part of the excavation. These consist of flint and pottery, some of them attributed to Bell beaker culture. During the Early Bronze Age II, geomorphology shows that the floodplain is well established and fords are appearing: near two of them, more than 250 sling stones have been discovered indicating a possible fight between Bronze Age populations. Two circular houses and several quadrangular structures (granaries?) associated with metallurgy indices (slag and melting-pot), revealed the presence of a Second Iron Age occupation. Finally, 15 m wide ditches with a depth of 4 to 5 meters have been identified in the north-west of the excavated area and belong to a quadrangular fortification. This feature is related to the fortified bridge of Pont de l'Arche (Eure), built between 862 and 873 AD by Charles the Bald, to stop the Viking raids. Ovens and a funerary group of 11 individual graves dated by C14, to the 9th century AD are completing this Early Middle Age occupation, whereas metallic finds (coins, weaponry, and clothing) show the use of the area until at least the 16th century AD.

The archaeological investigation carried out by INRAP at Alizay has widened the scope of the prehistoric and historic knowledge of Normandy while it improved methods in preventive archaeology by developing an original data recording process, which could be applied to other big scale archaeological projects.

References

AUBRY B., TOMANN A., BIARD M., BEMILLI C., SANTIAGO-LARA V., TESSIER V., PETIT P., HONORÉ D., 2011, *Une occupation du Tardiglaciaire : Alizay-Igoville (Eure)*, Journées archéologiques de Haute-Normandie, Harfleur 2010, Publication des Universités de Rouen et du Havre, pp. 9-24.

MARCIGNY C., AUBRY B., MAZET S., 2013, *Au bord de l'eau ! Les fouilles du Port-au-Chanvre à Alizay et Igoville (Eure), présentation liminaire : méthodes, attendus, premiers résultats*, Journées archéologiques de Haute-Normandie, Rouen 2013, Publication des Universités de Rouen et du Havre, pp. 33-46.

Records and Spatial Representations in the Context of a Rescue Excavation: The Case of Quincieux (Rhône-Alpes, France)

Ellebore SEGAIN, Veronique VACHON, Bernard MOULIN, Cécile RAMPONI, Wojciech WIDLAK

Rhône, Institut National de Recherches en Archéologie Préventive (INRAP), Lyon, France,
ellebore.segain@inrap.fr, veronique.vachon@inrap.fr, bertrand.moulin@inrap.fr,
cecile.ramponi@inrap.fr, wojciech.widlak@inrap.fr

Abstract

In the particular framework of rescue archaeology in France, our aim is to identify and analyse the archaeological information from sites concerned by development projects. In the rescue problematic, cost and time are important brakes to the collection of the data.
In this context, our team developed a process which allows to record and study archaeological information (spatial and descriptive), and help managing the excavation.
This protocol has been experimented in the excavation of Quincieux 'Grange Rouge' (Rhône-Alpes, France). This approach is based on collecting and archiving information into a database (DBMS), using PC-tablet on the field, in association with a GIS implemented with the latest archaeological survey. The aim is to optimize the collection's and investigation's time.The result is a coherent system that helps the archaeologist to deepen his management and investigation, from the field to the exploitation of the data and the publication of the results.

Keywords: Data standardization, Intra-site GIS, PC Tablet, Field management, Rescue archaeology

1. Introduction

1.1. The French system

In France, in the case of important development projects (like motorways, buildings…) the state's representative service can demand an archaeological evaluation.

The goal of evaluation is to detect, characterise, plot and date archaeological deposits.

Only after the edition of the evaluation report, the state's representative service can order an excavation.

The developer must choose the archaeological contractor for the excavation: it can be INRAP (Institut National de Recherches Archéologiques Préventives), or any other accredited body.

Each year, more than 2000 evaluations and around 300 excavations are conducted by INRAP.

Evaluation and excavation must follow the specifications of the French state related to the heritage code.This code establishes standards for packaging and inventorying all the documentation gathered and all the archaeological material found.

Archives are the field original documents like excavation notebooks, index of facts, maps, drawings, pictures and so on … All these documents are numbered and labelled to include a certain amount of information, and records in an accurate inventory too, including the container where they will be stored. The storage warehouses depend on the State representative service.

Collected material is distinguished by: archaeological material, which is called 'artefacts'; and natural and biological materials, which are called 'ecofacts'. It must be washed, sorted and packaged by sort of material. It must be record in specifics inventories, according to the accurate demand of the French state, including exact administrative localization (for the owner of the site).

In this regulatory context, the evaluation and the rescue excavation are also constrained by three parameters.

First, the space is strictly bounded by the development's project.

Second, the time is defined with the developer according season, location, and more often developer's obligation about the project.

Third, the price of the excavation is an important parameter for the choice of one contractor rather than another.

These three parameters are omnipresent at any step of the archaeological work. They induce specific methodological procedures (including mechanization) for field and post-excavation stages. It explains the choice of one process rather than another, and the search of the most efficient process for not loosing time.

1.2. Usual way of doing

Until now, data was mostly recorded on the field with paper, then copied in a spreadsheet during post-excavation step.

Spatial information was surveyed and drawn quickly by a surveyor, and given to the staff. The complete map was drawn only in post-excavation step, using the 'manual draws' made by archaeologist and fitted to the surveyor's map. The final map was designed with Adobe Illustrator©.

During the excavation, only sketches were presents on the field, or some impressions on paper, with the difficulties of reading the information on a large excavation.

In this system, the spatial information was completely disconnected from the descriptive information.

1.3. Recent development

Integration of GIS and databases is a recent development in INRAP coming with the rise of mechanization in rescue archaeology, i.e. in a business environment.

The increasing activity in preventive archaeology led to the rise of the production of archaeological information. This dynamics required the use of new methods, as developed by our Scientific and Technical Direction (DST) in the 2010 CAA conference by X. Rodier (Rodier et al. 2010).

Moreover, the need of standardization and the search of methods for saving time increases with the constraints always more present in the every day life of the rescue archaeologist. In this specific context, the use of GIS and DBMS for research and management is generalized from intra-site level to regional level. The rise of open source software like Quantum GIS really helps.

Different experimentations take place in the INRAP Company, all around France. One, concerning the town of Noyon and the canal Seine Nord Europe (Oise, France), was yet published (Bolo et al. 2014; and Bolo et al. 2010).

This article aims to explain one other experimentation, carried out by the site of Quincieux 'Grange Rouge' (Rhône, France). This excavation was foreseen for testing a new type of process, which was thought but not experimented yet during an excavation.

1.4. Quincieux experimentation

Before the growth of the motorway 'A46' made by the APRR Company, evaluation detected settlement on the 'Grange Rouge' area, near the town of Quincieux, located at twenty kilometres north of Lyon, near the Saone river (Fig.1).

The high points of experimenting on this site are the following: different human settlements are expected (from Neolithic to Modern age), on a very large area, with an important density of remains. The process was planed before the beginning, with a motivated staff. Furthermore, the evaluation of the site had been made with a similar protocol.

The evaluation was made nearly a year before, with a database linked to a GIS. The evaluation concerned around 276 000 m². During 2.5 months, 381 test trenches were made, revealing 237 remains, concerning 8 large periods, from Neolithic to Modern Age.

After this step, a report was made, and the French's representative service of archaeology demands an excavation on the largest place were remains were found, which concerned 91 000 m².

The excavation (field work) took place from February to September 2013, with a staff formed by one excavation manager, two section managers, and a team which goes up to 45 archaeologists. The post-excavation work is still in progress.

After six month of excavation, 1 800 remains were discovered, numbered, tested, drawn, photographed, surveyed and recorded. Furthermore, there were 720 isolated artefacts (mostly flints). The remains could refer to a rural settlement and a necropolis. The periods concerned go through the Neolithic to the Modern Age.

The excavation of these remains involves the production of important number of documentation, and the discovery of a lot of archaeological materials, as we can see in Fig. 2.

At this scale and in this chronological context, the use of GIS and DBMS looked essential, but needed to be

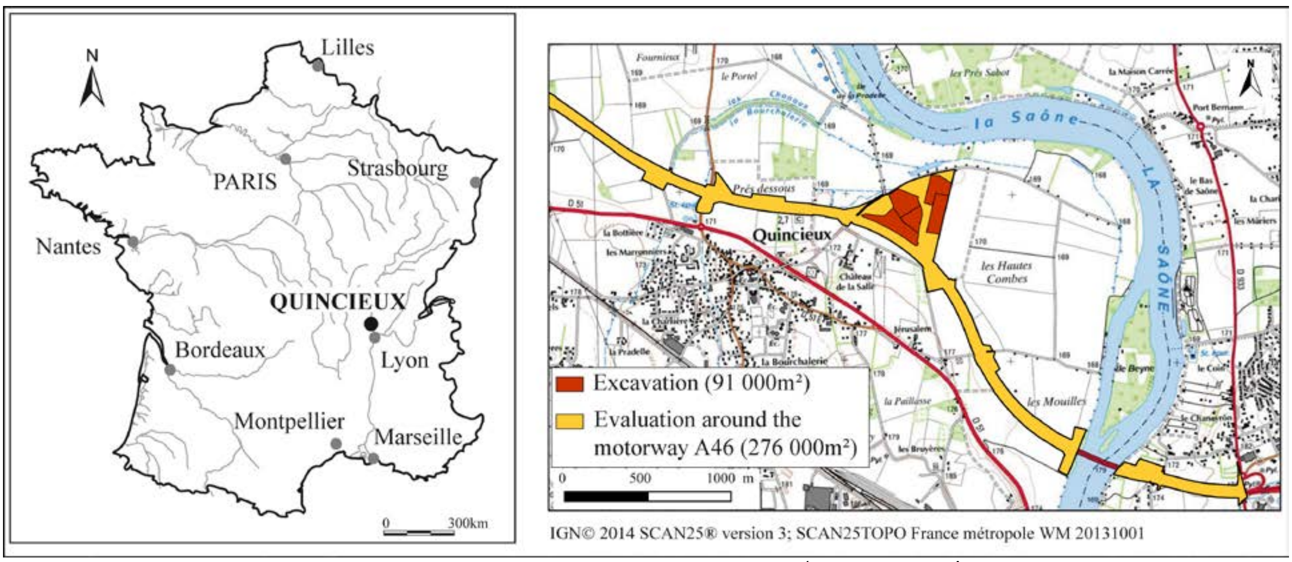


Figure 1: Localization of the site of Quincieux 'Grange Rouge'.

help for documenting	
number of nails	5460
documentation	
number of field draws	860
number of pictures	11230
survey and GIS documentation	
number of topographic surveys	65
total number of points in all topographic survey	23000
number of shape files	560
archaeological material	
number of bags of artifact (without paleobotany samples)	3770 (830kg)

Figure 2: Archaeological information produced by the Quincieux excavation.

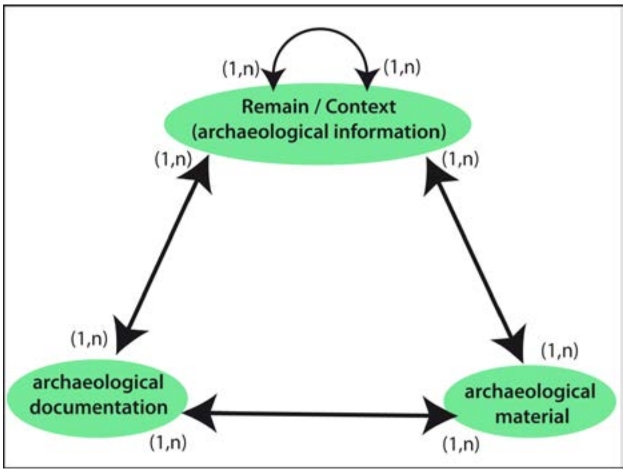


Figure 3: Basis of the system: the triangle of interrelation, as described by B. Desachy (Desachy 2008, p. 153).

well planned before and well organized to be efficient. The methods are described here, then the results of this experimentation, inducing discussions.

2. Methods

The process created for the excavation of Quincieux is structured around a device combining a GIS and a database dedicated to the recording and the management of archaeological information, both administrative and scientific information.
It was elaborated before the beginning of the excavation.

2.1. Everything is linked

They are relations and links between all the sorts of information given by an excavation.

First, we consider, as B. Desachy does in his thesis (Desachy 2008, vol 1 p. 153) that there is a triangle of interrelation between archaeological material, remain and documentation (Fig. 3).

They all can be inserted on a database as data, and they are always connected, and needed to be linked so as to allow the analysis work. The architecture of the database is based on this triangle.

Second, archaeological information is both descriptive and spatial information (Rodier et al. 2011). Archaeologist must deal with both for understanding his site, and he need to join them in a unique system.

Third, archaeological data are some spatial data with different scales of signification, and archaeological argumentation is based on space, especially in terms of proximity, connectivity, or relationship with geographical distribution (Rodier *et al.* 2011).

2.2. Spatial Dimension

The origin of the spatial information is the surveyor, who passed the information collected to the GIS operator who manages the spatial data.

The numbering of single entity is made by the archaeologist in order to eliminate duplicates for a successful integration into the database. The result is a codified survey.

This numbering allows a better differentiation of the points on the draw (Fig. 4): the way of 'numbering' the nails clearly distinguish them from the points dedicated to the remains, or the points dedicated to the isolated artefact. Nails have the letter 'C'(for 'clou', meaning nail in French) in front of the number, isolated artefact an 'I', and the facts no letter. This system allows multiplication of numbers without confusion.

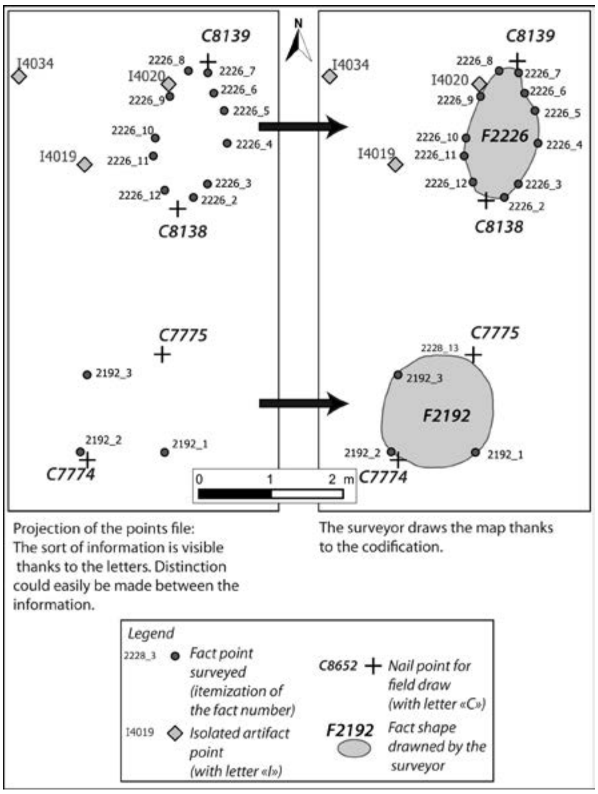


Figure 4: The numbering of the points drawn by the surveyor.

The surveyor also uses a codification to export different sorts of information about the object and its description. They are two sorts of code:

– Attributes describing the characteristics of identified elements (object class)
– Data showing the graphical representation of the object (closed graphic shapes, circle, ellipse...).

The information is described during the field work thanks to the codes that can be combined together, and allow the design of maps, and the export of files for GIS software.

2.3. Descriptive dimension

A unique ID is the heart of the archaeological recording system witch doesn't tolerate double. Each element, whether it's a fact, a stratigraphic unit, an isolated artifact, has a single ID, unique at the scale of the excavation (Fig. 5).

Archaeological information is record thanks to a part of the database based on a nested hierarchy where the 'Stratigraphic Unit' is the smallest entity, as you can see in the Figure 6.

This organization underlies the entire field approach, and the way of recording archaeological information.

The recording system allows to input any information support (pictures, drawings, comments and others), and keeps it related to the remains (see Fig. 5). It also uses a nested hierarchy (Fig. 7) for the fitting of the documentation in containers so as to store and archive them.

All official information needed and asked by the French state are included.

The database contains a material module: all the material can be input into the system, as data (see Fig. 5). It allows dating some structure, or providing information that would help the analysis of the remains, or managing the post-excavation treatment work.

The database is structured to harmonize and optimize the process and the restitution of archaeological information

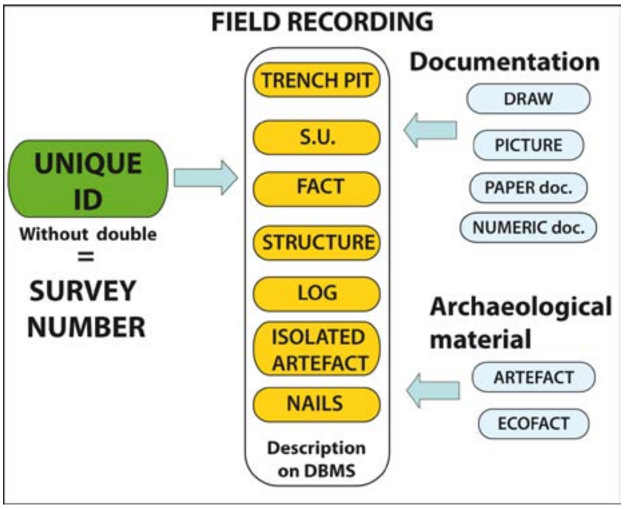


FIGURE 5: THE DATABASE MODEL.

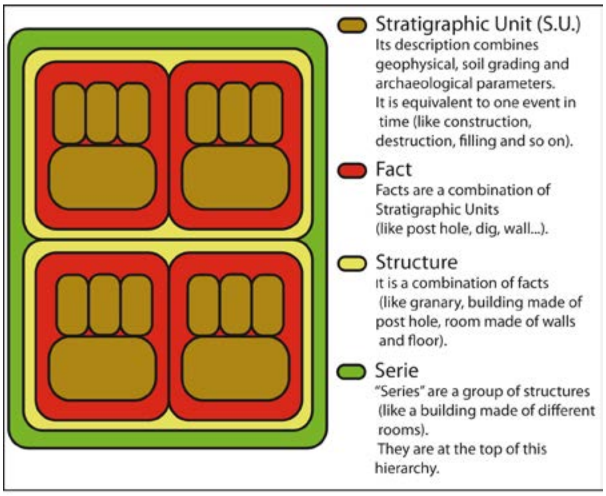


FIGURE 6: THE NESTED HIERARCHY OF THE ARCHAEOLOGICAL ENTITIES.

from the field to the legal inventories, by the standardization and the harmonisation of the information. It was thought also for enable thematic and chronological analysis later, keeping links between entities, documentations and artifacts.

2.4. Tools

The software and specific tools available for this experimentation are the followings:

– 3 PC Tablet 'Motion Computing F 5'(without wifi system)
– 'Access 2003 ©' for the DBMS

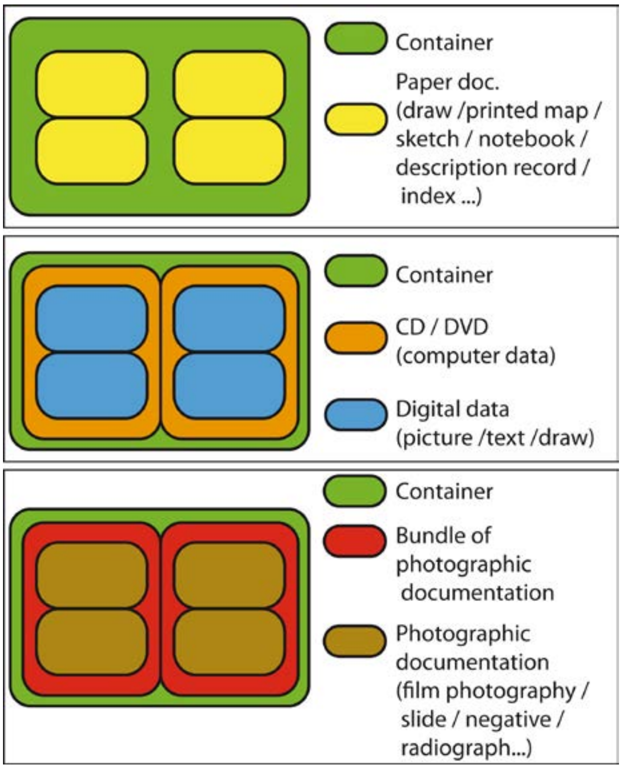


FIGURE 7: THE NESTED HIERARCHY OF THE ARCHAEOLOGICAL DOCUMENTATION.

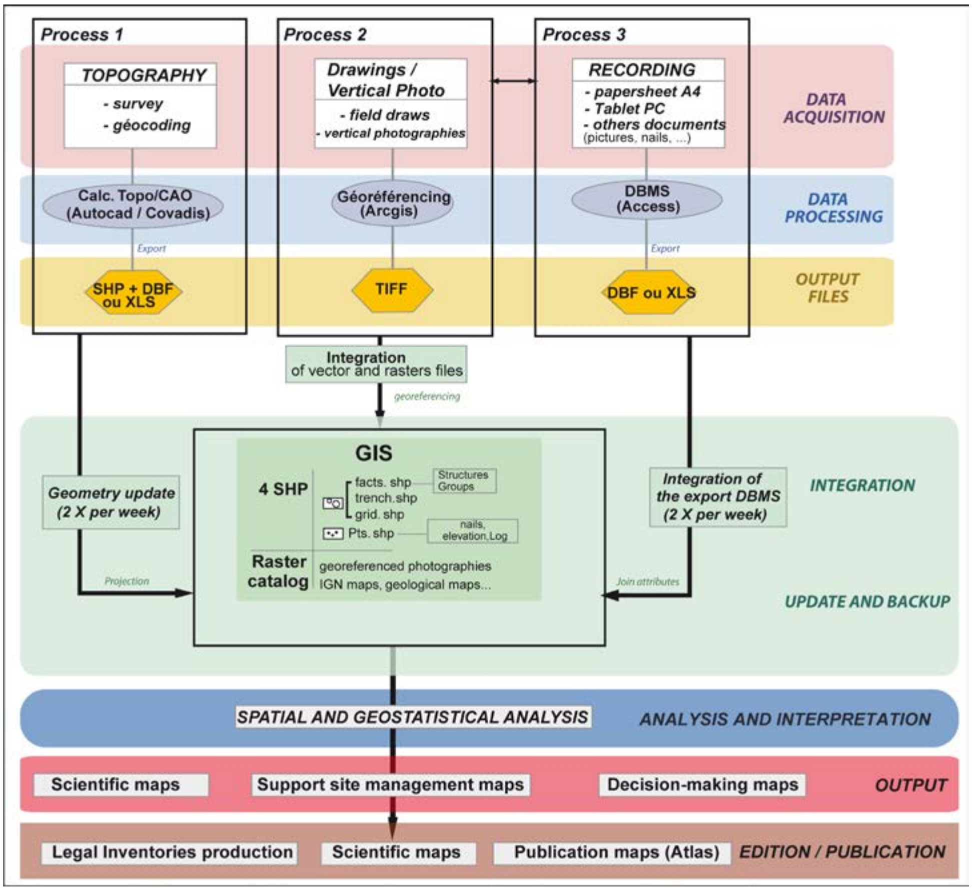


Figure 8: The process used during the excavation of Quincieux.

– ‘Excel 2003 ©’ for the export and import between DBMS and GIS (on the field)
– Covadis 13 © and AutoCad Map 3D © for the topographic survey
– ‘Leica TS06’ total station.
– ‘Quantum GIS’ (open source software), mainly used for consultation and ‘every days treatments’.
– ‘ArcGis 10 ©’ (1 license) occasionally used, for specifics treatments.

For the excavation, to record the information, an Access database is created based on the principles described above. This database is filled and implemented on two PC Tablets, in parallel with a record on paper. One more PC-Tablet was used for zenith picture.

Database was thought so as to be as ergonomic as possible for PC-tablet users (smaller screen; drop-down list so as to reduce input mistakes; …) and to avoid input repetition.

3. Results

As a result of this reflection made before the excavation, a process was created and used during the field work, considering the tools available and the constraints of rescue archaeology (space, time and cost).

3.1. The process

In figure 8, we can see the complete process, including the relation between the different actors. It would be developed and commented below. The process is made according to the F. Djindjian definition (Djindjian 2010).

The first process is about topography. The surveyor carries out the survey with a total station. During the topographic survey, the archaeologist works with his field book, saying remains numbers to the surveyor who realizes the geocoding.

At the office, the surveyor enters the field books calculated, using AutoCad-Covadis © and draws polygons and lines, automatically, thanks to the geocoding.

After each survey, the surveyor provides the shape and point files to the GIS operator and the management staff, in a very short time (Fig. 9).

On the field, the surveyor is present two days a week, so as to cover the 91 000 m² of the excavation. During scrapping time, surveys were given 24 hours after, and during testing time, 48 hours later.

At the end 23 000 different points are surveyed and directly linked to facts, isolated artefact or nails.

If the surveyor’s job is at the base of all the system, GIS is at the center.

The second process is about GIS. The GIS operator cleans and merges the latest shape files and the older ones, after cleaning the draw and checking the points and polygons. He also deals with the vertical photography of specific remains

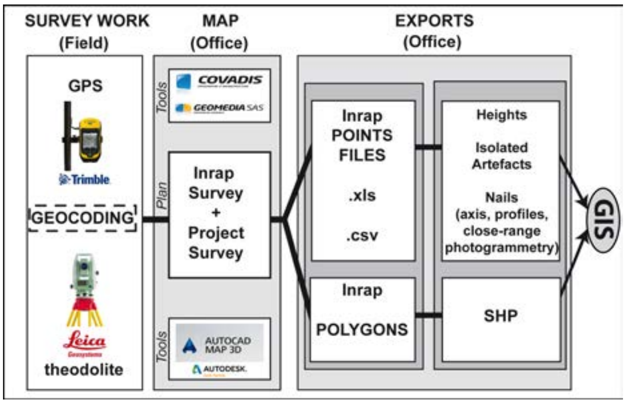


Figure 9: The topographic process.

like wall. He corrects the pictures and inserts them into the GIS so as to draw precisely the stones for the final edition.

He geo-references old maps like Napoleonic cadastral plan, and edits plans that allow preparation sketch for the next survey, or some documentation for public relation.

The GIS operator is the unique deposit of the back-up of the GIS and the only one that can change shape files. His work involves a permanent communication and adaptation with archaeologists and their demand.

If changes are necessary, archaeologists identify them thanks to the files send by the surveyor, publish a report, and send it to the GIS operator.

In all cases, it's the archaeologist who coordinates the process and validates all the plans.

The GIS operator works two days a week for the excavation too, near the archaeologists with which it interacts constantly.

At the end of the field work, near 65 surveys were concatenated and fixed. Spatial information were digitized, checked and cleaned at 99%. In total, it's nearly 560 working shapes files that were created, with 253 points files, 5 lines files, and 302 polygons files. The post excavation step is still generating shapes files.

The third process is about the descriptive dimension. The database was filled every day, by different members of the team, and not only the managers, according to the progress of the work. They operated often by group of two, one for describing, and another for writing, entering data for various remains and documentation. At any step, information about the team work (what is done) and the archaeological analysis (what is it) is put into the database. This moment allows, besides physical rest, reflection and comparison between remains.

The excavation area was divided into five sections of different sizes and there were only two PC-Tablets for the databases and no WiFi system. The two PC-tablets are in charge of two sections each one. The last section only worked with paper (with the same column than the database).

section	information digitized at the end of the field	detail
1	20%	paper
2	99%	
3	65%	started with paper, then PC Tablet
4	99%	
5	90%	

Figure 10: Archaeological information digitized at the end of the field step.

At the end of the excavation, most descriptive information is yet digitized (Fig. 10), but need to be checked and enriched.

During field work, each member of the staff and the GIS operator can link the new and complete shape file with the database of one (or more) section so as to use it as a geodatabase. More often, he links Excel files exported from the databases to the GIS.

This process allows creation of different sort of maps (see Fig. 8): management map (for example: remains done, remains to do) to help managing the excavation, or scientific map (interpreting the site).

At any step, GIS and database are still fundamental tools and their combination is essential for the progress of the reflection.

3.2. Evolution of the process

Very few evolutions appear during the field work.
The important quantity of data generated by the excavation on five sections induces a high rhythm of verification that cannot always be done during working time, but often during resting time (i.e. personal time).

The ease of use of the GIS and the high rhythm of testing the remains involves a demand of the managers about the GIS operator to send last managing plans more quickly, to fit always more with the daily work. At the end, these sorts of maps were made directly by the managers, and edit on the field.

Finally, information that, in the past, were available after the end of the field work, were now easily accessible, and generated a demand of almost instantaneousness.

4. Discussion

This system introduces interesting discussions in our staff and in our company.

The reflection about the benefits and issues of the GIS/DBMS system is maybe well known in Laboratories or Universities, but this process is a real innovation in rescue archaeology, because of the three constraints described in the introduction (space, time and cost) and the French heritage law. That's important to keep them in mind, reading the argumentation below.

4.1. Benefits

4.1.1. Management dimension

PC tablets: a real help for managing data
PC Tablets, used in a mechanization mindset, is a major archaeological evolution at least equivalent of the use of computers. With the help of the PC Tablets, the archaeologists are in direct interaction with the descriptive and the spatial information all along the field work.

Elements required for managing the operation on the field are available for consultation or updating at any time.

Office work and data analysis are simplified and more efficient because of the computerization and standardization of the data on the PC Tablet. Recurrent data records are optimized, meaning repetitive inputs are reduced.

Another benefit is that PC Tablet is easier to carry on the field than heavy paper documentation on a big site.

GIS and DBMS induce facilities of consultation and modification for everyone
At every moment, the site map was available, for every member of the team, and its consultation was fast and easy, no matter the knowledge of the GIS software.

The use of one single system provides at any moment for the all team an updated and interactive cartography. There is now a direct interaction between the archaeologist and the map, map that can represent all the data we need.
It's important to understand that, to run a big excavation like Quincieux was, a lot of people is involved in the same time, during the field of course, but also during the post-excavation step. All those people need to access the database or the map at the same time for updating.

Maps of typology of remains, or dating, can be made directly by section manager or specialist. Even if all the data wasn't yet input in the database, reflection could be started and represented.

The post-excavation work starts directly on the field
Imputing data in the database directly on the field is already a part of post-excavation work.

During all the field work, checks are made. But what is checked is not only the information written by the team on the field, but also the final data in the database, or the final map.

This step, appearing time-consuming, is in fact time-saving when you consider the post-excavation steps.

The process helps the field's management
Managing the team and the work was made before on paper plan. This system involved difficulties of reading information in a big area.

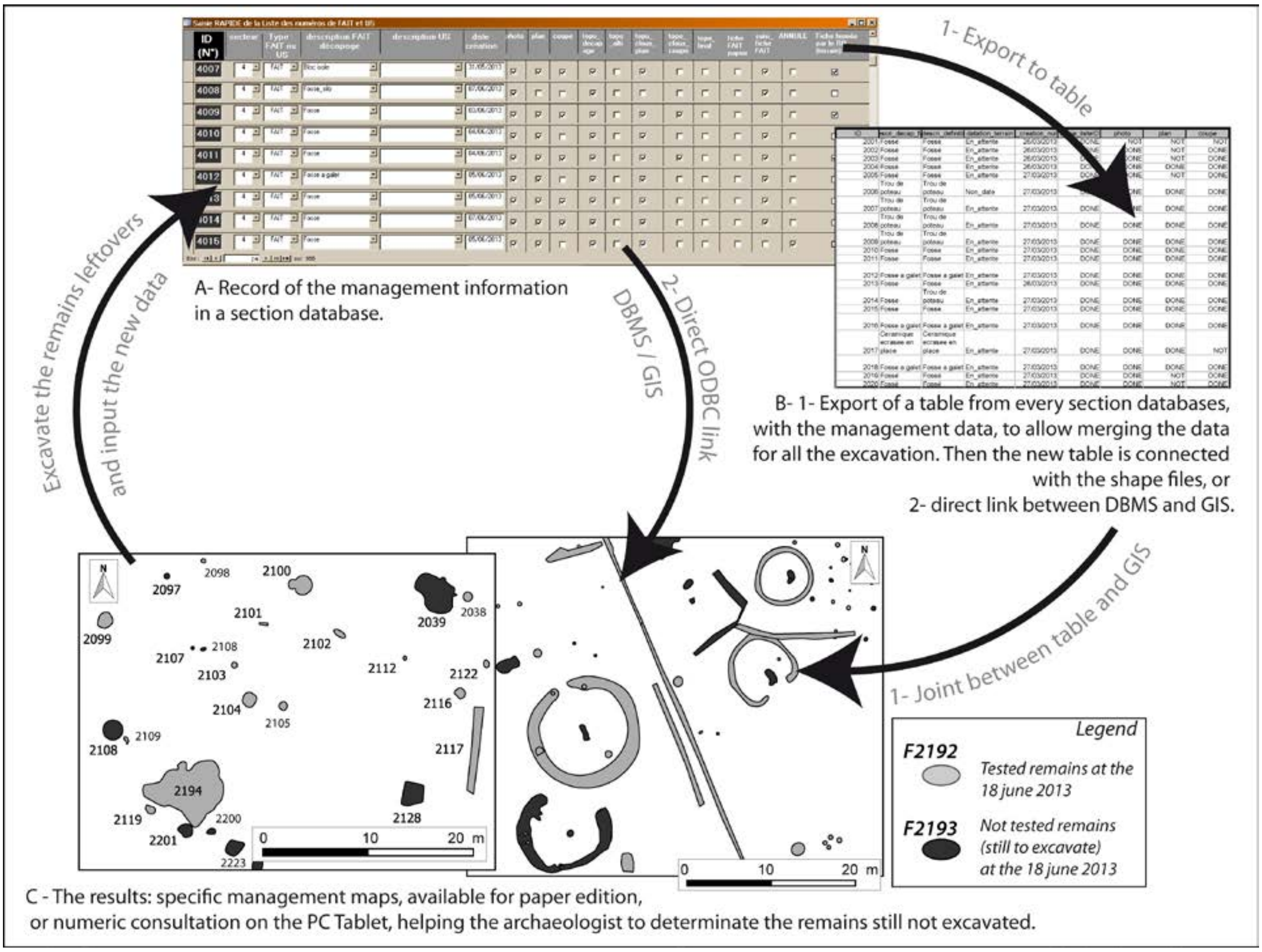


Figure 11: The process used during the excavation of Quincieux.

Thanks to the Quincieux process, at any step, information about team work is put into the database, and connected to the map. This interactive map, available on PC tablet, really helps managing the team, knowing what is already done exactly and what need to be done (see Fig. 11).

This process helps also to lead a scientific management, allowing identifying remains that can be treated differently, so as to help making excavation choices. For example, possible cremations were first excavated.

This allows management in real time of the excavation advancement, and optimization of the working time dedicated to the management.

4.1.2. Descriptive and Spatial dimension

The spatial and the descriptive dimension become easier to deal with, because of the GIS and the DBMS. Everything that induced in the past a lot of time for representing information and choosing what could be represented, can be made now quickly. The descriptive information can also be more efficient, as everything is linked.

The GIS helps the site analysis
Different analysis can be made thanks to the GIS, that couldn't be made before because of the time it required. Rescue archaeology is really constraint by this parameter.

For example, a work about the orientation of the graves helps us to confirm the existence of different steps in the medieval necropolis, even if the graves weren't already dated, and this existence of steps was only a hypothesis in the light of skeletons position.

Other example is the recovering of the Napoleonic map that helps identifying modern ditches.

The rise and the effectiveness of open source software like Quantum GIS is also a real helping factor without which we wouldn't have been able to do this experimentation.

The DBMS helps the study
As everything is linked into the DBMS, study of the site is easier to do, because the manager can find any information he/she needs. The manager can perform spatial queries or descriptive needed for the work at any time

The map edition can be made at any time
The use of GIS allows printing directly on the field a detail of an area, with the scale and information needed, for sketching.

It allows extracting, on demand, some complex map for different uses and goals: public relation conference, open-door days, site meeting with the contractor, coordination meeting with the staff, and so on.

Edition at any time is a real benefit for the management, but also for scientific uses.

An atlas system is used for a productive edition
For helping the edition of plan for the final report, or if needed, the maps produced are made following an atlas system. This allows harmonization of all the maps, no matter the subject.

More over, this allows production of standard maps, in a productive way for a big area.

Inventory management and production are made according to the heritage code
Like for the map, the edition of the data is possible at any moment, thanks to the database.

According to the 'French heritage code' mentioned before, all the lists (for artefacts, ecofacts, remains, pictures, and so on) can be extracted and produced with one click, and then inserted into the catalogue requested by the State, according to the print standards chosen by our company.

Archiving the spatial information is the permanent purpose
The files produced by our process have to be available for the future, for the public or for the scientists. The use of GIS shapes file in a standard version is a real progress instead of the Adobe Illustrator map made in the past, because they are less affected by the always more quick changes of the software version.

In addition, the entire process contributes to the harmonization and standardization of data across the site for archiving. Finally, storage is performed at more than 90% in a digital format which facilitates potential future researches.

The use of the process allows creation of recall for the archaeological information
Final backups are organized in two levels. First one are the inventories for the administration, at the regional scale. They follow the rules established by the French state. Seconds are for archaeologists, who can easily come back to the field data combined with the GIS, at local and regional scale.

For example, the same process, used during the evaluation step before Quincieux excavation, helped to estimate the potentiality of the site and allows to determinate the financial cost of the excavation. The estimation of the depth of topsoil and the numbers of remains expected, help to evaluate engine and team's time needed to excavate.

With this use of evaluation results, we can see that the result of one evaluation or excavation can help the field management of the next evaluation or excavation made in the surroundings areas. This process, once started, can easily be reused after the end of the operation.

4.2. Limits and Issues

PC Tablet constraints
As said before, only two PC-Tablets were available for the databases of the excavation, without WiFi system, and the last section only worked with paper.

This lack of PC Tablets came from their price. They are very efficient engine, waterproof, dustproof, and rugged, but they cost near 3 000 Euros each.

As the PC Tablet hadn't some WiFi system, we couldn't share one database on various PC-tablets, so we had to build five databases (one for each section) that we would join later. This system really increases the time dedicated to the maintenance work.

At the end of the field work, it represented nearly 800 remains yet to be digitized, that involved a very important work in post-excavation (and time). Fortunately, the other 1 000 were already done. With more PC tablet, we could imagine that all the data would have been digitized.

Another PC Tablet constraint is about the problem of visibility. The PC Tablet we used presents a hard to read screen during sunny days.

Rigour, time and big team

The system created here involved a lot of rigour, and this is difficult to require from a big team. During the excavation, the team reaches 45 archaeologists who worked on the site, and needed to be briefed at every step of the work. Then the specialists of different periods and material were added to this staff in post-excavation step.

Rigour involves a lot of work for the managers, so as to check everything, with evenness. It involves an important communication between the different members of the team, during a lot of meetings: daily meeting for the management staff, weekly then monthly meeting with the GIS Operator and surveyor, and regular meeting with all the staff, to explain and adjust the process. This organisation really needs to be well planned, to avoid mistakes, and induces a strong commitment of the staff.

The system clearly requires dedicated time for operating, a precious constant in preventive archaeology.

Reflection about thesaurus: a real need for national status

The use of only one database for a site, or a county, no matter what specialty is involved, should be a serious help for the archaeologist, but involves a huge reflection time, not only at regional level but also at national level, for the definition of the thesaurus and the computing field.

For the moment, this is not yet done, and harmonization is still hard to achieve.

Backup strategy

Working with GIS and database finally raises the issue of preservation and storage conditions for numerical data. Until today, legal State's saving data strategy in France is paper strategy, with the logic of 'if the numeric cannot be preserved, paper will remain'.

But in our case, the numerical data became too huge, mainly because of the size of the excavation, to be preserved only with paper. It concerns report, but also pictures, maps and inventories.

Formation

Using GIS involves working with people trained and opened to manipulation devices and new software. Training and testing time are necessaries. Currently, we are facing a generational problem inherent in the evolution of archaeological practices. An important work of monitoring and supports is necessariy to bring people and project to a new level, as explained by X. Rodier in the CAA Granada conference in 2010 (Rodier X. et al, 2010)

This is also the result of a more global evolution of archaeological practices and its perception in university formation in France. The learning about these fundamentals tools at this scale of practice is still lacking in France whereas the practice is going in the way of industrialization.

The problem of a national strategy

Other systems exist in France (Costa 2012), in other public company dependent on town or county, and they are all different and still in progress. They all work at town or regional scale. The first function of all these systems is their use as an archaeological map, i.e. a contextual map, for locating all archaeological sites in city or regional scale, but not in site scale. This scale occasionally comes after, like in the ALyAS project in Lyon (Hofmann et al. 2012) or the ArSol project in Tours (Husi et al. 2011).

In the case of INRAP, different systems and databases were developed, depending on regions and formations. They reproduced the situation described above. They were underlined in articles of members of the Scientific and Technical Direction of INRAP in 2012 (Ciezar 2012; Koehler et al. 2012).

Now, even if the use of these tools becomes easier and more common, national strategy (which is normally supported by the French representative service) is still being strengthened. Furthermore, this strategy is hampered by diversity of databases, traducing diversity of practices.

5. Conclusion

To conclude, various reflections led to the establishment of a process which allowed to record and study most archaeological information (spatial and descriptive), and helped managing the rescue excavation.

The process is structured around a device combining a GIS and a database.

It optimizes the collection and investigation time. Such a management of the collected data provides new opportunities to study in depth archaeological information.

Reflection about the architecture of the database represents a real scientific problematic, as the reflection about 'what we have to record on the field', in a rescue context, when time is always missing.

The system implies new methods for team management, and an important and permanent communication of the different parts of the team at any steps. This process questions also the traditional duality between excavation step and post-excavation step.

We hope that this experimentation, still in progress, can provide new directions for the management of an archaeological excavation and the collection and process of field information in the context of rescue archaeology.

This relatively new approach rises up a necessary reflection about the back-up strategy at any-scale from the site to the legal inventory and also in French administrations concerned by the management of archaeological information.

References

BOLO, A., FONT, C., FONTAINE, A., HULIN, G., MAZET, S., RASSAT, S. (2010). The CSNE (Canal Seine-Nord Europe) Project. In: Contreras, F., Farjas, M., Melero, F.S. (eds.). *Conference on Computer Applications and Quantitative Methods 'Archaeology Fusion of Cultures' (Granada 2010)*, BAR International Series, 2494. Oxford: Archaeopress.

BOLO, A., MUYLDER, M., FONT, C., and GUILLEMARD, T. (2014). De la tablette PC à la cartographie de terrain: exemple de méthodologie sur le chantier d'archéologie préventive de Noyon (Oise). In: Costa, L., Djindjian, F., Giligny, F. (eds.). *Actes des 3èmes Journées d'Informatique et Archéologie de Paris, JIAP 2012, Archeologia e Calcolatori. Supplemento 5*. 2014. Borgo San Lorenzo: Edizioni All'Insegna del Giglio s.a.s. p.247-256. Available from: URL: http://soi.cnr.it/archcalc/indice/Suppl_5/20_Bolo_et_al.pdf [Accessed: 27th August 2014].

CIEZAR, P. (2012). Acquisition et enregistrement de données à grande échelle en archéologie préventive. Observations autour des systèmes d'enregistrement à l'INRAP. In: Giligny, F., Costa, L., Djindjian, F., Ciezar, P., Desachy (eds.). *Actes des 2èmes Journées d'Informatique et Archéologie de Paris - JIAP 2010 (Paris, 11-12 juin 2010), Archeologia e Calcolatori. Supplemento 3*. 2012. Borgo San Lorenzo: Edizioni All'Insegna del Giglio s.a.s. p.215-228. Available from: URL: http://soi.cnr.it/archcalc/indice/Suppl_3/16-ciezar.pdf [Accessed: 27th August 2014].

COSTA, L. (2012). La mise en place d'un observatoire des pratiques géomatiques dans les organisations de l'archéologie. In: Giligny, F., Costa, L., Djindjian, F., Ciezar, P., Desachy (eds.). *Actes des 2èmes Journées d'Informatique et Archéologie de Paris - JIAP 2010 (Paris, 11-12 juin 2010), Archeologia e Calcolatori. Supplemento 3*. 2012. Borgo San Lorenzo: Edizioni All'Insegna del Giglio s.a.s. pp.265-278. Available from: URL http://soi.cnr.it/archcalc/indice/Suppl_3/20-costa.pdf [Accessed: 11th November 2014].

DESACHY, B. (2008). *Formalisation du traitement des données stratigraphiques en archéologie de terrain*. Thesis submitted in partial fulfilment of the requirements of Paris I University for the Degree of Doctor of Archaeology. Paris I University. Online since 18 July 2009. Available from: URL: http://le-nid-du-stratifiant.ouvaton.org/spip.php?article6 [Accessed: 27th August 2014].

DJINDJIAN, F. (2010). Quantifier les processus archéologiques. In: Biro, K.T. (ed). *Quantitative Methods for the Challenges in 21st Century Archaeology, Proceedings of the Annual Meeting of the UISPP 4th Commission (Budapest, Hungarian National Museum, 5-6 June 2009), Archeologia e Calcolatori*. XXI. Borgo San Lorenzo: Edizioni All'Insegna del Giglio s.a.s. p.233-247. Available from: URL: http://soi.cnr.it/archcalc/indice/PDF21/13_Djindjian.pdf [Accessed: 27th August 2014].

HOFMANN, E. and MANI, C. (2012). De la découverte au SIG: l'exemple de la base terrain du service archéologique de la Ville de Lyon. In : Giligny, F., Costa, L., Djindjian, F., Ciezar, P., Desachy, B. (eds.). *Actes des 2èmes Journées d'Informatique et Archéologie de Paris - JIAP 2010 (Paris, 11-12 juin 2010), Archeologia e Calcolatori. Supplemento 3*. 2012. Borgo San Lorenzo: Edizioni All'Insegna del Giglio s.a.s. p.141-152. Available from: URL: http://soi.cnr.it/archcalc/indice/Suppl_3/10-hofmann-mani.pdf [Accessed: 27th August 2014].

HUSI, P. and RODIER, X. (2011). ArSol: An Archaeological Data Processing System. In: Jerem, E., Redő, F. and Szeverényi, V. (eds.). *On the Road to Reconstructing the Past. Computer Applications and Quantitative Methods in Archaeology (CAA). Proceedings of the 36th International Conference. Budapest, April 2-6, 2008*. Budapest: Archeaeolingua. p. 86-92 (CD-ROM 278-284). Available from: URL: http://proceedings.caaconference.org/paper/cd36_husi_rodier_caa2008/ [Accessed: 27th August 2014].

KOEHLER, A. and TUFFÉRY, C. (2012). Harmonisation des méthodes et outils pour l'information archéologique à l'INRAP: constats, enjeux et perspectives pour un établissement national. In: Giligny, F., Costa, L., Djindjian, F., Ciezar, P., Desachy B. (eds.). (2012). *Actes des 2èmes Journées d'Informatique et Archéologie de Paris - JIAP 2010 (Paris, 11-12 juin 2010), Archeologia e Calcolatori. Supplemento 3*. 2012. Borgo San Lorenzo: Edizioni All'Insegna del Giglio s.a.s. p.229-238. Available from: URL: http://soi.cnr.it/archcalc/indice/Suppl_3/17-koehler-tuffery.pdf [Accessed: 27th August 2014].

RODIER, X., CIEZAR, P., and MOREAU, A. (2010). Using GIS in French Rescue Archaeology. The Choice of INRAP: a Tool for Research at the Scale of Excavation. In: Contreras, F., Farjas, M., Melero, F.S. (eds.). *Conference on Computer Applications and Quantitative Methods 'Archaeology Fusion of Cultures' (Granada 2010)*, BAR International Series, 2494. Oxford: Archaeopress.

RODIER X. (ed.), BARGE O., SALIGNY L., NUNINGER L., BERTONCELLO F. (2011). *Information spatiale et archéologie*, Coll. 'Archéologiques', Paris: Errance.

Cartography and Heritage: Past Practice and Future Potential for Mapping Scotland's Cultural Heritage

Peter MCKEAGUE

Royal Commission on the Ancient and Historical Monuments of Scotland, UK
Peter.mckeague@rcahms.gov.uk

Abstract

Since 1983, responsibility for the depiction of 'antiquities' on maps published by Britain's national mapping agency, the Ordnance Survey, has rested with English Heritage, and the Royal Commissions on the Ancient and Historical Monuments of Scotland and Wales. This relationship has been maintained through a period of significant change, as digital formats compete with paper maps. The monopoly of supply has been challenged by alternate mapping and imagery products. Developments in survey hardware have helped democratise the collection of highly accurate datasets and, encouraged by the growth of developer-funded archaeology in Britain, there are now many more archaeological companies undertaking accurate surveys. However, restrictions governing the selection and publication of archaeological information on paper maps hinder the potential for presenting the wealth of archaeological information in a digital environment, and new approaches are called for to realise the value of data gathered at great expense.

Keywords: Cartography, GIS, INSPIRE, Inventory

1. Introduction

Archaeological sites, or antiquities, have featured on maps produced by the Ordnance Survey (OS), Britain's national mapping agency, since the 1830s. From 1983, responsibility for the selection and survey of sites for depiction has rested exclusively with English Heritage, and the Royal Commissions on the Ancient and Historical Monuments of Scotland and Wales (RCAHMS and RCAHMW). Since then, digital processes have revolutionised survey, information management and map production, although their full impact has not been realised for mapping the historic environment, where current practice is routed in 19th century cartographic approaches. This paper describes the limitations and necessarily selective nature of traditional cartographic approaches and explores how digital processes can reinvigorate the large-scale mapping of the historic environment.

2. The role of the Ordnance Survey

The OS is responsible for the topographic survey and publication of the natural and manmade features of Britain. Maps, in both paper and digital formats, range from small-scale products at 1:250,000, 1:50,000 and 1:25,000 – ideal for exploring the countryside and Britain's rich heritage – through to large-scale (Basic Scale) maps at 1:10,000 for moorland and mountainous areas, 1:2,500 for small towns and rural areas and 1:1,250 for town centres, suited for professional use. In 1992 OS Land-Line, the first nationally available large-scale vector map was released, and by 1997, except for print-on-demand services, publication of large-scale paper maps ceased. OS Land-Line was tile-based, and had fairly limited attribution. It was superseded by OS MasterMap™ in 2001, which included polygon features for the first time, enhanced attribution, and unique reference numbers, the TOID (TOpographic IDentifier), for individual objects on the map. The map was no longer a map but a seamless spatial geo-database (McKeague and Hart 2003).

2.1. Cadastres

The OS maps boundary features but does not determine boundary ownership, record land valuation or land use, although these functions are carried out by other government agencies. Britain is unusual in most major economies in that it does not maintain a cadastre (Grover 2008). These large-scale maps were accompanied by registers recording the extent, value and ownership of land. Each land parcel has a unique number on the map with a corresponding number and entry in the register. These early examples of joined-up geography were generally resisted in Britain. For instance the Tithe Commutation Act (1836) replaced the ancient system of tithes, traditionally a tenth of the produce of the land levied by the Church of England, with a money payment. The Act required boundaries to be mapped at large scale to determine the value of the land and therefore the amount of tax to be raised. Despite recognition by those undertaking the mapping, that this work should form the basis of a national cadastre, the legislation was amended to prevent Tithe Maps from being used for this purpose (Grover 2008). Although 25-inch (1:2,500 scale) OS maps assigned field parcel numbers consecutively within each parish, with the acreages of fields and often the type of land use recorded in separate Parish Area Books or Books of Reference, they did not define land ownership. Instead, other institutions

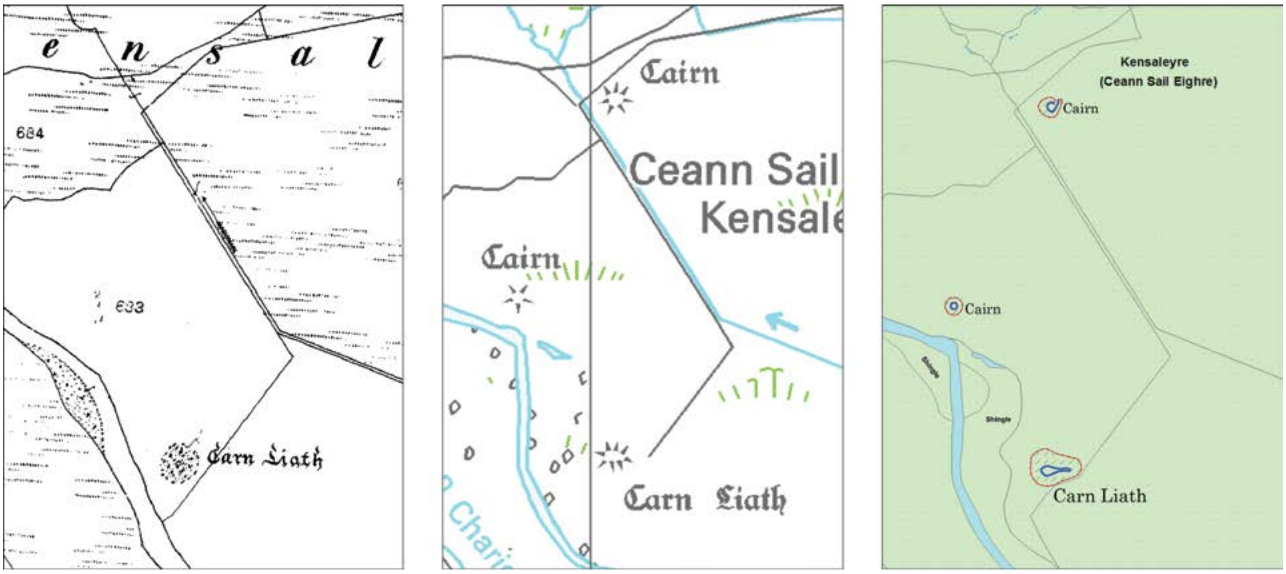


Figure 1: The depiction and annotation of antiquities on maps changes over time: Càrn Liath, Isle of Skye, depicted on the 1st edition OS 6-inch map (Ordnance Survey 1881),
current 1:10,000 edition OS map and OS MasterMap. Note the changes in annotation for burial cairns added on later editions of the map and the simplification of hachures representing slopes, to simple lines on the current vector map.
Current edition OS maps

hold information for their own particular task, creating information silos with the map base simply forming a background. This is an approach that unnecessarily disarticulates the written from the visual and hinders the effective sharing, and combination of data - something that can be redressed through digital mapping approaches.

2.2. Archaeology on Ordnance Survey Maps

The systematic inclusion of archaeological sites on national maps in Britain dates back to the mid 18th century. 'Greatly embarrassed for want of a proper Survey of the Country' (Committee of the Royal Scottish Geographical Society, 1973) during the Jacobite Rebellion of 1745-6, the military authorities commissioned William Roy, assistant Quartermaster at the Board of Ordnance, to undertake a Military Survey of Scotland. The Roy Map at one inch to 1,000 yards (1:36,000) represents 'a landmark of cartography, exemplifying emerging standards in military surveying and cartography, holding a seminal influence on British military mapping in the later 18th century... holding a crucial influence on the early work of the British Ordnance Survey itself' (Fleet and Kowal, 2007, 195). The Military Survey stimulated Roy's interest in antiquities and in particular Roman military monuments, leading to the publication of a volume of large-scale surveys of Roman fortifications (Roy, 1793). He believed that 'the relatively unchanged Scottish landscape provided a line of communication and sympathy between the military men of his own time and the Roman generals who stalked that same territory 1,500 years previously' (Hewitt, 2010, 36). If the locations of fortifications provided the strategist with an insight to commanding and controlling a landscape, more pragmatically antiquities provided essential way marks and reference points in an unfamiliar landscape.

The importance of antiquities in the context of military mapping ensured that the specifications of the National Survey, drawn up under the threat of a French Invasion in 1791 and leading to the foundation of the Ordnance Survey (OS), recognised the need to systematically depict these features. Since then archaeological sites, or 'Antiquities' have been published on OS maps. As part of a military organisation, army staff were responsible for the survey but relied on local knowledge – from landowners, the clergy and historians – for the correct orthography information recorded in the Object Name Books. Information recorded included the (published) name as written on the Plan, various modes of spelling the same Name, the authority for those modes of Spelling and the Descriptive Remarks or other General Observations which may be considered of Interest. For instance, the entry for Càrn Liath on the Isle of Skye (Figure 1), cites the Reverend Darroch, the Reverend Martin and Alexander McLeod as the authorities for the name and notes that 'This name applies to a circular cairn of stone about 22 yards in diameter by about 15 feet high it is situated at the head of Loch Eyre and the northern Bank of the River Haultin' (Ordnance Survey, 1876-8). The entry was subsequently annotated in red 'Supposed to contain the bones of the rival clans who fought here in 1539'. The information recorded in the Object Name Book was used to inform how sites were labelled on the map; an integral part of the process of map production, providing the authority for the published name.

Together the mapped depictions and the accompanying descriptions provide the 'events', the specific observations collected at a single point of time, that first describe and define many of the monuments recorded in national and local inventories today.

Until the end of the First World War identification and interpretation of antiquities took place in consultation with local 'authorities' but in 1920 the OS appointed its first archaeological officer, OGS Crawford eventually leading to the creation of an archaeology branch within the OS. In the tradition of Roy, OGS Crawford produced a thematic map of Roman Britain in 1924 (Ordnance Survey, 1924). After the Second World War under OGS Crawford's successor, Charles Phillips, the Ordnance Survey Archaeological Division developed a card index providing a comprehensive record of the known archaeology of Britain. Although the primary purpose of the card index was to inform the selection and management of antiquities for publication on OS maps, the card index also served as a national non-intensive Inventory of the archaeology of Britain. Site locations were annotated on OS 6-inch maps and cross-referenced to a card index which detailed the location and type of monument, supported by descriptive accounts based on field observations from Ordnance Survey field investigators, and précis of published accounts, including Commission inventory entries, and supported by bibliographic references. Often, if an antiquity was considered for publication, an inked measured survey drawing (Figure 2), usually at the appropriate mapping scale to help inform draughting map detail for publication, would be attached to the card. Following a review of the role and procedures of the Ordnance Survey (Serpell, 1979) the responsibilities of the Archaeology Division were transferred to the three Royal Commissions (RCAHMS, RCAHMW and RCHME (now incorporated within English Heritage)). In Scotland, the information on the card indexes had been computerised by 1992 and now forms the core of the Canmore records published online. The organisation of information on these card indexes also informs much of the data and data management required for modern computerised systems.

2.3. Ordnance Survey Specifications

The depiction of sites on OS maps is governed by cartographic conventions. Cartographic principles govern the selection and depiction rules for antiquities on OS maps, inevitably presenting a selective view of the nation's heritage. Depiction of modern mapped detail and annotation takes precedence over historical features. The OS defined antiquities as extant artificial features over 1m in vertical height, although they also regularly published earthworks of lower relief. Features would be exaggerated in order to depict the monument adequately. For a similar reason, inferred features could be shown to indicate the 'site of' or 'course of' – a particularly useful device for depicting the line of a Roman road. Until 1989, when the rule was relaxed, 'antiquities' had to predate AD 1714. Selected sites are usually labelled in 'Antiquity type' an Old English Gothic font – although a sans serif 'Roman' font was adopted to annotate Roman sites and until the late 1950s antiquities post-dating 1066 were labelled with a gothic font (Oliver, 1993). Early editions of the OS maps often included the locations of find-spots of portable objects, the locations of destroyed monuments, nationally important battlefields and natural features associated with historic events. The depiction of features first portrayed in the 19th century has persisted through to modern

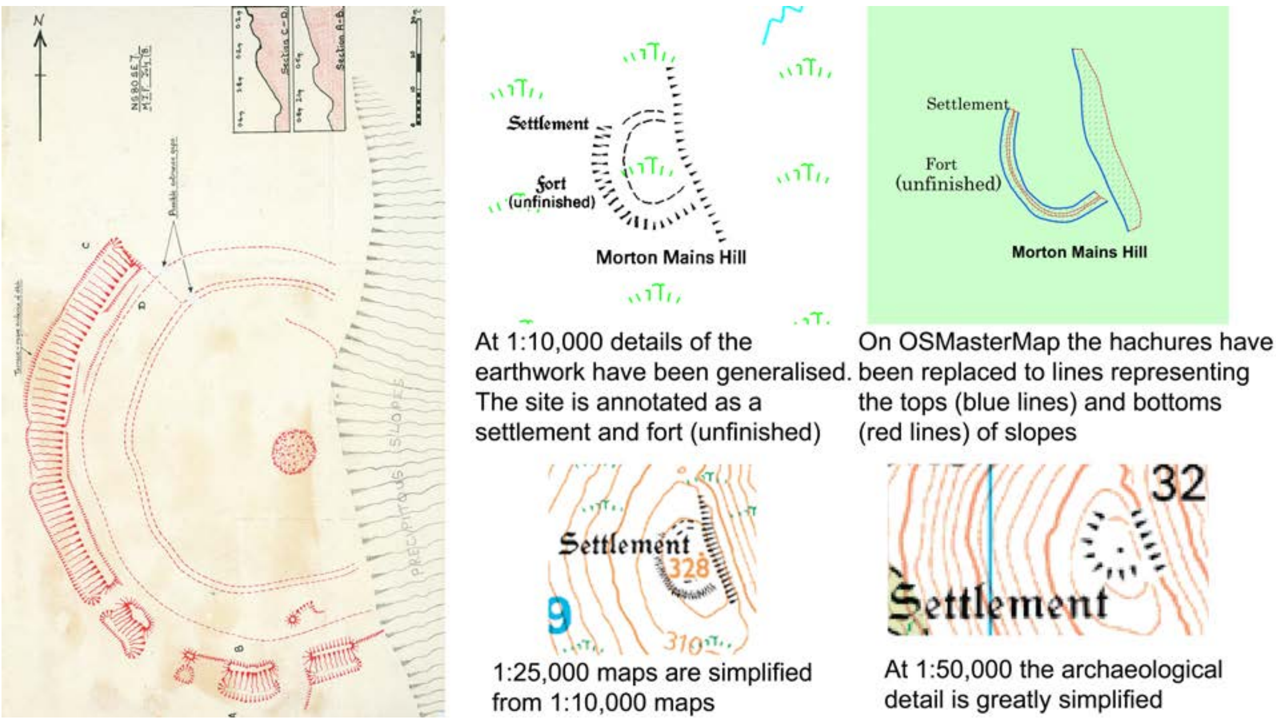


Figure 2: Details from the 1978 survey drawing of Morton Mains fort (NS80SE 7), from the OS Card Index held at RCAHMS are generalised for publication on OS Maps. Gradually site representation becomes highly stylised. Image SC 866919. © Crown copyright: RCAHMS. Licensor www.rcahms.gov.uk, current edition OS maps © Crown Copyright and database right 2014, All rights reserved OS Licence No. 100020548.

maps until quite recently, even though these categories of antiquity were outside OS specifications. Maps were rarely updated with new detail.

Selection of antiquities can be informed by their national significance. However, not all nationally significant, or statutorily protected, sites will be depicted on the map. For the built environment, most listed buildings and industrial monuments are depicted as topographic features in their own right, but they may be labelled in Antiquity type. However, selection is limited to those of exceptional interest or open to the public and, without exception, modern features take precedence over historic detail. As publisher, the OS retains editorial control over the map.

2.4. Questions of scale

The depiction of antiquities across the suite of OS maps is governed by scale and convention, and for archaeology is necessarily selective. There are over 145,000 archaeological records in the Scottish national inventory, and it is unrealistic to expect all to be published on OS maps. Using the selection criteria outlined above, some 25,000 antiquities are published on Basic Scale maps, with a similar number featuring on 1:25,000 scale maps, originally a product derived from the 1:10,000 maps. On 1:50,000 scale maps the number of sites featured reduces to about a third and the level of detail becomes very stylised (Figure 2). Only the most significant sites, usually well known tourist attractions, are depicted at 1:250,000.

The representation of an archaeological site on a map is just that: a representation. It is governed by cartographic rules that influence how much detail can be depicted. The following example of an earthwork survey at a palisaded settlement and unfinished fort at Morton Mains, Dumfriesshire (Figure 2) illustrates the generalisation of archaeological detail for cartographic purposes. Originally surveyed at 1:2,500 in 1978, the plan shows the enclosing bank and ditch of the fort as well as the twin palisade, one side of the gateway feature and the remains of what may be a cairn.

For publication at 1:10,000 scale, the breaks in the rampart and ditch have been generalised to indicate an almost continuous arc of earthwork. Despite the 1978 survey, rather misleadingly the hachures have been reversed suggesting that the feature is primarily a ditch. This depiction references the original depiction of the earthwork on the 1st edition OS 6-inch map (Ordnance Survey, 1861). Within, the twin palisade is reduced to a single line on all but the N side, where the absence of the external rampart allows for the depiction of the segment of the outer palisade. The transverse palisade slot of the gateway visible on the original survey drawing has been omitted. Derived from the published 1:10,000 map, the 1:25,000 map retains the same level of detail as previously published, although the annotation has been reduced from 'Settlement' and 'Fort (unfinished)' to simply 'Settlement'. At 1:50,000 scale, the map depiction has been further generalised to provide an overall impression of the site. With the transition from paper maps to digital formats, the hachures depicting slopes have been simplified to lines representing the top and bottom of slope and the slope symbolised. The meaning conveyed by the hachures has been lost although they can be generated through automated programs (Regnauld, MacKaness and Hart 2002).

2.5. Depiction and annotation issues

Archaeological features on OS maps are represented on the published paper maps because, by and large, they are topographic features but they also reflect antiquarian interests. Early editions of the OS maps also included find spots and notable events. The depiction of features first portrayed in the 19th century has persisted through to modern maps until quite recently even though these categories of antiquity were outside OS specifications. On 1st edition maps, antiquities were more often annotated with their distinctive name 'Càrn Liath' rather than labelled with their descriptive name or classification 'Cairn' (Figure 1). A 'descriptive' name following standardised terminologies was preferred for antiquities added on later editions. Again these labels often persevere onto modern editions of the map alongside additional detail simply annotated with their classification. Depictions may try and convey the shape of a site, or they may simply be symbolised. In Figure 1 the modern map attempts to indicate the shape of 'Càrn Liath' and the cairn at the north edge of the extract whilst a standard mound symbol is used to indicate the position but not the shape of the cairn in between. When the standard mound symbols were converted into a digital format, creation of linework to indicate the tops and bottoms of slope inferred a misleading geometry for these symbols.

2.5. Bias

Information published on OS maps must meet technical specifications. As the agency responsible for providing new and updated information to the OS, RCAHMS is unable to guarantee that information reported by other archaeologists meets the required specifications. For instance RCAHMS has records for 1,849 burnt mounds in Scotland; the known distribution of these sites is uneven nationally and marked by discrete dense clusters of sites reflecting patterns of fieldwork (Cowley, 2011). Only those surveyed by OS staff and latterly RCAHMS are considered for publication, and many of these do not meet OS specifications. As a result only 632 burnt mounds are published at Basic Scale, presenting a skewed representation of our knowledge of the monument type.

Not all monuments are treated equally on the map. Although the map specifications clearly define the selection criteria, exceptions are made for antiquities of Roman date. A separate font type is used to annotate Roman sites across the map scales. On Basic Scale maps the locations of Roman military sites are often depicted as non-obstructing features, even though there are no visible earthworks. The separate treatment of Roman antiquities references the military background of the OS and the interest William Roy showed in mapping military antiquities.

3. Digital opportunities

Depiction of archaeological sites on OS paper maps is governed by questions of scale, standards and specifications that inform the selection process. The result is a cartographically pleasing product, highlighting selected sites that are interesting topographic landmarks or real-world objects. These features can be depicted and annotated in a range of ways appropriate to the map scale but functionality is greatly restricted by the medium. What works for a paper map, does not work in a digital environment. As the large-scale maps evolve into searchable spatial geo-databases, the incomplete picture of the historic environment can be misleading. The supply of 'antiquities' for publication on maps also needs to be re-evaluated in light of wider technological developments.

3.1. The impact of technological change

Arguably the greatest impact of technological change on cartography is socio-economic. The audience for paper maps is relatively select – people who have made a conscious decision to buy maps. Today, through navigational mapping and location-based services delivered over the Internet to multiple platforms or downloaded onto portable devices, people are much more familiar with and comfortable using maps, even if the map detail is often greatly simplified.

Although there is continued demand to purchase recreational maps in both paper and digital formats, where there is a need for large-scale mapping, these products are licensed rather than purchased. The impact of licensing on knowledge creation and sharing should not be underestimated. In a case study explored at CAA 2013 (Hamilton and McKeague, forthcoming), 'whilst the Intellectual Property Rights (IPR) of information referenced against paper maps was not overtly an issue for paper maps, it quickly became a dominant theme with the new business models emerging with the establishment of the OS as a trading fund and the release of vector-based mapping on an annual license'. Although the rules governing shared IPR are gradually relaxing, the onward use of data directly copied from base map features for inclusion in another dataset is greatly restricted.

There is also now effective competition: although the OS still provide premium map products, they are no longer exclusive providers. OpenStreetMap provides an alternative large-scale crowd-sourced map, whilst commercially available ortho-imagery products offer alternative options. Although lacking the detail provided by large-scale mapping, products like Bing or Google Maps as well as OS Open Data, underpin most map-based internet applications.

Technological developments have also helped democratise data gathering. Recreational, hand-held GPS and even GPS-enabled mobile phones have lowered the bar for capturing data, enabling anyone to locate something to within a few metres. Higher specification equipment enables highly accurate mapping of archaeological detail independent of local map control, whilst laser scanning technologies capture real world features as three dimensional point clouds.

Coupled with the widespread availability of technological tools to simplify data capture in Scotland there has been a marked growth in both developer-funded archaeology during the 1990s and community-led projects to use these tools. It is no longer the preserve of state organisations to map the archaeology of the nation though, for purposes of quality assurance and level of metrical accuracy, it remains their exclusive right to publish surveyed material on OS maps.

With so much choice in background mapping products, and much more accessible tools to gather data, Application Programme Interfaces (API), Web Map Services (WMS) and Web Feature Services (WFS) offer solutions to publishing a vast range of information against those backgrounds. Through OS Open Space (Ordnance Survey n.d.), the OS offer guidance on using their API to create applications immediately democratising the process of publishing spatial information in a web map browser (see Krüger n.d). Publishing data as WMS or WFS allows far richer datasets to be displayed against the user's preferred background layers, than could ever be realistically achieved through traditional mapping approaches. Through a WMS for RCAHMS data, the user can immediately access the locations of over 300,000 records (albeit as point data) remotely at their own workstations; this is in contrast to the 25,000 sites depicted on a traditional map. With WMS and WFS, users are not restricted to a static view of monuments. The services are dynamic, updated as and when new data is added. Services may also be combined, presenting information from different sources in a single portal. For instance, PastMap (http://pastmap.org.uk/) offers a single entry point for data about the historic environment across Scotland. It combines in a single browser the constraint mapping for designated sites from Historic Scotland; trigger mapping from HERs highlighting areas of potential; and inventory data from Canmore.

3.2. Richer data

Established cartographic practices inevitably focus on presenting information about topographic features, but users can access far richer datasets about the historic environment through downloadable datasets and web services. For instance data from the RCAHMS Historic Landuse Assessment (Millican and Middleton, forthcoming) may be viewed in a map browser (http://hla.rcahms.gov.uk/) or, subject to licensing, downloaded (http://hla.rcahms.gov.uk/content/data-download) for use remotely.

Much more needs to be done with the data created and collected digitally through fieldwork. In a field project, fieldwork extents and features are usually mapped, collated and analysed in a GIS, then published as part of the project report and perhaps eventually deposited in a suitable archive. During the process of completing the project, the spatial value of the data, gathered at some effort and expense, is gradually stripped out and reduced to an illustration in the report. Conventional publishing presents a myopic view of the archaeology as static images without realising the wider spatial potential of the data. It locks information down, preventing discovery, reassessment and reuse of that data in other contexts. Digital data presents an opportunity to take a fresh look at how spatial data is used in archaeology, and to unlock the value trapped on the pages of reports. Doing so would help fulfil the vision of Scotland's GI strategy (Scottish Government, 2005) and the EU INSPIRE Directive (INSPIRE 2007) to improve the effectiveness of how we share spatial information. Data should be collected once and maintained at a level where this can be done most usefully and effectively. Users should be able to combine spatial datasets from different sources and share that data between many users and applications, and the spatial data required for good governance should be available under conditions that do not restrict its extensive use.

3.2. Governance and Technical specifications

Whereas OS paper maps and RCAHMS own survey plans have established conventions to ensure a consistent product (McKeague and Cowley, 2013), there is as yet little standardisation in the digital workplace. Key to delivering INSPIRE is the establishment of a Spatial Data Infrastructure or SDI: the policies, human resources and related activities needed to acquire, process, distribute, use, maintain and preserve spatial data. Work on national SDIs should inform, support and underpin the delivery of cultural heritage spatial information. However, national SDIs may limit access, or place conditions on information to those working in public administration, restricting engagement with academic and commercial organisations outside that 'club' – the very people creating and potentially benefiting from a cultural heritage SDI.

INSPIRE defines the technical specifications for 34 themes but does not specify how data should be visualised on the map. Data providers can define how they want their data to appear, but with so much data from so many providers flexibility is key. What works for one combination of layers might not work with a different range of datasets.

3.4. The third dimension

The growth of digital technologies in creating highly accurate, realistic three dimensional models of the historic environment extends the potential of established cartographic approaches to mapping. Three dimensional models present a more complete representation of real world features and through interactive tools allow the user to visualise the impact of noise and change on the historic environment and protect important vistas (Dorffner, 2005, Dokonal, 2010).

3.5. New life for Old data

Through geo-referencing, new technologies can also breathe life into old data. In the course of fieldwork RCAHMS staff may produce measured survey drawings of individual sites to accompany descriptive accounts of sites. Selected detail from these plans may then be put forward for inclusion as an antiquity on OS maps. Plans from the inventory may be copied and taken out onto the monument, or the OS maps used to locate the most prominent features to assist in on-site relocation of archaeological features, though for the non-specialist user, understanding the relationship between the mapped detail and physical remains may be challenging. For instance RCAHMS surveyed a group of hut circles and field system at An Sithean, Islay in 1975 for the Argyll Inventory (RCAHMS, 1984) (Figure 3).

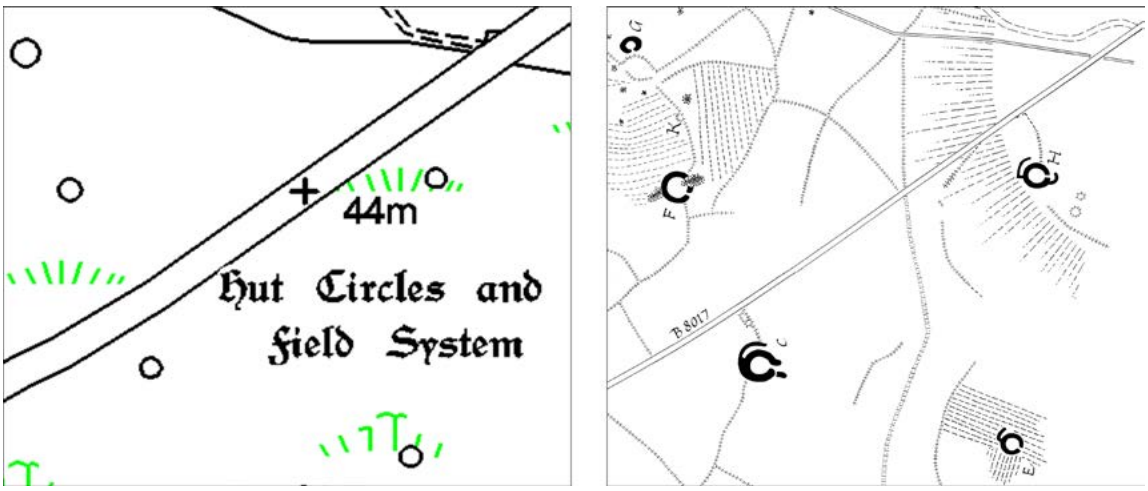


Figure 3: Comparison between published OS 1:10,000 map and an extract from the RCAHMS survey plan (1975) originally drawn at 1:1,000 of the hut circles and field system at An Sithean, Islay. The survey provides a more complete picture of the archaeology of the area than can ever be achieved with traditional paper mapping where only selected monuments are published. Image SC 1364845 © Crown copyright: RCAHMS. Licensor www.rcahms.gov.uk, (SC 1364845), current edition OS map © Crown Copyright and database right 2014, All rights reserved OS Licence No. 100020548.

The area surveyed lies in a protected bird reserve and the vegetation cover is gradually obscuring the archaeological features. Providing a geo-referenced plan as a WMS could assist the reserve staff, unfamiliar with archaeological sites, in relocating features so that the archaeology may be managed more effectively. The WMS can be consumed into the reserve's GIS, viewed against a base map or ortho-imagery, and combined with specialist datasets required for the management of the reserve. Through web services the value of the data collected to document the historic environment is released. Data is also more easily found through metadata discovery services and it is more easily used in a format that assists those charged with the stewardship of the land, including the care of the historic environment.

Creation of WMS for detailed survey plans drawn up to publication standard could redress the decline in the art of cartography, observed with the transition from paper to vector-based products. With the introduction of vector maps the reduction of carefully crafted depictions of earthworks to simplistic lines renders the depiction almost meaningless. Through a WMS, original survey plans could be published without simplification or loss of detail that publication at a mapping scale inevitably requires. This approach offers a number of benefits: it presents the user with a more detailed interpretation of the site than could possibly be conveyed through traditional publication; it delivers efficiencies, saving time taken to prepare separate illustrations for the map base and eliminates the workflow between the supplier (RCAHMS) and the publisher (OS) reducing the time taken to publish the survey detail; moreover, it acknowledges the reality that RCAHMS is no longer the sole survey agency in Scotland – subject to IPR, content could be provided by any of the archaeological companies working in Scotland today.

3.6. Spatial Geo-databases

The evolution of official mapping in Britain, with the resistance to cadastral systems, has encouraged the development of organisational data silos where data is managed for a specific purpose. The potential added value delivered through sharing often remains unrecognised. With the release of OS MasterMap and the introduction of TOIDs in 2001, the mapped detail is no longer simply a visual tool but the core component of a complex spatial geo-database. Complex attribution provides flexibility in presentation of the mapped detail, allowing multiple visualisations of the data. For instance antiquities could be represented thematically, or by period. They could be annotated by their distinctive name or by their descriptive name, drawn from semantically enriched vocabularies such as those published through www.heritagedata.org. The introduction of TOIDs allows the user to create associations between their datasets and the mapped object, potentially reconnecting the mapped feature with the descriptive accounts (McKeague and Hart, 2003). The next challenge is to think outside the map; to detach the spatial component from the map and treat it as part of the narrative account of a record in the same way that the user views images or plans associated with a site record.

4. Summary

Paper maps can only ever present a highly selective view of the nation's heritage. Both the number of sites and the level of detail shown is scale dependant. Nevertheless there is still a place to publish antiquities on OS maps which, after all, reach a far bigger audience than most archaeological publications. However the view presented is partial and for business use grossly under-represents the historic environment. Advances in digital technology challenge traditional approaches to mapping. With the range of mapping products available, more accessible tools to map our heritage to a high standard and growth in archaeological practitioners skilled in survey, the existing model of sole supplier and sole publisher appears very outdated. With APIs, WMS and WFS, data is no longer in the map but layered over the map. These services provide more complete and richer datasets, introduce additional layers representing archaeological constraints or fieldwork extents that can be accessed remotely, and can be layered and combined with multiple other datasets and backgrounds for the benefit of the remote user. Through the power of spatial geo-databases, integration of the descriptive with the visual offers flexibility impossible with paper maps. Digital maps challenge established cartographic practices in representing the nation's heritage and whilst perhaps not quite as visually pleasing a product, offer considerably more potential for publishing data about the historic environment. Whilst these ambitions may present a utopian vision of digital mapping in the future, initiatives such as the INSPIRE Directive provide an impetus and roadmap for spatial data harmonisation.

Acknowledgements

I would like to thank Robin Turner and Kirsty Lingstadt for comments on earlier drafts of this paper and Georgina Brown for her assistance in providing the illustrations.

References

Committee of the Royal Scottish Geographical Society (1973). *The Early Maps of Scotland to 1850*, Edinburgh:The Royal Scottish Geographical Society.

COWLEY, D. (2011). Remote sensing for archaeology and heritage management – site discovery, interpretation and registration. In: Cowley D. (ed). *Remote Sensing for Archaeological Heritage Management. Proceedings of the 11th EAC Heritage Management Symposium, Reykjavik, Iceland, 25-27 March 2010*. EAC Occasional Paper No.5. Occasional Publication of the Aerial Archaeology Research Group No.3, Brussels: Europae Archaeologia Consilium (EAC). pp. 43-55.

DOKONAL, W (2010). Creating and using 3D City Models. *In Architecture and Modern Information technologies* [Online] 1 (10) 2010. Available from:

http://www.marhi.ru/AMIT/2010/1kvart10/dokonal/dokonal.pdf [Accessed: 30th July 2014].

DORFFNER, L. (2005). Das 3d Stadtmodell von Wien – Grundlage für Planungsaufgaben und Visualisierungen. In: Forschungsgesellschaft Wiener Stadtarchäologie (ed) *Archäologie und Computer Workshop 10, 2005*. Vienna: Museen der Stadt Wien – Stadtarchäologie.

FLEET, C. and KOWAL, K.C. (2007). Roy Military Survey map of Scotland (1747–1755): mosaicing, geo-referencing, and web delivery. *e-Perimetron*. [Online] 2 (4), Autumn 2007. p. 194-208. Available from: http://www.e-perimetron.org/Vol_2_4/Fleet_Kowal.pdf [Accessed: 17th July 2014].

GROVER, R. (2008). Why the United Kingdom does not have a cadastre – and does it matter? *Annual Meeting 2008 FIG Commission 7 Verona, Italy, September 11-15, 2008*. [Online] Available from: https://www.fig.net/commission7/verona_am_2008/papers/12_sept/7_2_grover.pdf [Accessed: 17th July 2014].

HAMILTON, S. and MCKEAGUE, P. (forthcoming). Reality Bites: Reviewing a decade of online collaboration. In: Traviglia, A. (ed.). *Across Space and Time, Proceedings of the 41st Computer Applications and Quantitative Methods in Archaeology Conference CAA 2013 Perth, Australia*. Amsterdam: Amsterdam University Press.

HEWITT, R. (2010). *Map of a Nation. A biography of the Ordnance Survey*. London: Granta.

INSPIRE (2007). *Directive 2007/2/EC of the Council of the European Union and European Parliament setting up an Infrastructure for Spatial Information in the European Community (INSPIRE)*, Official Journal of the European Union, L108/1 25.04.2007. (2007) [Online] Available from: http://eur-lex.europa.eu/legal-content/EN/TXT/HTML/?uri=CELEX:32007L0002 [Accessed: 17th July 2014].

KRÜGER, W. (n.d.). *Scottish Highland Bridges / Brücken in den Schottischen Highlands* [Online] Available from: http://www.kweimar.de/SCMap.html [Accessed: 17th July 2014].

MCKEAGUE, P. and COWLEY, D. (2013). From Paper to Digital, and Point to polygon – the Application of GIS in a National Body of Survey and Record. In: *International Journal of Heritage in the Digital Era* 2 (4). pp. 677-694. [Online] DOI: 10.1260/2047-4970.2.4.677

MCKEAGUE, P. and HART, G. (2003). OS MasterMap: Future Directions in Mapping Britain's Past. In: Doerr, M. And Sarris, A. (eds). *The Digital Heritage of Archaeology. Computer Applications and Quantitative Methods in Archaeology, Proceedings of the 30th conference, Heraklion, Crete, April 2002*. Athens: Hellenic Ministry of Culture, p.p 361-367.

MILLICAN, K. and MIDDLETON, M. (forthcoming). The end game: As Scotland's Historic Land-use Assessment project reaches completion what have we learned? In: Traviglia, A. (ed.). *Across Space and Time, Proceedings of the 41st Computer Applications and Quantitative Methods in Archaeology Conference CAA 2013 Perth, Australia*. Amsterdam: Amsterdam University Press.

OLIVER, R. (1993). Ordnance Survey maps: a concise guide for historians. London: Charles Close Society.

ORDNANCE SURVEY (1876-8). *Original Object Name Book, Inverness-shire (Skye)* volume 7, 173 [Online] Available from http://www.scotlandsplaces.gov.uk/digital-volumes/ordnance-survey-name-books/inverness-shire-os-name-books-1876-1878/inverness-shire-skye-volume-07/173 [Accessed: 17th July 2014].

ORDNANCE SURVEY (1861). Dumfriesshire 1st edition 6-inch map, sheet XXII. [Online] Available from http://maps.nls.uk/view/74426653 [Accessed: 30th July 2014].

ORDNANCE SURVEY (1881) Inverness Skye 1st edition 6-inch map, sheet XVII [Online] Available from http://maps.nls.uk/view/74427207 [Accessed: 30th July 2014].

ORDNANCE SURVEY (1924). *Map of Roman Britain.* Southampton: Ordnance Survey.

ORDNANCE SURVEY (n.d.). *Innovate with OS Open Space*. [Online] Available from: http://www.ordnancesurvey.co.uk/innovate/innovate-with-open-space.html [Accessed: 17th July 2014].

RCAHMS (1984). *The Royal Commission on the Ancient and Historical Monuments of Scotland. Argyll: an inventory of the monuments volume 5: Islay, Jura, Colonsay and Oronsay*, Edinburgh: HMSO. pp. 123-6, No.237.

REGNAULD N., MACKANESS, W. and HART, G. (2002). Automated Relief Representation of Archaeological Monuments and Other Antropogenic Forms. *Computers, Environment and Urban Systems* 26 (2-3). pp. 219-239.

ROY, W. (1793). *The Military Antiquities of the Romans in North Britain*. London: Society of Antiquaries.

SCOTTISH GOVERNMENT (2005). *One Scotland, One Geography; a geographic information strategy for Scotland*. Edinburgh: Scottish Government. [Online] Available from: http://www.scotland.gov.uk/Publications/2005/08/31114408/44098 [Accessed: 17th July 2014]

SERPELL, D. (1979). *Report of the Ordnance Survey Review Committee*. London: HMSO

Visualization Based on the Norwegian University Museum Database

Espen ULEBERG, Mieko MATSUMOTO

Museum of Cultural History, University of Oslo, Norway
espen.uleberg@khm.uio.no, mieko.matsumoto@khm.uio.no

Abstract

The project Dynamic Distributions investigates changing relations between humans and landscape during the Stone Age in South-Eastern Norway. The archaeological data is collected and reused from the database at the Museum of Cultural History at the University of Oslo.

Artefacts that are selected on background of accuracy of coordinates and dating are mainly stray finds dated typologically. They are analysed in relation to landscape regions in three different time periods; Late Mesolithic Nøstvet phase, Early Neolithic/Middle Neolithic A, and Late Neolithic/Early Bronze Age. The accuracy levels decide the scale of analyses that can be applied, while the time scale depends on the duration of the archaeological periods. The geographical coordinates of the artefacts are categorized according to different accuracy levels. Principles from Exploratory Data Analysis are applied to enhance temporal and geographical relationships. The sudy visualizes a continuation and gradual intensification of the use of inland landscapes.

Keywords: Museum database, Landscape regions, Exploratory Data Analysis, Visualizations, Norwegian Stone Age

1. Introduction

Dynamic Distributions is our project collecting and reusing data from the database at the Museum of Cultural History at the University of Oslo. The project aims at investigating the changing relation between humans and landscape during the Stone Age in South-Eastern Norway.

The present paper is mainly concerned with methodological aspects of the project. Considerable data sources are becoming digitally available, and archaeologists can work with large datasets that open new possibilities of analyses and visualization. It is desirable to be able to find all information online. At the same time, the totality can become overwhelming if the datasets are not filtered or presented in ways that make it possible to present them in visualizations like thematic maps and tables. Exploratory Data Analyses (EDA) is a way of looking for patterns through a series of different types of analyses. It is an iterative process, where initial observations will give ideas to new constellations that again can lead to new types of analyses. Online datasets presents opportunities to work with EDA without the initial, time-consuming creation of datasets from literature and archives.

2. The basics

2.1. The database

The six Norwegian university museums cooperate to create common database solutions through MUSIT (www.musit.uio.no). The database is event based and is developed in line with the CIDOC-CRM concept (Jordal *et al.,* 2012:256). The artefact catalogues, both hand-written and printed, were digitized in the 1990s. The text was then SGML-tagged (Holmen *et al.,* 1996) and later imported into the database. In this way, the original terms used for artefacts, materials and places were registered.

The majority of the artefact terminology on the website is still the same as used in the original catalogues. Some of the terminology could be normalized with synonym lists, but in other cases only a re-cataloguing will give a precise result.

The database preserves all old versions of the terminology when the information is revised. The objects used in Dynamic Distributions have been re-catalogued, and the information added to the main database. In this way both the database and the web page are constantly updated. This procedure is chosen rather than creating a separate database for the project. It is preferable that the results become available at once without a later phase when the main database is to be updated.

The database has since 2005 been used for cataloguing and registering of finds in the archaeological collections in Norway. It is a well-functioning museum database system with an acquisition protocol, artefact catalogue, conservation and loan modules, and a module for photography. The content is consecutively published on the website (www.unimus.no) and is freely available to all. The website has export facility to text format, and links to full catalogue texts and images. The text export includes coordinates with accuracy levels. The images are in high resolution and can be downloaded and used under a Creative Commons license (CC BY-NC 3.0).

The original catalogue texts also contained cadastre information that is outdated. Due to later normalization, it is now possible to query the material using updated cadastral information. Most of the objects have been given a geographic position through coordinates with varying accuracy levels. An important aspect is the relation between archaeological inquiries and the accuracy of the coordinates. The different accuracy categories reflect the provenance of the artefacts. Objects that are not integrated in the museum collection soon after their discovery will often have provenance with low accuracy. This does however not mean that they cannot be used at all, but it restricts their use to analyses at coarser scales. When addressing the national or European distribution of an artefact, a rather low accuracy will suffice. When looking at distance from water, height above sea level or preferred soil types, a higher level of accuracy is necessary.

The main categories of accuracy in the database are site, three levels of cadastral units, two levels of ecclesiastical units, municipality, and county. In addition there are place names that have varying accuracy, according to the size of the area they refer to. The ordering of place names follows the standard of the Norwegian Mapping Agency (Kartverket). A place name that indicates a vast area, like a large valley or mountain, will have a low accuracy. The name of a small lake or inlet will have a higher accuracy. A verbal description that indicates a limited area within a farm can be at a level near the site accuracy. For this paper, the individual place names have been classified on an ordinal scale to make them comparable to the other categories and used appropriately in the analyses.

The selection of database entries chosen for this project consists of 4000 units (Fig. 1). Only 9% of the artefacts have a site level accuracy, while 62% have the accuracy of the Cadastral entity 2 (Farm). Visualizations including sites and all three cadastral entities can use 89% of the units. Each object is assigned a coordinate. A cadastral unit can be a large area covering different ecological zones, but the coordinate indicates a point in the farm yard. Concerning ecclesiastical units, the coordinates for the church are chosen. Similarly, the town hall is chosen instead of a geographical midpoint for a municipality, as this gives a better indication to the reader that it is a coordinate with low accuracy. For artefacts from known sites, modern excavations and surveys, the coordinates are very precise.

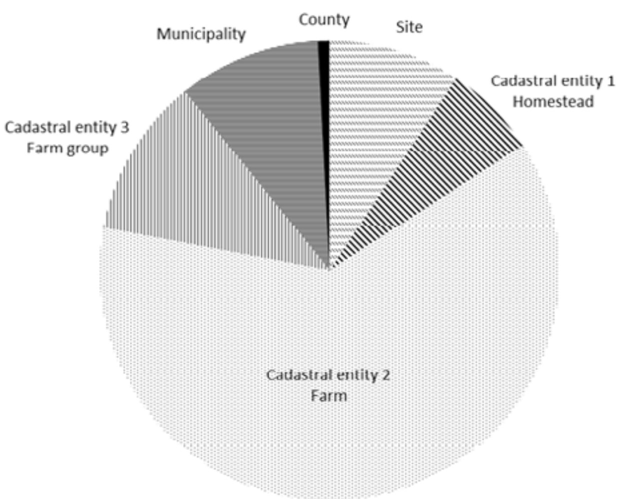


Figure 1: The accuracy levels of find provenance .

The national Norwegian sites and monuments register, Askeladden (askeladden.ra.no), has also information about places where artefacts were found (Berg, 2007). The reason why Askeladden does not contain all finds is mainly because it will only include sites with an accuracy level of site or in some cases cadastral farm. For a research database, it will in some instances be useful to know the overall position of as many objects as possible, even only at the level of a municipality. The MUSIT database has therefore maintained its own set of coordinates.

2.2. Landscapes

The archaeological finds chosen for this paper are from six counties in South-Eastern Norway; Akershus, Oslo, Hedmark, Oppland, Buskerud, and Vestfold (Fig. 2). They cover the landscapes from the Oslo fjord in the south through agricultural areas, valleys and forests up to high mountain areas in the North and Northwest. The archaeological finds are combined with maps of landscape types, and larger lakes and rivers.

The landscape types presented here are based on the work by Oscar Puschmann (Puschmann, 1998; 2005) at the Norwegian Forest and Landscape Institute. He has defined four levels applicable at different scales. The uppermost has five landscape types, reflecting today's agricultural areas. It is constructed for a scale level of 1:2 million. The next level has 45 different landscape regions. They are coastal areas, different types of forest and agricultural areas. The accuracy of this mapping is for a scale level of 1:500 000. The third level (1:250 000) consists of 444 landscape sub-regions, and can in size be compared to municipalities. The fourth level is not described in detail, but is a further division of sub-regions into landscape areas. One of the main objectives of Puschmann's work was to operationalize the landscape regions as geographical units of analysis and obtain a clear indication of the geographical extent of environmental problems. The three dimensional contents of the landscape has been of central importance. The classification is based less on traditional science and cartography but more on a multidisciplinary understanding and holistic evaluation (Puschmann, 1998:1).

The Landscape reference system is hierarchical, and the divisions are made on basis of six variables; Major landform, Geology, Water and waterways, Vegetation, Agricultural areas, and Buildings. The Major landform is the dominant form of the landscape. The Geology takes both bedrock and deposits into consideration and adds geological detail to the major landform. The variable Water and waterways subsumes lakes, fjords and sea as well as streams, rivers, and waterfalls. The Vegetation pattern concerns natural, semi-natural and managed vegetation. The variable Agricultural areas concerns suitability for agricultural and land use; meadows, fields, and pastures. The variable Buildings also covers technical installations, and is perhaps the variable most reflecting modern use of the region and least relevant for an archaeological study. The description of the regions has been made on basis of thematic maps and data registers (Puschmann, 2005).

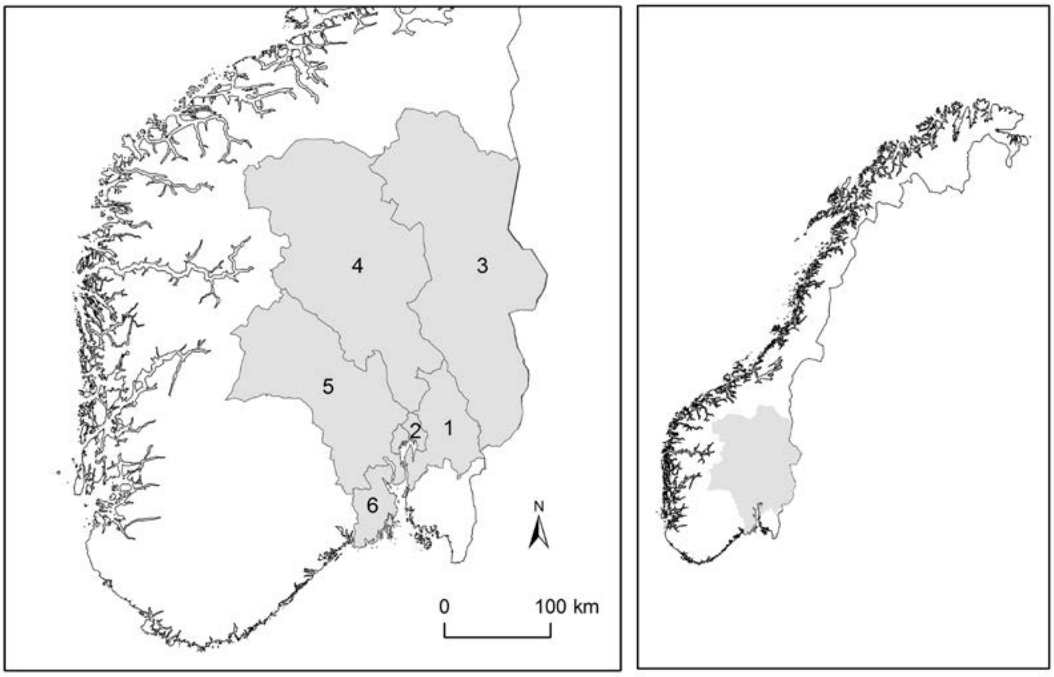


Figure 2: *T*he six counties in South-Eastern Norway; 1. Akershus, 2. Oslo, 3. Hedmark, 4. Oppland, 5. Buskerud, 6. Vestfold

The divisions reflect the present use of the landscape. Nevertheless, the modern agricultural landscape is dependent on the overall basic limitations and conditions. This landscape division has provided a useful starting point for analysis of prehistoric settlement. Solheim (2012:52–55) used the intermediate level for an aggregation into three zones in Eastern Norway. Matsumoto et al. aggregated 10 intermediate levels into seven when describing Neolithic finds in parts of South-Eastern Norway (Matsumoto et al. in print).

The six counties cover seventeen of Puschmann's intermediate level landscape types. Two additional factors are necessary to interpret and visualize the archaeological artefacts in these landscapes. One is the eustatic land rise after the Ice Age, and the other is an archaeological dimension reflecting the distribution of archaeological sites and our general understanding of the landscapes. As a result, seven landscape types are identified. At this scale, it is possible to use sites that have an accuracy level of cadastral unit or better.

The land in South-Eastern Norway was pressed down by the heavy Scandinavian ice sheet. After the end of the Ice Age, around 10 000 BP, the land rise was quite rapid, and the old coast line is now as high as 220 m asl in the inner part of the Oslo fjord (Sørensen, 1979). Consequently the elevation around the Oslo fjord is an important element in understanding the use of landscapes in the Stone Age. What is now an inland settlement was a site at the coast or fjord. The Digital Elevation Model used in this paper is from the 50 m grid provided by Norwegian Mapping Authority. Analysis that involves the land rise should mainly use finds with a site level accuracy.

The archaeological dimension of the landscapes reflects the distribution of sites, i.e. the use of the landscapes in prehistory. The initial mapping of raw data on all seventeen landscape types in a traditional manner indicated that the landscape divisions should be adjusted to our research questions. Landscape regions were thereupon subsumed in fewer and larger groups according to two criteria: basic similarities of the regions and homogeneous find distribution. One example is the Mountain area with low find density in the different landscape regions. It is highly probable that prehistoric people perceived these regions as having the same affordances. Another example is the Oslo fjord which includes regions with high find density and affected by sea level changes during the Stone Age. In this way, the number of sub-regions is reduced from seventeen to seven; Oslo fjord area, Lake and Silurian areas, Clay soils, Lower valleys, Higher valleys, Woodlands, and Mountain areas (Fig. 3). In this revised landscape context, artefacts with a geographical accuracy of cadastre or better can be included. This ensures coherence in the choice of landscape categorization, the geographical accuracy levels of the provenance, and the map scale. Both geographical and temporal scale must be chosen in relation to the affordances of the material. This will be addressed in the next section.

2.3. Time

The presentation of prehistoric landscapes also involves the time dimension. The time scale chosen should make it possible to visualize processes and developments in the landscape use pattern. The scale of explanation must relate to the scale of observation (Holdaway, 2006). South-East Norwegian Late Mesolithic, Neolithic and Early Bronze

Age are the limits in time and space for this paper, and the time scale within this is chosen in accordance with the date ranges of the archaeological material and an aggregation that will provide new insight. The large stone artefacts in the museum collection are mostly stray finds that can only be dated typologically. Among several types of large stone artefacts, Late Mesolithic Nøstvet axes, and Neolithic–Early Bronze Age axes, sickles and daggers are selected. A total of 1963 entries in the database of these artefact types with sufficient geographical accuracy are included.

Dating of Mesolithic sites in South-Eastern Norway is closely connected to the eustatic land rise. Sites are generally assumed to have been on or near the coast, and subsequently the dating of land rise will date the site. This idea is generally corroborated by C14-datings from excavations. In this way, the time axis has a pendant in the geographical z-axis, the height above the present day sea level. The Late Mesolithic Nøstvet phase is generally dated to 6300–4400 BC (Glørstad, 2010:35-36). Early in this period, 5900 BC, the sea level was around 40 m higher than today in the Southwest, and about 60 m higher in the inner part of the Oslo fjord (op. cit.:41). In the Northeastern part of the Oslo fjord, the sea level dropped from 70–66 m to 42 m above the present sea level during the Nøstvet phase (Berg, 1997:22).

The Norwegian Neolithic typology is to a large part based on Danish studies. Denmark has the same artefact types and in far higher numbers. The Norwegian Neolithic can be dated to the period 3800–1800 BC. The different artefact types, axes, sickles and daggers, have rather wide date ranges. The time scale should be large enough for patterns to emerge, or in other words, the time slice used to aggregate finds on the time axis should be large enough to give an interesting view of the connection between landscape and artefact type. The time slice should reflect the dated artefacts and a continuous landscape use. Artefacts with a narrower date range will therefore be joined with artefacts with a wider date range to create a dataset in a proper size and time scale. The aggregation reflects a perceived change in the archaeological material. The small group of Early Neolithic faceted battle axes which are grouped with other axes from the Early Neolithic and Middle Neolithic A. The Late Neolithic and Early Bronze Age is treated as one entity in the analyses because flint daggers of type VI (Lomborg, 1973) and some types of the simple shafthole axes can be dated as late as Early Bronze Age. The decision to aggregate in this way can obscure patterns that could be visible at a more fine-grained level, but it gives a resolution appropriate for the archaeological material and the geographical area covered by the analyses. The next step is to investigate different combinations of these elements and explore possible patterns.

2.4. Exploratory Data Analyses (EDA)

The availability of larger datasets presents new opportunities for Exploratory Data Analyses (EDA). EDA (Andrienko *et al.*, 2006) is to do several types of analyses and to present them with different types of visualizations. It is pattern enhancement out of various and vast data volumes, but not necessarily in relation to maps. It can also be used with statistics, presentations of tables and diagrams. However, with the easy availability of GIS, it is good to use maps as a major part of the data presentation. Andrienko et al. refers to Salichtchev's concept of the Cartographic research method (op. cit.: 169). The Cartographic research method is to apply maps for the scientific description, analysis, and comprehension of phenomena. A vital point is to explore the map instead of exploring reality, and the map becomes in this way both an instrument and a subject of investigation. Given the nature of archaeological material, and the interest in the relation between prehistoric societies and landscapes, EDA in a map context combined with the MUSIT database can provide new understanding.

The EDA is an iterative process, where initial observations will give ideas to new constellations that again can lead to new types of analyses. Salichtchev's representation of the method presents an iteration of preparing collected data for mapping, and subsequently to use information from the map to extend the information and enhance the comprehension of the real world. In the process of mapping, unnecessary information is excluded, and it can be possible to acquire new knowledge through information processing, inductive and deductive reasoning. The map construction should be done not only as illustration but in a way that increases understanding. That will make it possible to obtain new information from the map through map reading, analysis and information processing. The interpretation of this new information will in turn lead to better comprehension of the real world (Andrienko *et al.,* 2006:169–170).

There has been an enormous development in the use of GIS in archaeology (Scollar, 1999). By the mid-nineties there where already a high number of GIS papers at CAA-conferences. At later CAA-conferences, GIS has become widely acknowledged analytical tool.

Archaeological excavations conducted by the Norwegian university museums have developed steadily towards digital documentation. The museums have now decided on a common documentation standard and data system. The excavation results shall be archived digitally within the MUSIT system. All reports and articles presenting results have for several years contained maps prepared in a GIS. Nevertheless, the need to go beyond pretty maps is still present. GIS could and should not be used only for presenting results, but as a part of the research process, and preferably for analyses during an excavation and not just in the final publication. It is, however, not right to reduce the presentation of an archaeological material on a map to just a pretty picture. There is new insight to be gained when a find distribution is presented in a new way. What really can create new understanding is the iterative process of putting known facts together on the map, to try new constellations and combinations at different scales. In this way, EDA with maps can challenge the traditional map construction.

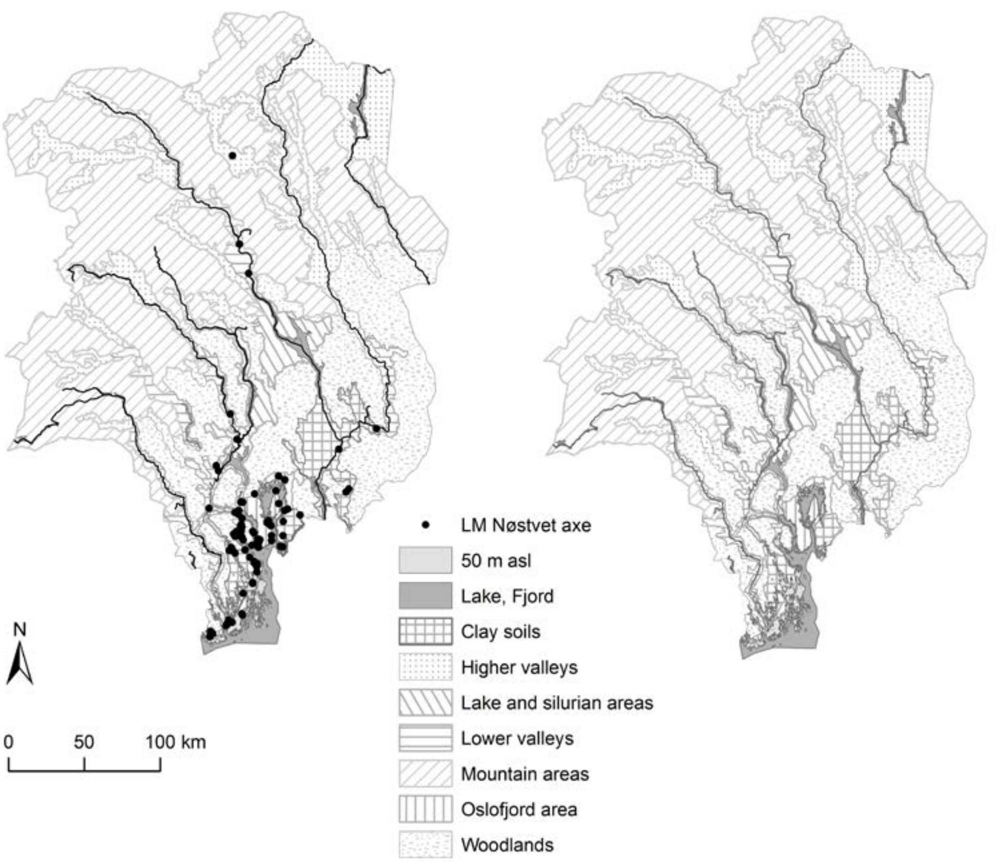


Figure 3: Late Mesolithic Nøstvet axes, grouped by height above sea level.

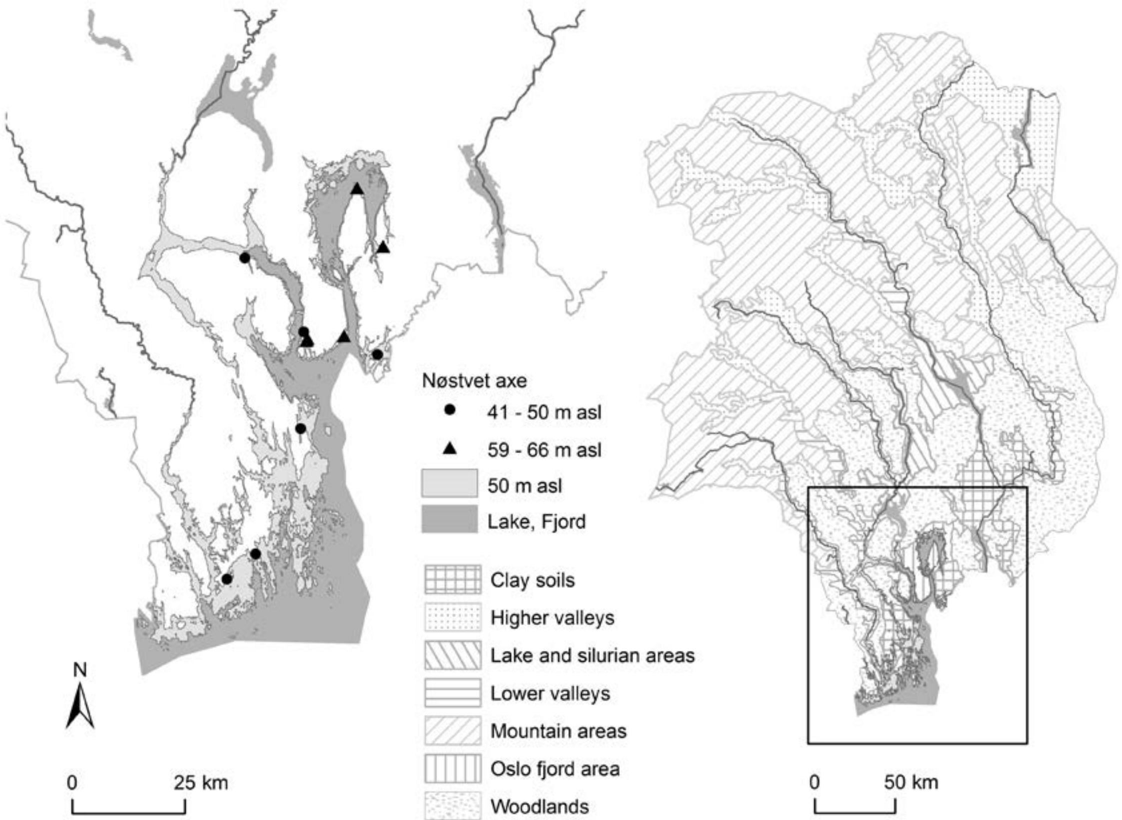


Figure 4: Late Mesolithic Nøstvet axes, sea level 50 m asl

EDA is to present the data in a way that visualizes the totality of information to the viewer. The information should be appropriately represented, so that the required information is perceivable from the display items. This does, however, not mean to present raw data without manipulations. The analyst should use existing knowledge to divide the data into subsets reflecting substantial differences in the behaviour. This use of domain knowledge is important to achieve a good visualization (Andrienko *et al.*, 2006:579–583). Domain knowledge is what makes it possible to choose good levels of abstraction and aggregation and to group and sort facts in a meaningful way in the presentation. This may imply that the principle of totality must be abandoned for the sake of clarity. The seven landscape categories and the three time slices chosen in this paper is an example of aggregation based on domain knowledge.

To realize that domain knowledge is significant makes us aware that there is no such thing as an objective presentation of archaeological material. Post-processual archaeology has clearly stated that all choice and description of archaeological material is in itself the result of a series of assumptions and subjective decisions. Not that there are no facts, but the selection and subsequent presentation in visualizations is a subjective process.

3. Visualizations

The visualizations in figs 3–4, 6-7 present different aspects of the Norwegian Stone Age. In one way, the presentation is rather traditional – it combines a find distribution with landscape variables. The background is, however, not a series of geological and vegetational single variables, but a three dimensional classification of landscape regions based on a multidisciplinary understanding and holistic evaluation. This is a better starting point for archaeological analysis than a traditional scientific description of landscape. For the sake of clarity, archaeological domain knowledge is used to reduce the number of landscape regions.

The typologically dated Stone Age artefacts chosen for this work consist of nearly 2000 database entries. The accuracy of the spatial and temporal information varies from exact sites coordinates to the quite coarse county level. The accuracy of dating are intervals ranging from 400 to 1100 years. The immediate impression of the sites plotted on the map is that the majority of finds are in areas close to the Oslo fjord; the South-Eastern part contains most of the finds. The artefacts mainly concentrate in the regions Oslo fjord, Lake and Silurian areas, Lower valleys, and Clay soils.

Coordinate accuracy	Site	Cadastral unit
n	39	209
Up to 70 m asl	31	156
Up to 50 m asl	8	93
Up to 30 m asl	6	77

FIGURE 5: COORDINATE ACCURACY AND SITES ACCORDING TO SEA LEVEL.

The late Mesolithic find distribution (Fig. 3), as presented through the Nøstvet axes, shows a clear affinity to the coastal landscape of the period. There are 39 entries in the database with site accuracy. 22 are found between 70 and 50 m asl 25 between 70 and 30 m asl. This agrees well with the range of sea levels within this period. It is also possible to group the sites in two categories, with the intervals 41–50 m asl and 59–66 m asl (Fig. 5).

The first category has a Southern and Western distribution while the other are found in the Northern and Eastern parts. This coincides well with the sea level gradient with the sea level early in the Nøstvet phase, 5900 BC, ranging from 40 m asl in the Southwest to 60 m asl in the inner part of the Oslo fjord (Glørstad, 2010:41).

The distribution of Nøstvet axes can also illustrate the point of including only sites with a high accuracy level in this analysis. When all Nøstvet axes with an accuracy of cadastral unit or better is included, the total number increases to 209. The overall distribution pattern remains the same, but number of axes below 30 m is higher, and the relative difference between the groups is not as clear as when only the axes with site accuracy is included (Fig.5).

The shore line displacement curve is not given the same importance in Neolithic as in Mesolithic studies. This can be ascribed to two reasons. First, the landscape change is not so dramatic in this period as in the preceding Mesolithic. Nevertheless, at the end of the Mesolithic, the shore line in the inner part of the Oslo fjord is still at a level of 38 m above the present shoreline (Berg, 1997). Accordingly, the shoreline at 30 m above the present is chosen as an approximation for the whole area at the beginning of the Neolithic period. Second, the change from Mesolithic to Neolithic generally implies a change of subsistence. The Neolithic mode of subsistence is by definition agricultural. It has therefore been closer at hand to discuss the sites in relation to fertile soils than to the coastal environment. Consequently, it has not been the same tendency to date Neolithic sites in relation to height above sea level.

The distribution of axes from the Early Neolithic and the first part of the Middle Neolithic (EN and MNA) (Fig. 6) shows a continued preference for areas close to the fjords. At this scale level, the finds with an accuracy of cadastral unit or better is included. Compared to the distribution of Nøstvet axes, one can also observe a general increase in finds, especially in the Southwestern part of the Clay soils region. There are also finds from more interior areas from this period. The number of finds in the interior is however small, especially considering the time span of 1100 years that the EN/MNA covers.

The distribution of axes, sickles and daggers from the Late Neolithic and Early Bronze Age shows a further expanding and intensified use of landscapes. The Oslo fjord and Lower valleys have many sites in preceding periods, so the increase

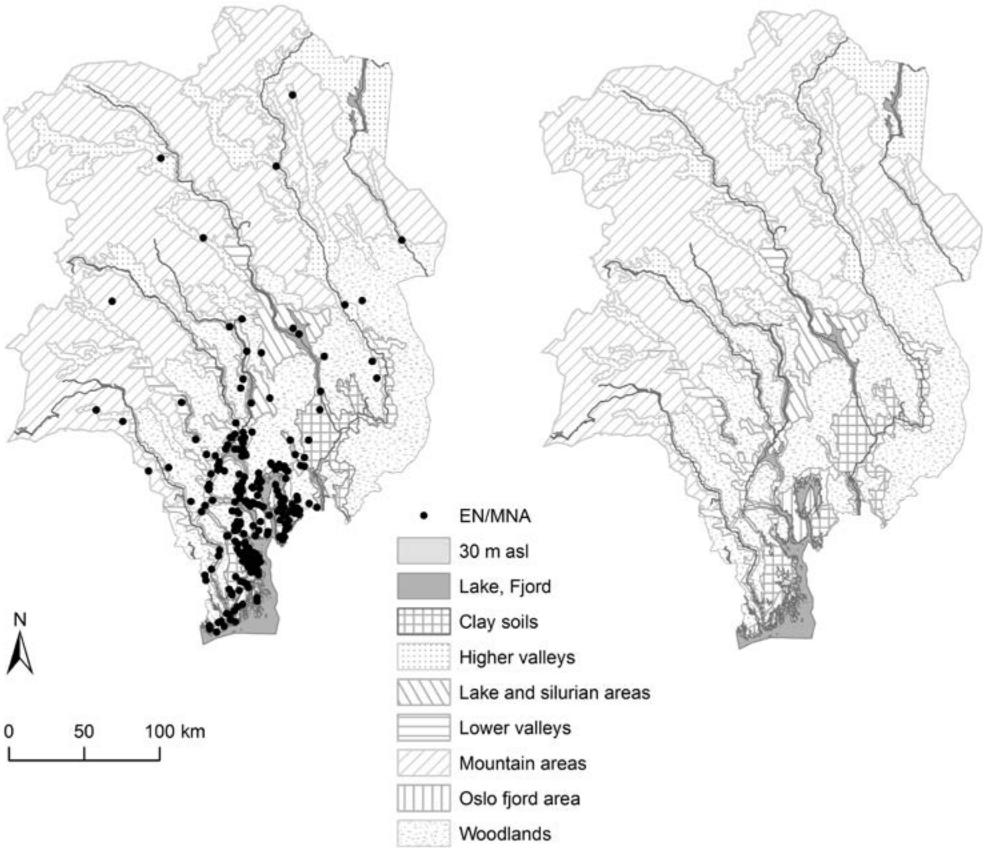


Figure 6: Axes from Early Neolithic and Middle Neolithic A, sea level 30 m asl.

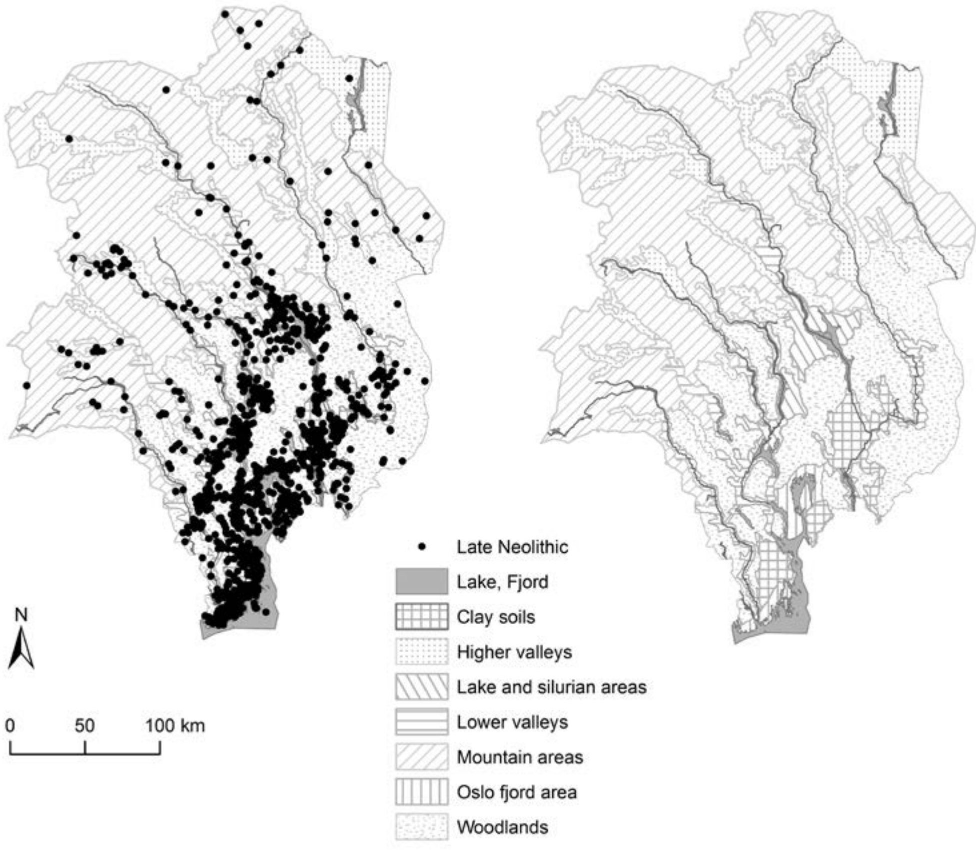


Figure 7: Axes, sickles and daggers from Late Neolithic and Early Bronze Age.

in the number of sites is relatively higher in the Lake and silurian as well as the Clay soils areas. There is also a marked increase in the distribution of finds in the interior.

The presentation of Neolithic finds could use the cadastral level for all finds. The eustatic land rise in South-Eastern Norway after the Ice Age gives a different situation for the Mesolithic finds. The earliest coastal sites in this part of Norway are now at more than 100 m asl so that many Mesolithic coastal sites are now inland sites. The impression of artefact distribution in the landscape changes when the visualization includes an old shore line level, and the analysis should rely on objects with site accuracy. The land rise is similarly important in the understanding of Early Neolithic sites. Although the difference is smaller, the result of a shore line at 30 m asl gives a different understanding of the Early Neolithic site distribution. The maps that present the relation between sites and landscapes have included sites with an accuracy level of cadastral units and better.

4. Discussion

The possibility of easy access to information from archaeological museum collections opens up new venues for wider research. Representativity and how the material has been collected will always be important questions to address when using existing sources. The collections of the Norwegian university museums have a reasonably good chance of being representative, since the museums can legally claim all archaeological finds, and they conduct all excavations of prehistoric sites. The museums also have a tradition dating back to 1866 of publishing annual overviews of the acquisitions to the collections. The net publishing can be seen as a continuation of this tradition through new media. Most of the items are presented with the artefact terminology used when they were first described at the museum. The event based database makes it possible to update information and at the same time preserve the old entries. The total material would be even more readily available if information was regularly updated when a group of objects was analysed by a researcher. In the project Dynamic Distributions, the MUSIT database is updated because it is used for the analyses.

Stone artefacts like axes, sickles and daggers are for the most part stray finds. The information about the find context can in some cases be sufficient to categorize them as finds from depots, sacrifices or graves. It is reasonable to believe that these artefacts are a kind of inalienable possessions that materialize ideas and insures stability and security in everyday life. Because of their importance, there have presumably been rules for their production, use and disposition (Reitan, 2009:46). These large Stone Age artefacts can therefore represent important aspects of life in the Neolithic period. Although most of them are stray finds, their deposition indicate the structuring of the cultural landscape (Lekberg, 2002:29ff.). This paper presents Stone Age artefact distribution at a coarse scale level and in relation to landscape types and height above sea level. Another possible approach is to discuss them in relation to water catchment areas (von Hackwitz, 2012:61–63). Major lakes and waterways have been included in the visualization, because they give a possible explanation to single finds. Earlier studies of artefact distribution supports the notion of different cultures in different landscape types, independent of major watersheds (Amundsen, 2011).

The transformation from the Mesolithic to the Neolithic has customarily been described quite dramatically, with terms like the Neolithic revolution, indicating an abrupt change in the subsistence economy and landscape use. Norway has been on the outskirts of the Neolithization. There is hardly any evidence of Early Neolithic agriculture in pollen diagrams, and there are no bones from domesticated animals before the early Middle Neolithic (Prescott, 2009:198, Solheim, 2012:60–61).

The present point of view is that there has been a gradual change, from the Mesolithic period onwards, and groups of people living in present day South-Eastern Norway have been in contact with people further south and exchanged ideas. The contact might have been in the form of a series of short distance travels along the coast (Glørstad, 2012:8–9). This social contact and exchange of ideas led to the introduction of artefact types to Norway, a direct influence from further south. The social contacts were established and parts of the material culture changed, but this did not lead to simultaneous changes in the subsistence pattern. The economic basis changed much more slowly, with a much more gradual change from a Mesolithic to a Neolithic economy (Reitan, 2009:41–46, Solheim, 2012:62–67). The visualization presented here indicates the same development. There are no dramatic changes in the distribution pattern from Late Mesolithic Nøstvet phase to the Early Neolithic/Middle Neolithic A. When the sites are seen in relation to a sea level 30 m above the present shore line, a preference for the areas close to the fjord is even more evident. There are, however, some indications of a higher preference for areas further from the coast (Østmo *et al.,* 2006), but it is mainly a continuity from the Mesolithic through the beginning of the Neolithic.

The beginning of the Late Neolithic is seen as the start of a shift in economy and social system in Norway (Prescott, 2009). The map of axes, sickles and daggers from the Late Neolithic (Fig. 7) shows intensification and expansion further inland. There is a strong increase in the use of all regions, but the most marked is in the Lake and Silurian areas as well as the Clay soils region. These regions are well suited for agriculture and pastoralism, and the observation supports a shift in subsistence. The shift in site density is not easily seen at this scale, but could be presented in more detailed analyses of the area.

The scale levels show broad developments during this part of the Stone Age. In further work we will extend the study area to all of South and East Norway. The material collected here is also a good starting point for more detailed analysis. Different subsets can be selected and used in studies at varying scales with the appropriate selections of geographical accuracy and time scales.

References

AMUNDSEN, H. R. (2011). *Mot de store kulturtradisjonene – Endringsprosesser fra tidligneolitikum til førromersk jernalder mellom Mjøsa og Femunden*. Ph.D. dissertation. Oslo: IAKH, Det humanistiske fakultet, Universitetet i Oslo.

ANDRIENKO, N. and ANDRIENKO, G. (2006). *Exploratory Analysis of Spatial and Temporal Data*. Berlin, Heidelberg: Springer Verlag.

BERG, E. (1997) *Mesolittiske boplasser ved årungen i Ås og Frogn, Akershus. Dobbeltspor/E6-prosjektet 1996*. Varia 44. Oslo : Universitetets Oldsaksamling.

BERG, E. (2007). Using a GIS-based Database as a Platform for Cultural Heritage Management of Sites and Monuments in Norway. In: Clark, J. T, Hagemeister, E. M. (eds.). *Digital Discovery. Exploring New Frontiers in Human Heritage. CAA2006. Computer Applications and Quantitative Methods in Archaeology. Proceedings of the 34th Conference, Fargo, United States, April 2006*. Budapest: Archaeolingua. pp. 345-351 (CD-ROM).

GLØRSTAD, H. (2010). *The structure and history of the late Mesolithic societies in the Oslo fjord area 6300–3800 BC*. Mölndal: Bricoleur Press.

GLØRSTAD, H. (2012). Traktbegerkulturen, kysten og det tidligste jordbruket – et problem for periferien? In: Solberg, A., Stålesen, J. A. and Prescott, C. (eds.). *Neolitikum. Nye resultater fra forskning og forvaltning*. Nicolay skrifter. 4. Oslo: Nicolay arkeologisk tidsskrift. pp. 7-18.

VON HACKWITZ, K. (2012) The Creation of Regions: An Alternative Approach to Swedish Middle Neolithic Boundaries and Cultures. *Norwegian Archaeological Review* 45 (1). pp. 52-75.

HOLDAWAY, S. J. and WANDSNIDER, L. (2006) Temporal Scales and Archaeological Landscapes from the Eastern Desert of Australia and Intermontane North America. In: Lock, G. and Molyneaux, B. L. (eds.). *Confronting Scale in Archaeology. Issues of Theory and Practice*. New York: Springer Verlag.

HOLMEN, J. and ULEBERG, E. (1996) The National Documentation Project of Norway – the Archaeological Sub-Project. In: Kamermans, H. and Fennema, K. (eds). *Interfacing the Past. Computer Applications and Quantitative Methods in Archaeology CAA95*. Vol. I. (Analecta Praehistorica Leidensia 28). Leiden: Institute of Prehistory, University of Leiden. pp. 43-46.

JORDAL, E., ULEBERG, E. and HAUGE, B. (2012). Was It Worth It? Experiences with a CIDOC CRM-based Database. In: Zhou, M., Romanowska, I., Wu, Z. Xu, P. And Verhagen, P. (eds.). *Revive the Past. Computer Applications and Quantitative Methods in Archaeology (CAA). Proceedings of the 39th International Conference, Beijing, April 12-16*. Amsterdam: Pallas Publications. pp. 255-260.

KVAMME, K. L. (1989) Geographic Information Systems in Regional Archaeological Research and Data Management. In: Schiffer, M. B. (ed.). (*Archaeological Method and Theory* 1. Tucson: University of Arizona Press. pp. 139-202.

LEKBERG, P. (2002). *Yxors liv, människors landskap: en studie av kulturlandskap och samhälle i Mellansveriges senneolitikum*. Uppsala: Department of Archaeology and Ancient history, Uppsala University.

LOMBORG, E. (1973). *Die Flintdolche Dänemarks: Studien über Chronologie und Kulturbeziehungen des südskandinavischen Spätneolithikums*. In: Nordiske Fortidsminder. Serie B; bd 1. Copenhagen: Universitetsforlaget.

MATSUMOTO, M. and ULEBERG, E. (forthcoming). Scale dependent patterns in South-Eastern Norway. Traviglia, A. (ed.). *Across Space and Time, Proceedings of the 41st Computer Applications and Quantitative Methods in Archaeology Conference CAA 2013 Perth, Australia*. Amsterdam: Amsterdam University Press.

.PRESCOTT, C. (2009). History in prehistory – the later Neolithic/Early Metal Age, Norway. In: GLØRSTAD, H. and PRESCOTT, C. (eds.) *Neolithisation as if history mattered*. Uddevalla: Bricoleur Press.

PUSCHMANN, O. (1998) *The Norwegian landscape reference system – use of different sources as a base to describe landscape regions*. NIJOS report 12/98. Ås: NIJOS.

PUSCHMANN, O. (2005). *Nasjonalt referansesystem for landskap. Beskrivelse av Norges 45 landskapsregioner*. NIJOS rapport 10/2005. Ås: NIJOS.

REITAN, G. (2009). Offer det? *Primitive tider* 11. pp. 31-50.

SCOLLAR, I. (1999). 25 Years of Computer Applications in Archaeology. In: DINGWALL, L., EXON, S., GAFFNEY, V. LAFLIN, S. and van LEUSEN, M. (eds.). *Archaeology in the Age of the Internet. CAA97. Computer Applications and Quantitative Methods in Archaeology. Proceedings of the 25th Anniversary Conference, University of Birmingham, April 1997* (BAR International Series 750). Oxford: Archaeopress. pp. 5-10.

SOLHEIM, S. (2012). En historie om møter og motsetninger. Introduksjonen av jordbruk til Øst-Norge. *Viking*. LXXV. pp. 53-74.

SØRENSEN, R. (1979). Late Weichselian deglaciation in the Oslofjord area. *Boreas*. 8. pp. 241-246.

ØSTMO, E and SKOGSTRAND, L (2006). Nye funn av traktbegerkeramikk ved Oslofjorden. Børsebakke og Vøyenenga. *Viking*. LXIX. pp. 69-90.

Proceedings of the 42nd Annual Conference on Computer Applications and Quantitative Methods in Archaeology
CAA 2014 - 21st Century Archaeology/F. Giligny, F. Djindjian, L. Costa, P. Moscati, S. Robert (eds.)

An Inventory of Lucanian Heritage

Alain DUPLOUY,[1] Vincenzo CAPOZZOLI,[1] Alessia ZAMBON[2]

[1] Université Paris 1 Panthéon-Sorbonne, Paris, France
[2] Université de Versailles Saint-Quentin-en-Yvelines, France
alain.duplouy@univ-paris1.fr, vincenzo.capozzoli@univ-paris1.fr, alessia.zambon@uvsq.fr

Abstract

The project 'Ancient Lucania: archaeology and national heritage' aims at investigating the process of Lucanian ethnogenesis using, beside texts, an archaeological approach. In order to achieve this project, we are compiling both a 'Gazetteer of archaeological sites of ancient Lucania' and an 'Inventory of Lucanian antiquities'. To manage and organize our data, we needed to develop a relational database providing a simple and efficient end-user experience, as well as an online collaborative work platform meeting the needs of several contemporary accesses. These requirements had to be associated with a low-cost development process demanding no specific programming skills and sustainable ensuing maintenance costs. Accordingly, we opted for a proprietary database management system. FileMaker products swiftly appeared as both practical and sustainable solutions. This paper presents the technical choices made to conceive a relational database, to implement its structure in FileMaker and to add mapping tools in the resulting Ancient Lucania Database.

Keywords: Relational database, proprietary DBMS, FileMaker, mapping tools, ancient Lucani

In the last decade, the Chair of Greek archaeology of the University Paris 1 Panthéon-Sorbonne has acquired a growing expertise in developing web-related GIS and databases in fieldwork archaeology. More specifically, within the context of the international Itanos survey project, we developed an online database aimed at solving the typical problems of a collaborative work involving many archaeologists scattered around the world (Costa et al. 2008, Duplouy et al. 2009). Benefiting from this experience, our research team is now involved in a new project related to 'Ancient Lucania: archaeology and national heritage', which is funded by the City of Paris through the 2012-2015 Emergence(s) scheme. For more than four decades, archaeologists of Paris 1 have been involved in various Lucanian archaeological fieldworks, exploring most of the Lucanian territory and a diversity of historical, archaeological and landscape situations.

1. Lucania and Lucanians: an archaeological perspective

Lucania is a region of South Italy, which covers modern southern Campania, the most part of Basilicata and northern Calabria (Fig. 1). It is the home of various pre-Roman civilizations known through ancient authors. In the archaic period (late eighth to early fifth century BC), the region was inhabited by various indigenous tribes known as Oenotrians, Chones or Serdaioi, whereas Greek colonists founded several cities in the coastal plains (Poseidonia, Eleia, Laos, Sybaris, Siris and Metapontum). Towards the mid-fifth century BC, the region seems to have been progressively occupied by newcomers. According to ancient authors, Lucanians, an Oscan-speaking people related to the Samnites, entered the area from the North. Despite the fierce opposition of the principal Greek cities, the Lucanians soon occupied the whole region and ruled over it. However, they subsequently suffered by choosing the losing side in the various wars on the peninsula in which Rome took part, being brutally punished by the Romans and eventually reduced to subjection.

The transition between the archaic indigenous and the Lucanian phases of southern Italy has been explained in two very different ways. On the one hand, ancient authors

Figure 1: Map of ancient Lucania (© Lucanie antique, Université Paris 1 Panthéon-Sorbonne).

and various modern historians subscribed to the idea of a migration of Samnite tribes coming from the North and expanding into Lucania. On the other hand, however, many archaeologists favour the idea of an internal evolution, named 'Lucanian ethnogenesis', and imagine the maturation of a Lucanian identity within the local populations of archaic Lucania (e.g. Pontrandolfo 1982). The literary model of Samnite expansionism is actually never completely dismissed and is regularly reintroduced, consciously or not, in archaeological interpretation, which has often proven to be unable to leave the choice between Samnite expansionism and indigenous internal evolution.

Beyond ancient texts, which insist on feuds between Greeks and Lucanians, material culture remains the most significant documentation available on pre-Roman Lucania and offers the best prospect of gaining an understanding of the full complexity of local situations in its chronological depth, without reducing it to ethnicity matter. Instead of opposing Greeks and indigenous populations in the historical pattern of southern Italy, all these people must actually be considered in the wider network of Mediterranean history (Horden & Purcell 2000, Morris 2005). As recent literature tends to show, patterns of evolution during the whole pre-Roman period are related to multiple transformations in the material culture that are not necessarily ethnically bounded, but are also typical of a global Mediterranean evolution in which Lucania, as a region located at the crossroads of multiple cultural influences, took part.

Labelling cultures is not an end in itself. But characterising them is a necessity. To be specific, archaic indigenous material cultures and Lucanian as well are actually made of specific sets of attributes, but also marked by significant internal variations (see Morgan 1999). In order to assess the process of 'Lucanian ethnogenesis' during the fifth and fourth century BC within the wider spectre of Mediterranean context while taking into account the diversity of local patterns and the various substrata of archaic Lucania, it could be worth using a polythetic model. As formalized by David Clarke (1968: 668), the polythetic model offers an interesting tool to explain archaeological diversity by allowing variations in the definition of a specific culture. The point is not to define a set of distinctive attributes that would be ethnically relevant, but rather to explain the variability of that material culture.

2. Ancient Lucania: archaeology and national heritage

This is not the place to present or discuss all these historical questions, and we shall focus on the instruments built in order to conduct the research. We decided to gather all pieces of archaeological evidence and to undertake the collation of a 'Gazetteer of archaeological sites of pre-Roman Lucania'. There is plenty of archaeological data available on ancient Lucania. The first major discoveries date back to eighteenth-century antiquarians, but the creation of the Soprintendenza della Basilicata in 1964 and its establishment in Potenza marked the beginning of systematic archaeological researches. Archaeological activity was intense in Basilicata in the last decades, with plentiful operations of rescue and preventive archaeology beside planned fieldworks. Half a century of activity has also produced a profusion of reports, which are unfortunately dispersed in local periodicals, museum catalogues and regional conference proceedings that are not always easily available outside of Italy. One of the aims of the project is to gather in a comprehensive database all pieces of information related to archaeological sites and material excavated in the territory covered by pre-Roman Lucania. The objective is both to provide a convenient access to scattered information, but also to take a critical stance, particularly on earlier researches, that takes into account recent discoveries and interpretative models. By the end of the project, the Ancient Lucania Database will eventually contain a large number of evidences, making it easily available for analytical purposes and archaeological synthesis.

Another aspect of the research programme is linked to the Lucanian heritage in Paris. Lucanian heritage is intended here as consisting of both Lucanian objects, that is objects made in Lucania or issued from a Lucanian workshop whatever the place of their discovery, and objects discovered in Lucania whatever their actual fabric. Although rather heterogeneous, the category of Lucanian heritage offers the possibility to study both Lucanian products (including their possible diffusion outside Lucania) and Lucanian archaeological contexts (including imported ware). Since the City of Paris funds the project, the notion of Lucanian heritage has initially been restricted to the Parisian component of this worldwide heritage. During the eighteenth and nineteenth century, as part of their Grand tour, numerous young foreigners reached the city of Naples, and some dared to venture further South to Calabria and Basilicata, which were less easily accessible (Settembrino & Strazza 2004). A handful of Frenchmen are among the adventurous who pushed their journey to the remote territories of Lucania. Testimonies collected by these travellers (diaries, books, drawings) recount their explorations and discoveries, and highlight the importance of French antiquarian tradition on Lucania. Many objects found during their excavations or acquired by these men have been brought back to France and now adorn the Parisian museums, such as the architectural terra-cottas excavated in Metaponto by the Duc de Luynes in 1828 and donated to the Cabinet des médailles in 1862 (fig. 2). Others were subsequently acquired on the art market. We will attempt to document the full history of these objects, from their present location and as far as possible in the history of collections back to their place of discovery. In association with curators, we are therefore collating an 'Inventory of Lucanian antiquities', paying attention to the way these pieces arrived in Parisian Museum.

3. Choosing a solution: FileMaker as a DBMS in the cloud

Providing a simple and efficient end-user experience is an essential aspect of the project. The Ancient Lucania Database should offer a user-friendly interface for disseminating the results of the research programme, and require no special skills in building complex search queries.

Figure 2: Architectural terra-cotta with lion-head spout (Paris, Cabinet des médailles, Luynes 780) (© Lucanie antique, Université Paris 1 Panthéon-Sorbonne).

It should also provide a collaborative work platform that meets the needs of several contemporary accesses on the same set of data, both in capturing, analysing and searching the data. These requirements had to be associated with a low-cost development process, demanding no specific programming skills, and sustainable ensuing maintenance costs as well. In 2006 indeed, the creation of the already mentioned Itanos Survey Database (http://prospection-itanos.efa.gr) with open-source technologies – typically an Apache-MySQL-PHP system – required specific development by a computer engineer, and its maintenance over the years has proven to be costly and discontinuous. Capitalizing on this previous experience, it has been decided not to replicate this solution. Actually, open source software is in no way free for archaeologists who want to focus on the archaeological part of their job without turning into computer programmers… even if they are 'geeks'.

Considering these prerequisites, we opted for a proprietary database management system (DBMS), and FileMaker products swiftly appeared as both practical and sustainable solutions. Although an expensive alternative to open source software from the sole perspective of the purchase of the licences, this solution actually complies with all the requirements of the project. Building the Ancient Lucania Database with FileMaker offered the following advantages: (a) simple structuring of our dataset with a minimum of technical knowledge, (b) easy access by non-specialists, and (c) instant sharing over the internet by multiple users.

a) FileMaker offers quite an easy, though powerful, implementation of relational datasets. Since its initial launch, FileMaker's relational model has been considerably enriched. Establishing relations throughout the dataset has been built at the core of the managing capacity of the software, allowing complex relations to be created with a minimum of technical knowledge. Despite some peculiarities of the FileMaker technology as a DBMS in managing tables and relationships (see below), FileMaker's relational model offers multiple tables per file and a graphical relationship editor that displays and allows manipulation of related tables in a manner that resembles canonical entity-relationship modelling.

b) As a subsidiary of Apple Inc., FileMaker Inc. has always been careful to associate his database engine with a graphical user interface (GUI) compliant with the expectations of non-specialist users. In developing the Ancient Lucania Database we were very keen on offering the same user-friendly experience to end-users, immediately providing them all relevant information. That is to say that user requirements and areas of interest were considered right from the start in the design of the tool against an archaeological background interest. There are indeed plenty of queries that can be performed against the whole collected set of data, but only some of them are really relevant to an archaeological research on ancient Lucania. In other words, visualization of relevant data has been prioritized against the whole series of possible interactions with the dataset. Instead of performing any kind of on-demand queries with more or less complex search criteria, FileMaker offers these oriented visualisation capabilities, turning end-users toward specific areas of research by referring to previously created relations in the relationship graph (RG). If FileMaker doesn't have the flexibility of SQL – even if plugins and recent implementations in the latest versions of the software have brought a full set of SQL capabilities –, it is actually built on a different concept.

c) A database is useless if it cannot be accessed widely and easily. The Ancient Lucania Database had to provide a collaborative work platform that meets the needs of end-users for capturing and searching data. Simultaneously accessing the most up-to-date information could only be achieved through a centrally hosted database on a server. Cloud technologies therefore appeared as a necessity. Accordingly, our choice to use FileMaker has also been determined by the need to interact with other researchers and databases through the FileMaker Server capabilities. In this perspective we benefitted from the technological infrastructure offered by the Huma-num Very Large Facility, which has been established to facilitate the digital turn in humanities and social sciences among French research operators. Huma-num provides the capabilities of FileMaker Server at no cost for our team.

Considering their potentialities and knowing their restrictions, FileMaker products nevertheless appeared as practical and viable solutions to be used as a DBMS in the cloud for the Ancient Lucania Database. We are fully aware that entrusting our data to propriety database systems can be risky, as it already happened that database companies went under, leaving scholars with datasets difficult to access on new computers that couldn't run the older database software. However, although the whole set of data and the structure of the Ancient Lucania Database will be enclosed in a proprietary file system, the formal organisation of a FileMaker database can still be transferred, with various adaptations, to open source solutions in the possible (but still unlikely) event of a breakdown of FileMaker Inc. and the discontinuity of its development.

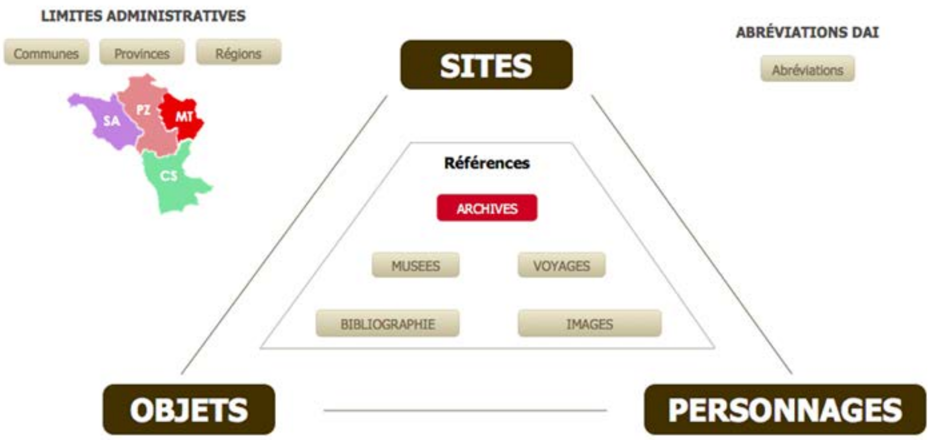


Figure 3: ER model of the Ancient Lucania Database (© Lucanie antique, Université Paris 1 Panthéon-Sorbonne).

The Ancient Lucania Database was created using FileMaker Pro 12, with a later upgrade to FileMaker Pro 13, which allows a much broader set of solutions for the graphical user interface and upgraded web publishing capabilities through the WebDirect technology. It has been hosted on a Huma-num FileMaker Server 12 since the summer 2013, and then migrated to a FileMaker Server 13. Since FileMaker natively manages multiple users with specific accounts and privilege sets, data are automatically signed when captured or modified. Beside registered users, a more general audience of scholars interested in ancient Lucania can also fully access the data as simple guests with no writing permissions. The database can be accessed through a FileMaker client on a desktop or laptop, through the FileMaker Go app on an iOS device, or through a simple web browser on any computer through our website (http://lucanie-antique.univ-paris1.fr). FileMaker technologies indeed provide a type of quick, easy and secure web publishing by creating the web form and other codes needed to access the database as a website, with no development skills required. The database accessed via the web uses an interface that mirrors the one used in the desktop app, and changes made through the web are updated to the FileMaker Server-hosted database in real time.

All this allowed us to spare weeks of specific development by a computer engineer. Even archaeologists with a reasonable knowledge of FileMaker are able to modify the structure of the data, the layout of the interface or the way the DBMS works. This could never be achieved with an externalised development process, except with time-consuming backwards and forwards exchanges between archaeologists and engineers.

4. Building the database: from data to end-users

The first task was to describe the collected data and their processing requirements as an elementary entity-relationship model (ER model) (Fig. 3). Considering the ambitions of our programme towards archaeology and cultural heritage, the conceptual data model has been organized around three main objects: the archaeological sites of Lucania, the Lucanian antiquities and the people related either to the exploration of Lucania or to the collecting of Lucanian antiquities (travellers, collectors, art dealers or even famous curators). These entities are characterized by specific attributes and documented by what we may call references (hundreds of bibliographic records, unpublished archive references, administrative boundaries, and images). They are also linked with each other by relationships that express the dependencies and requirements between them. The relational database has been implemented moving from this elementary conceptual model, which has been kept pro memoria as a gateway to the Ancient Lucania Database.

4.1. Defining and naming the tables

Each entity of the conceptual ER model is represented as a table, with data fields being the attributes. The key principle in building a relational database is that one-to-one relations are included as fields in the current table, one-to-many relations need a related table with match (or key) fields, and many-to-many relations require an intermediary table between two related tables. To be specific, to one archaeological site correspond several objects, but every object is linked to only one site. For instance, among the numerous antiquities discovered in the Achaean colony of Metaponto, the proto-classical architectural terra-cottas kept at the Cabinet des médailles in Paris were excavated in the urban sanctuary (the so-called temple of Apollo Lyceum, actually the temple of Hera) and brought to France by the Duc de Luynes. In that case, the relation between a single site and the objects found there is a one-to-many. Conversely, different persons (travellers, curators, art-dealers, etc.) can be related to the same sites. For example, Aubin-Louis Millin, traveling in southern Italy between 1811 and 1813, explored various sites in Campania, Calabria, Basilicata, Apulia and elsewhere (D'Achille et al. 2012a and 2012b), but these sites where also visited by other travellers. In that case, the relations are many-to-many relationships, and intermediary tables are necessary to establish the links between sites and people.

Although there are only three main objects on which the Ancient Lucania research project focuses, 23 tables were actually necessary to implement the conceptual ER model and organize every piece of information and their relations (Fig. 4).

Despite the lack of specific requirements and the general flexibility of FileMaker coding, adopting a consistent

ABR_Abreviations_DAI	List of periodicals and their usual DAI abbreviations
ARC_Archives	Archives on Lucanian antiquities, people or sites
AXO_ArchivesXObjet	Archives related to a specific object
AXP_ArchivesXPerso	Archives related to a specific person
AXS_ArchivesXSite	Archives related to a specific site
BIB_Bibliographie	Bibliographic records on Lucanian antiquities, people or sites
BXO_BiblioXObjet	Bibliographic records related to a specific object
BXP_BiblioXPerso	Bibliographic records related to a specific person
BXS_BiblioXSite	Bibliographic records related to a specific site
COM_Communes_2011	List of Italian municipalities in 2011 (according to ISTAT)
CXV_CommunesXVoyage	Municipalities related to a specific journey
IMA_Images	Images of Lucanian antiquities, people or sites
MUS_Musees	Museums hosting Lucanian antiquities
OBJ_Objets	Lucanian antiquities (i.e. antiquities from Lucania)
OXP_ObjetsXPerso	Lucanian antiquities related to a specific person
PAR_Parametres	List of invariable technical parameters
PER_Personnages	People associated with Lucanian antiquities or sites
PXV_PersosXVoyage	People related to a specific journey
PRO_Provinces_2011	List of Italian provinces in 2011 (according to ISTAT)
REG_Regions_2011	List of Italian regions in 2011 (according to ISTAT)
SIT_Sites	Archaeological sites of Lucania
SXP_SitesXPerso	Sites related to a specific person
VOY_Voyages	Journeys in Lucania

Figure 4: List and definition of the required tables of the Ancient Lucania Database.

naming convention is especially important when working in a team environment. All main tables are thus named with an uppercase three-letter acronym (TLA), followed by underscore and the full descriptive name. Intermediary tables (establishing a relation between two main tables) are named using the first letter of both main-table TLAs joined by an intermediary X indicating the relationship.

4.2. Organizing the relationship graph

Elaborating the relationship graph (RG) is the focal point of a relational database development. Considering various peculiarities of the software as a DBMS, implementing the conceptual ER model of the Ancient Lucania dataset into a FileMaker RG actually implied various adaptations. Of course, FileMaker allows the creation of a relationship between any two tables in the RG, but the relationship must never create a cycle, or closed loop between tables, which is actually the way our ER model was conceived in order to link archaeological sites, objects, and people between them. Because FileMaker does not allow creating cycles within the RG, any attempt to create a cycle requires generating duplicate but uniquely-named table occurrences (TOs). Indeed, the 'entities' on the graph are not tables, but TOs, i.e. instances of each table. Since FileMaker only allows a single relationship between any two TOs, what may seem like duplications of tables on the RG actually ensures that there can be no ambiguity when referring to a TO from the context of another TO. Many developers therefore subscribe to the 'Anchor & Buoys' model in organizing the RG.

The 'Anchor & Buoys' model is based on relationships organized into horizontal groups not touching each other within the RG: the so-called Table Occurrences Groups (TOGs). Each TOG consists of one 'anchor' at the left, serviced with related data via any number of threads of 'buoys' strung off rightward. In the Ancient Lucania Database there are three main TOGs, beginning with the TO of the three main tables (SIT, OBJ, and PER), plus several general reference TOGs beginning with the TO of the various reference tables (ARC,

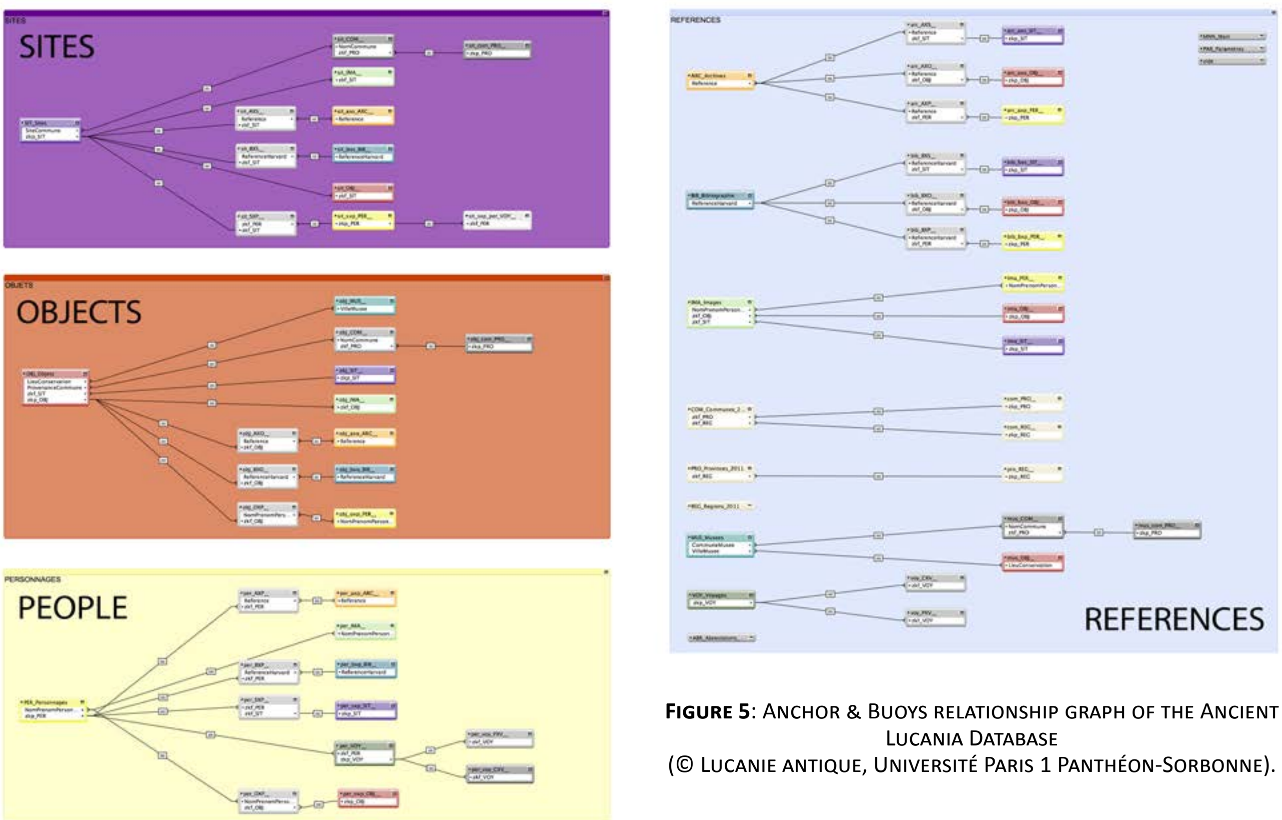


Figure 5: Anchor & Buoys relationship graph of the Ancient Lucania Database
(© Lucanie antique, Université Paris 1 Panthéon-Sorbonne).

BIB, IMA, COM, PRO, REG, MUS, VOY). The RG (fig. 5) has been graphically designed in order to vertically align the different 'anchors' as well as the 'buoys' (including intermediary and related tables).

Of course, the 'Anchor & Buoys' model is not without various restraints. Such a hierarchical model lends itself indeed to the requirement of only one path between any two TOs. It means that the ensuing RG does not support bi-directional relations inherently, so that expressing right-to-left data requires creating another TOG with another 'anchor' as initial point. The necessity to create the same relationship in different TOGs to replicate the same functionality in different contexts nevertheless ensures every relationship is created from the context of each specific TO. To some degree, the 'Anchor & Buoys' model was developed for the FileMaker developers' convenience, in order to ease the randomness and difficulties of finding one's way around and of knowing one's context. Being formally codified, such a method therefore also ensures that all archaeologists ever implied in (future) developments of the Ancient Lucania Database will find their way out, even if the RG cannot be automatically replicated in an open source DBMS. Whereas the conceptual ER model is the main structuring expression of all dependencies and requirements between the various entities of the database, the RG is actually nothing more than a visual tool helping the developer to execute the relationships that are needed to express the required links between the various parts of the dataset and that are offered to end-users queries.

4.3. Designing layouts and end-user interface

According to the 'Anchor & Buoys' model, layouts may only be based on the 'anchor' (leftmost) TO. The 'buoy' TOs exist solely to feed data to any layouts based on the 'anchor'. Although there are as many layouts as TOGs, allowing a full and detailed access to the whole dataset, the database however favours a direct access, both in capturing and searching mode, to the three main objects of the research. These main layouts allow to visualize at a glance the whole relevant data, even if stored in related tables. Moreover, the main layouts have been elaborated in order to always reproduce the same graphical and conceptual pattern.

The Sites layout is here presented as an example (Fig. 6). If the SITES table contains the main attributes about the archaeological sites of ancient Lucania, related data (such as the references, or attendant antiquities and people) are stored in other tables. The Sites layout nevertheless provides access to the whole information in a unique page, organized in distinct areas and tabs. The upper strip of the layout provides the name, place and geographical information for the correct identification and positioning of the site. On the left side of the layout, four tabs gather together all the peculiar characteristics of the actual site: general overview and topography, chronology and phases of occupation, detailed information on the investigated structures, and history of researches. Below these tabs, two portals summarize the main objects of the research that are related to the actual site, that is all the antiquities and people associated with the site that are here displayed in successive rows, referring to the associated TOs in the context of the Sites TOG. At the right end of each row a button allows to access the full related records of these foreign tables, respectively the Object records according to a one-to-many relationship and the People records through an intermediary table according to a many-to-many relationship. Incidentally, thanks to the structure of the TOG in the RG, relating existing records or creating new records in intermediary tables within the interface of the Sites layout is easily achieved. On the right

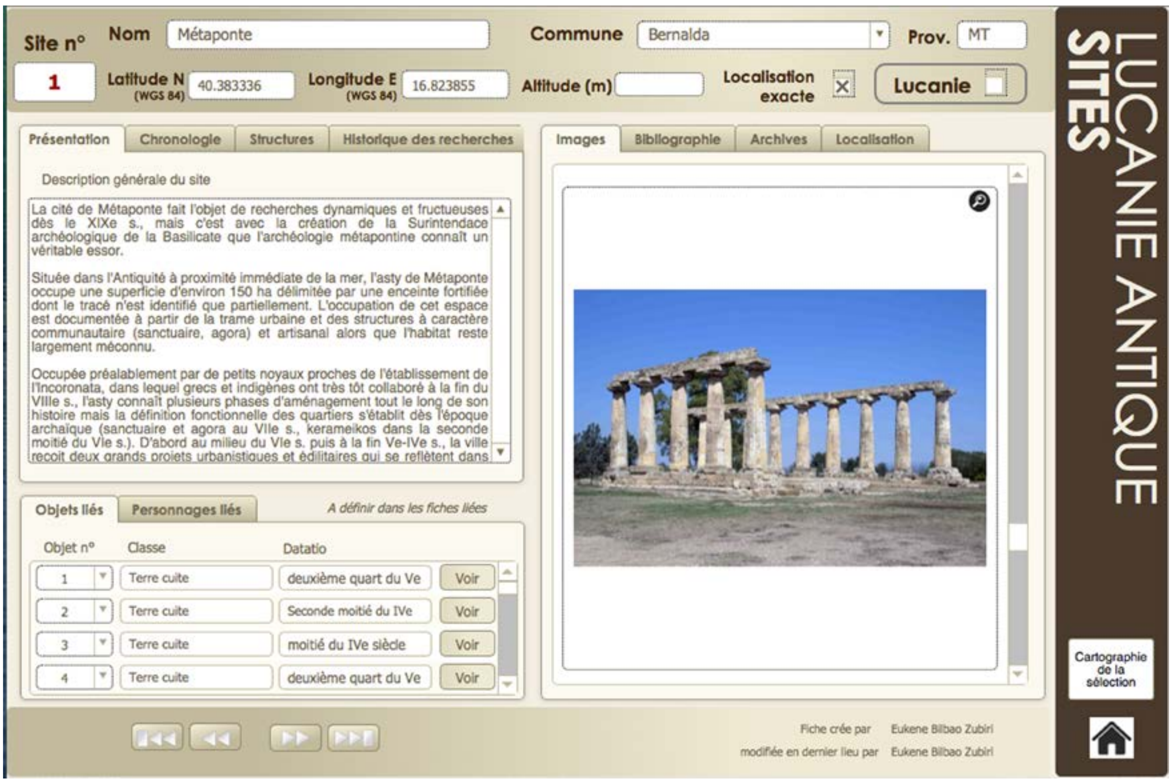


Figure 6: Sites layout of the Ancient Lucania Database (© Lucanie antique, Université Paris 1 Panthéon-Sorbonne)

side of the layout, three tabs correspond to the three sorts of related references, i.e. images, bibliographic and archive records, while a fourth tab offers a basic location of the site on a static physical map centred on today's Basilicata. The 'Images' tab offers thumbnails of all images related to the site according to a one-to-many relationship (a button brings directly to the related images in full definition within the Image layout). The 'Bibliography' and 'Archives' tabs correspond to two portals listing the bibliographic and archive records in successive rows, sorted by chronological order. Again, at the right end of each row a button gives access to the full record through the relevant intermediary table, allowing particular quotation of all referred works.

The same layout has been replicated for the two other main objects of the research, Objects and People: at the top, the identification of the actual object or person; on the left side, peculiar information relevant to it and two portals summarizing the related records; on the right side, tabs corresponding to the related references, i.e. images, bibliographic and archive records. Specific layouts have been designed for the various reference TOGs, according to the specificity of the pertaining information. Finally, various layouts have been created for some TOs belonging to specific threads of 'buoys', mainly for intermediary tables that actually do not have their own TOG. They are not essential to the basic user experience, but they can supplement the information by giving further details or search options without loading down the main layouts with secondary information.

As a result of an idiosyncratic RG, designing layouts has been processed with the end-user experience of the Ancient Lucania Database as a focal point. Paraphrasing G. Orwell's *Animal Farm*, we could say that 'all data are equal, but some data are more equal than others'. From the perspective of the Ancient Lucania project, various pieces of information are indeed more important than others. Contrarily to a generic database, which normally flattens the actual relevance of each table, some data are here prioritised and presented in records according to this heuristic hierarchy. Applying the various operators allowed by FileMaker queries, large parts of the dataset can also be searched according to the predefined layouts and their inherent hierarchy. It means that queries that can be theoretically performed on such a relational dataset but that are not fully relevant to the project archaeological aims are not actually available to end-users. Finally, a particular attention has been given to navigation possibilities between the various layouts. Since the 'Anchor & Buoys' model guarantees that relationships are always made in the specific context of a particular TOG, numerous buttons could be added to the end-user interface, allowing to navigate between related tables. For instance, users can search the People data with specific criteria, then visualize all the Objects or Sites related to the results of the initial query.

5. Adding mapping tools

Mapping the results of a specific query can be very convenient for end-user experience. Basic requirements, such as the location of single or multiples sites, do not need a full GIS-solution with elaborate features of spatial

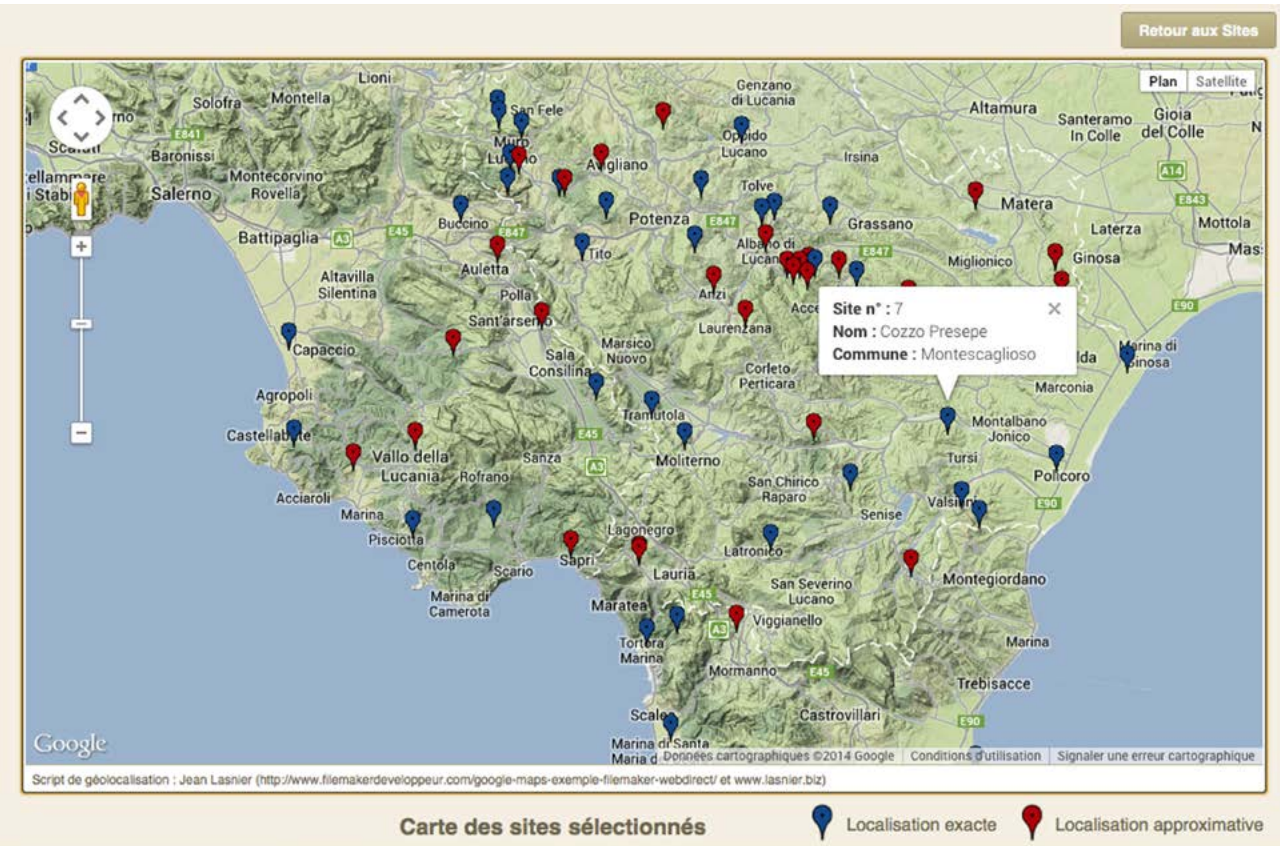


Figure 7: Multiple-sites mapping tool of the Ancient Lucania Database showing the distribution of fortified settlements in ancient Lucania (© Lucanie antique, Université Paris 1 Panthéon-Sorbonne)

analysis. Although these might be of some interest for an on-going research project on ancient Lucania, they are not the main purpose of the actual database. Simple mapping tools are nevertheless a plus that can be added to the user interface thanks to various technologies built in within FileMaker products or offered on the web by Google Maps.

As already mentioned, the location of every archaeological site is provided as a component of the Sites layout. The site is located on a static map thanks to the FileMaker web viewer tool and the Google Maps API, using the geographic coordinates manually included in the description of the site. In compiling the records, archaeologists are indeed asked to enter the location of each site according to Google Maps geographic coordinate system (latitude and longitude in decimal degrees, according to the specific 3857 EPSG code allocated to the Google Maps datum). They also have to specify the accuracy of their cartographic information, by choosing between an approximate location – when no precise location is available in the extant literature – or an exact location (i.e. less than a 100m radius) – thanks to a published map or to personal on-site experience.

The main interest of a mapping tool is actually to provide a map of multiple sites matching definite attributes, such as the result of a query performed within the Sites layout. Generating a map of all fortified settlements, for instance, can be very meaningful to users of the Ancient Lucania Database. This implementation nevertheless required a bit more of specific development and coding, even if there are several examples provided on the internet with the associated code and scripting documentation. To be specific, our work benefitted from the generic script generously provided on his website by Jean Lasnier, a developer of FileMaker solutions (www.filemakerdeveloppeur.com/google-maps-exemple-filemaker-webdirect/), which has been adapted to the peculiarities of the Ancient Lucania Database. The solution consists in writing a JavaScript code recorded as a field of the Parameters table (PAR), whose variables are dynamically modified thanks to a FileMaker script defining an array of figures with the coordinates (and their accuracy level) of the sites to be mapped. The JavaScript code is sent to Google Maps API with the relevant data and then visualized in a FileMaker web viewer (Fig. 7). Sites whose location is approximate are mapped with a red pin, sites with an exact location are marked with a blue pin. To each marker is attached an info window that displays the site number, its name and the municipality in which it is located. The web viewer displays by default a physical map based on terrain information, which can be switched to other map types provided by Google (road map view, Google Earth satellite images, or an hybrid map combining the latters). Of course, the map can be zoomed in or out, and it natively provides an access to the Google Street View service. In substance, end-users can map on demand the result of any query performed against the data displayed in the Sites layout, just by clicking on a button launching the whole process of map making. Moreover, since the Objects and People tables are related to the Sites table in the RG, the result of any query made in the associated layouts can be visualized as a map, such as the distribution map of Lucanian red figures vases hosted in Parisian museums – provided that the relevant provenance information exists – or the itinerary of travellers in southern Italy.

Eventually, the Ancient Lucania Database will offer to all scholars interested by the archaeology and cultural heritage of this southern-Italian area a powerful, but yet user-friendly research tool gathering all useful information, with a critical perspective, as well as elementary mapping functionalities allowing cartographic visualization of the principal search queries.

References

CLARKE, D. L. (1968). *Analytical Archaeology*, London: Methuen.

COSTA, L., DUPLOUY, A., EYRAUD, F. and SARRIS, A. (2008). Vers la mise en réseau des données et des chercheurs : le système d'information de la prospection d'Itanos (Crète orientale). *Archeologia e calcolatori*. 19. pp. 103-122.

D'ACHILLE, A.-M., IACOBINI, A., PRETI-HAMARD, M., RIGHETTI TOSTI-CROCE, M. and TOSCANO, G. (2012) *Viaggi e coscienza patrimoniale : Aubin-Louis Millin (1759-1818) tra Francia e Italia. Voyages et conscience patrimoniale : Aubin-Louis Millin (1759-1818) entre France et Italie*. Rome: Campisano.

D'ACHILLE, A.-M., IACOBINI, A. and TOSCANO, G. (2012). *Il viaggio disegnato : Aubin-Louis Millin nell'Italia di Napoleone, 1811-1813*. Rome: Campisano.

DUPLOUY, A., SCHNAPP, A., SCHNAPP-GOURBEILLON, A., TSIGONAKI, C. and VIVIERS, D. (2009). Recherches archéologiques à Itanos (Crète orientale). *Revue archéologique*. pp. 208-219.

HORDEN, P. and PURCELL, N. (2000). *The Corrupting Sea: A Study of Mediterranean History*. Oxford: Blackwell.

MORGAN, C. (1999). The Archaeology of Ethnicity in the Colonial World of the Eighth to Sixth Centuries BC: Approaches and Prospects. In: *Confini e frontiera nella Grecità d'Occidente. Atti del trentasettesimo convegno di studi sulla Magna Grecia, Taranto 3-6 Ottobre 1997.* Taranto: Istituto per la storia e l'archeologia della Magna Grecia. pp. 85-144.

MORRIS, I. (2005). *Mediterraneisation.* In: Malkin, I. (ed.). *Mediterranean Paradigms and Classical Antiquity*. London: Routledge. pp. 30-55.

PONTRANDOLFO, A. (1982). *I Lucani. Etnografia e archeologia di una regione antica*, Milan: Longanesi.

SETTEMBRINO, G. and STRAZZA, M. (2004).*Viaggiatori in Basilicata* (1777-1880), Potenza: Consiglio regionale della Basilicata.

Webography

DUPLOUY, A. (2006-2014). *Prospection Itanos. Base de données en ligne*. [Online] Available from: http://prospection-itanos.efa.gr/. [Accessed: 23rd July 2014].

DUPLOUY, A. (2013-2014). *La Lucanie antique: archéologie et patrimoine.* [Online] Available from: http://lucanie-antique.univ-paris1.fr [Accessed: 23rd July 2014].

LASNIER, J. (2011-2014). *Développement simple en FileMaker*. [Online] Available from: http://www.filemakerdeveloppeur.com. [Accessed: 23rd July 2014].

Integrating Complex Archaeological Datasets from the Neolithic in a Web-Based GIS

Kai-Christian BRUHN,[1] Thomas ENGEL,[1] Tobias KOHR,[1] Detlef GRONENBORN[2]

[1] i3mainz - Institute for Spatial Information- and Surveying-Technology, University of Applied Sciences Mainz, Germany
[2] Römisch-Germanisches Zentralmuseum (RGZM) , Germany
bruhn@fh-mainz.de, thomas.engel@fh-mainz.de, tobias.kohr@fh-mainz.de, gronenborn@rgzm.de

Abstract

This paper presents DASIS (Distributed Archaeological Sites Information System) that serves as a virtual research environment for settlement structure analyses. In the past decade heterogeneous datasets of the West-Central European Neolithic have been generated in the context of several research projects. A key challenge is the integration of these complex project-specific data models with one-dimensional data tables of a Web GIS.

It will suggest potential ways to join archaeological data and to combine relational data structures providing spatial access on the data to enable further analysis. This scenario entails multiple problems that are prototypical and still lack a general solution. It proposes a modular data concept that is transferable to similar projects. At the same time it reviews and evaluates qualified frameworks that allow for implementing an individually tailored data model into an archaeological information system with GIS capabilities.ment of different methods for sustaining contributor participation through time and a discussion of their implications for the sustainability of the MicroPasts project and (potentially) other archaeological crowd-sourcing endeavours.

Keywords: WebGIS, Archaeological Information System, Neolithic, Data models

1. Introduction

Through a joint research project the i3mainz and the Römisch-Germanisches Zentralmuseum Mainz (RGZM) develop DASIS (Distributed Archaeological Sites Information System), an archaeological information system serving the analyses of complex landscape-based settlement patterns. In the past decade numerous master theses, dissertations and various research projects generated a series of heterogeneous datasets of the West-Central European Neolithic. A key challenge is the integration of these complex project-specific data models with one-dimensional data tables of a Web GIS.

The following components have been identified as vital for the envisaged information infrastructure:

- a Web-based interface compliant to W3C standards
- solid syntactic and semantic indexing and disclosure of information (metadata and ontology)
- role based data editing/management capabilities
- data rights management, while enforcing the use of consistent licenses
- reporting capabilities

The resulting infrastructure should provide interfaces to open standards, a flexible, easy to adapt data structure implemented with mature open source technologies. It should be possible to semantically enrich the research data (using e.g. CIDOC CRM) during or after integration. As the system is planned to be available for long-term use, sustainable technologies are chosen. To achieve this, a thorough review of earlier approaches, as well as a valuation of available platforms, provide a strong conceptual scaffolding for developing an adequate solution.

2. Earlier Research/Research History

Many archaeological research projects have faced similar issues in the attempt to build web portals for data management and analysis recently: First initiatives (Haskiya, 2002; Seleh *et al.*, 2001; Dam *et al.*, 2004; Berg, 2001) started to use the web in the early 2000s to digitally publish research data, making the information available in static tables and scanned documents.

By the end of the decade, web portals indexing archaeological geodatasets (e.g. Mantegari *et al.*, 2006 for the Italian bronze age; Sarris *et al.*, 2008 for Crete), web portals offering map based access to archaeological geodata (e.g. Berg, 2007 for the Norwegian Askeladden database), semantic web portals providing linked geo data sets (e.g. Jeffrey *et al.*, 2007 for the 'common interface environment' and Archaeobrowser of the archaeological data service (ADS); Mantegari, Mosca & Rondelli, 2008 for the Po-basin) evolved. Some of these portals started using international Open Geospatial Consortium (OGC) standards from the GIS world to disseminate data and allow simple geographic and thematic queries within the Web GIS.

However, while most projects have strived to tackle challenges within specific domains, this paper takes a closer look at different frameworks that might be adequate to create a more general approach. A high percentage of projects reviewed do not exist any more. While seeking a general solution for typical problems that archaeological WebGIS`s are facing, it is crucial to identify the reasons that generally lead to their discontinuation. Doing so, common issues are evident on both the organizational as well on the technical side. On the organizational side, temporary limited project funding, along with a frequently problematic migration into institutional environments are the main reasons for discontinued projects.

Technically, proprietary systems without open interfaces, non-standard based systems and systems with restricted licensing have been proven to be negative assets for the survival of Web GIS projects. All of these factors can be summarized with the term 'non-openness'. Furthermore, rapid technological progress, especially if the framework of the project relies on a high amount of self-developed code, limit the sustainability and therefore the lifetime of projects, especially after the main funding of a project ran out (see above). A factor that can be noted as organizational and technological is the presence of a dedicated (Online-) community, which has proven to be an essential element for the success.

Summarizing, it seems crucial to get away from the development of another 'Portal'. These, often project funded, non-community driven projects cannot be kept up to date technologically and organisationally in the long run.

3. Archaeological Background

3.1. Research objectives

The archaeological research objective of the project is to establish a fine-graded and chronologically high-resolved archaeological site database for a succession of periods, in this case the early to middle Neolithic of western Central Europe. This data base could be complementary to larger data bases and serve as a highly resolved model base for regions which are less thoroughly researched. Moreover, it could also be used for a number of questions regarding settlement dynamics, soil preferences, reactions to climate fluctuations, alignments along natural communication routes or changes in settlement complexity.

3.2. Origin, characteristics and integration of the archaeological datasets

The archaeological datasets originate from research- and state crm-generated databases, covering different periods of the Neolithic as well as extending over a variety of areas. Research initiatives include the MK Project (MK Projekt n.d.) and nlDynamics (nlDynamics n.d.) which have both created extensive data sets, while focusing on domain specific research questions. Various master- and doctoral theses, as well as extractions of the state-crm databases of e.g. Hessen, Rhineland Palatinate and Baden-Württemberg provide further data.

Each of these data sets were created to meet individual academic or public concerns, which is reflected in differing heterogeneous modelling of archaeological information, in particular regarding space and time. While one dataset focuses on the question of settlement continuity, others have been assembled to tackle state preservation objectives. Integrating these concepts is a complex task that needs a deep understanding of the underlying data models and attribute values.

Therefore, data integration is necessary for identifying common values in which the archaeological data is stored, the categories they are respectively based on and the history of those categories. The integration process for the legacy datasets identified the following challenges:

> One of the common tasks is the integration of legacy data belonging to an archaeological findspot, which is modelled and interpreted in different levels of detail. For example the number of ceramic fragments belonging to a certain period can be stored numerically or through a description ('yes', 'many' etc.). In the latter case this has to be treated as an interpretation of an initial interpretation (= that ceramic fragments are belonging to a certain period). Nonetheless, together with their temporal dimension, the information is still both valuable and legitimate and should therefore be stored in a Web GIS for further analysis. Additionally the nature of some of the datasets make it necessary to also value the certainty of the information. A feature or find of a findspot could for example 'definitely exist', but it is 'uncertain' if it belongs to phase 'x' or 'y'. This should especially be considered when planning within a research environment that also takes uncertainty into account.
>
> Moreover, the spatial dimension of the legacy data is far from being homogeneous. All of the findspots have spatial coordinates that can be re-projected into a common coordinate-reference-system. However the weakness of many of the coordinates is their historic dimension. For example the sources of digital findspots in federal crm management systems were - until recent times - just markings on a topographical map in the scale 1:25000 or even 1:50000, with the information frequently based on weak evidence such as an interpretation of a place description. It is often impossible to get more detailed information. Today, these coordinates imply an accuracy they cannot represent. It is crucial to identify such cases and to make this information available as well. Therefore, a system should also be able to include the (un)certainty of this precision.
>
> The archaeological datasets also include different levels of detail in their temporal resolution. While the MK (Michelsberg Culture) can not be separated into more than a few coarse periods, the LBK (Linear Pottery Culture) can be divided into up to 15 periods, even representing generations. The resolution also depends on the spatial dimension of the datasets. For instance, regionally different geological conditions could impact archaeological evidence. Working with such temporal concepts including a spatial dimension therefore presumes the possibility not solely to model hierarchical concepts of time, but to also include the relation of time-concepts, e.g. archaeological periods with each other.

4. System Requirements

4.1. Use cases

For DASIS we propose six use cases (Table 1):

– Researchers capture and edit archaeological data (acquisition).

- Users need to be able to query the collected data (discovery).
- Users can define a subset of data to share with others for editing (dissemination).
- A supervising researcher has the ability to assign and manage research tasks (management).
- The system administrator is able to import large datasets in a predefined format (import).
- Users can upload guest (geo–) data for visualization purposes (upload).

User role	Available use cases
Admin	discovery, acquisition, dissemination, management, import, upload
Manager	discovery, acquisition, dissemination, management, upload
Editor	discovery, acquisition, dissemination, upload
Public	discovery, upload

Table 1: Use cases.

According to the present use cases we distinguish between four user roles: The general public has rights to discovery and upload external data for visualization. In addition an editor role can make use of acquisition and dissemination. Furthermore, a manager has access to data and rights management and the administrator to import.

4.2. Functional Requirements

Functional requirements directly arise from the use cases above. Non-functional requirements rather depend on the type of the system. We will address the ones that seem important for a generic framework that we are in search of within the evaluation of these frameworks (Section 5).

One key challenge is to combine complex project-specific data models with one-dimensional data tables of common Web GIS. Archaeological data are usually stored in relational databases with complex data structures that need an interface to be accessed. At the same time OGC Web services like the Web Feature Service (WFS), which are consumed by Web GIS, are designed to provide one-dimensional datasets only. This presents a prototypical problem that seeks a general solution.

In addition to the demands that relate to Input, Management, Analysis and Presentation (cf. Heipke, 2002) functionalities, we consider the following requirements as important: The system needs user management. We pursue that users can share datasets with others and define privileges on these. Meanwhile this entails the necessity for data versioning to track changes, along with their authors. Finally, the GUI of a framework that can be applied in different contexts and countries should support multiple languages.

4.3. Non-functional Requirements

In systems engineering, a non-functional requirement is a requirement that specifies criteria that can be used to judge the operation of a system, rather than specific behaviours. They are often called the qualities of a system.

On top of classical non-functional requirements defined in various ISO standards, community requirements were added as they turned out to be vital for the long-term success of a framework/platform (Raja & Tretter, 2011). These are measurable to a certain degree, which is the main difference to the ISO requirements that cannot be measured without a product used under real conditions (ISO 2014, Chapter 4.2.4). In this case immeasurable criteria have been evaluated without having a final product. To decide if an ISO criterion is to be valued as positive, neutral or negative, the projects have been investigated closely, taking into account that this only reflects a personal perspective (Table 4).

5. Evaluation of Open Source Frameworks

An assessment identified potential frameworks and platforms, that incorporate as many functional and non-functional requirements as possible in order to minimize the development efforts. We have identified 3 key projects with a differing focus for a thorough evaluation: HeuristScholar, Geonode and Arches.

5.1. HeuristScholar

Arts eResearch of the University of Sydney launched Heurist in September 2005. Heurist (Johnson, 2008) is an open-source data-management application for Arts and Humanities scholars (Organisation, Person, Web site, etc.), and thus primarily targets researchers. It follows a central approach, providing a server with one instance of the system, where users can design and create new databases and work with them.

5.2. GeoNode

GeoNode (Geonode, 2014) emerged from the geospatial open source community in 2010 and is driven by the company Boundless (formerly OpenGeo). The software is a complete open-source SDI framework that has a strong focus on the dissemination of spatial datasets of little complexity. It distinguishes between layers (datasets) and maps (compilation of layers). OGC Web services are a key component in the architecture and are used for all interaction with the data (manipulation, querying and visualization).

The application consists of a portal that allows filtering of an underlying catalogue and a Web GIS client that offers basic GIS capabilities such as querying and measuring. GeoNode extends the fundamental publish-find-bind concept of an SDI with collaboration functionality. This enables the user to not only comment on, rate and share content, but also to define access and edit privileges on data which he or she contributed.

Technically, the framework is built on the three tiers of a PostgreSQL/PostGIS database, GeoServer and a JavaScript (OpenLayers, GeoExt, ExtJS) Web GIS client called GeoExplorer. This paper evaluates GeoNode version 2.0.

The database can then be populated with all kinds of data, textual information, images, geographic and temporal data, bibliographic records etc. Heurist provides collaboration through a range of communication channels with features like tagging, rating, notations and bookmarking.

Technically, Heurist uses MySQL on the database level, PHP on the server-side and JavaScript in the Web client. We use the current version 3.1.6 for evaluation.

5.3. Arches

Arches (Myers *et al.*, 2012) is the youngest of the three projects released in a first version in October 2013 by Farallon Geographics. The Getty Conservation Institute (GCI) and World Monuments Fund (WMF) initiated the project to develop a web based inventory system for cultural heritage that primarily targets public authorities. Originating from the MEGA-Jordan project (Myers & Dalgity, 2012) its aims for providing similar functionality through a general framework under an open source license. Arches models archaeological data in graph concepts complying to CIDOC CRM, both on the site and artifact level. The graphical user interface consists of two parts, a map view and a search mask. Users can explore, query and manipulate data.

Technically, Django provides the backend of the Web portal, PostgreSQL/PostGIS acts as the database, and the JavaScript libraries OpenLayers and ExtJS form part of the Web client. In contrast to GeoNode, there is no Web service level between the client and the database. However the Arches developers propose to extend the application via GeoServer to provide selected data through interoperable OGC services. The version used for evaluation is Arches 2.0.

5.4. Comparison and Results

The three evaluated frameworks emerge from different backgrounds and so have their focus on diverse capabilities. Table 2 summarizes the main characteristics of the frameworks. Table 3 and 4 compare functional and non-functional requirements for the defined use cases.

HeuristScholar is a flexible system, which makes it suitable for many use-cases. Data acquisition, management and consumption functionalities as well as collaboration features are exhaustive and match our needs, however there might be issues with spatial and temporal reference models as well as the visualization of the spatio-temporal relations of our archaeological datasets. Also, there are apparently no plans to make the system multilingual. Unfortunately, HeuristScholar does not provide interfaces to open standards like defined by the OGC or set by CIDOC CRM (although the latter would be possible to implement manually with existing tools), which was defined as an essential precondition. The non-functional requirements are to be valued positive, except non-existing mailinglists. Being a mature system, the roadmap as well as the activity numbers show there are still lots of efforts going into development.

GeoNode, relies on one-dimensional data tables when it comes to WebGIS functionality. Technically it is possible to work with db views to manage relational research data in a WebGIS. This would result in focusing on development efforts in that domain, especially when planning for a semantically enriched data integration. Naturally providing a good support for spatial reference.

6. Implementation of a common Use Case

6.1. Data import and extending Arches

Based on our defined requirements and the subsequent evaluation we choose Arches to develop DASIS. The most complex task when using Arches with legacy data is to map existing datasets to the Arches data model. However, the software handles all adaptations within modularized packages, which keeps it flexible and allows the mapping of pretty much any legacy data. The default CDS package of Arches proposes 13 different resource types that are modelled conforming to CIDOC-CRM. Administrators can (de-)activate these resource types. In the given use case, for instance, the vast majority of data can be mapped to one single resource type (ARCHAEOLOGICAL HERITAGE (SITE).E27).

Once the necessary resource types are identified, they need to be compared to the legacy data. While most likely the legacy data might not contain information for all entities they can be useful for data recorded in the future. Contrariwise if the legacy data cannot be mapped completely to the resource type, the graph needs to be extended. In our use case this is the case for three attributes (Figure 1):

– the certainty that a site belongs to a particular period.

	Heust	GeoNode	Arches
Primary audience	Researchers	General public	Public authorities
Data focus	Arts and humanities	Spatial	Cultural Heritage
System type	Single instance	Framework	Framework / Platform
Key feature	Bib. management	Collaboration	Data modelling

Table 2: Quick comparison of the evaluated systems.

Category	Functional requirement	Heurist	GeoNode	Arches
Data acquisition	capture and editing arch. data	yes	restricted	yes
	capture and editing of metadata	yes	yes	yes
	controlled vocabularies	yes	restricted	yes
	import of heterogeneous archaeological datasets	yes	no	yes
Data management	storing complex data structures	yes	no	yes
	metadata concepts and models	Yes	yes	no
	spatial reference models	restricted	yes	restricted
	temporal reference models	restricted	no	yes
Data consumption	interfaces for dissemination of data among researchers	yes	yes	restricted
	compliance to standards of international authorities (OGC/W3C)	no	yes	yes
	searching and filtering capabilities, metadata	yes	yes	yes
	spatial querying capabilities	no	restricted	restricted
	visualization of spatial data	restricted	yes	yes
	visualization of domain specific data	restricted	restricted	yes
Collaboration	rights management on various levels (visualization, editing, managing, etc.)	yes	yes	no
	sharing data with others (working groups/public)	yes	yes	no
Versioning	tracking data modifications by user	no	no	no
Multilingualism	localisation of the GUI should be possible	no	yes	yes

Table 3: Comparison of functional requirements.

Category	Non-Functional requirement	Criteria	Heurist	GeoNode	Arches
Community	maturity	first release - most recent release	2005	2010	2013
	future developm.	date of most recent release	Dec 13	Jan 14	Mrc 14
	size	quantity of developers	5-7	25	7
		quantity of contributors	10-15	66	10
	openness	licence	GNU GPL	GNU GPL	GNU AGPL3
	reliability	most recent emails (dev + user mailinglists)	n/a	23.3.2014 / 19.3.2014	24.03.2014
	activity	most recent commit	Apr 13	May 14	May 14
		commits in last 12 months	3	1667	57
	longterm outlook	impression of planned roadmapn campaigns, etc.	+	o	+
	documentation	extensiveness	large	large	large
		completeness	complete	alm.comp.	complete
		quality	good	good	good
		up-to-date	yes	fairly	yes
ISO/IEC 25000	usability	e.g. efficent design of queries	+	-	+
	portability	operation system / software modules	o	+	+
	stability	productive systems	+	o	o
	security	user roles / data security / technical security	+	+	o
	(administrative) scalability		+	+	o
	reusability		o	o	+
	maintainability		o	o	o
	extensibility		o	o	o
	interoperability		o	+	+
	supportability		+	+	+

Table 4: Comparison of non-functional requirements.

– the certainty that a find or feature was found or identified at a particular site.

– the quantity of finds or features that were identified at a particular site.

The ambiguous ways of modelling circumstances in CIDOC have been discussed in various papers (cf. Binding, 2008). In addition Martin Doerr (Doerr *et al.*, 2014) recently proposed a CIDOC extension for modelling scientific observations. Without digging too deep into CIDOC modelling, we pursue a relatively simple model for our system that keeps a balance between detail and feasibility. This model defines certainty as a type (E55) that is attached to a component (find / feature, time period) via 'has type' (P2). The quantity of identified components is modelled in form of a measurement that has a dimension (E54). The entities measurement type (E55), unit (E58), description (E62), and value (E60). Values can be stored both numerical and descriptive.

In addition, the integrated data comprises the information, how a site was investigated. Arches provides a separate resource type (INVESTIGATION.E7) to manage such information. However, since the migrated data does not possess more detailed information on the investigation the use of a proper resource type would mean a large overhead and impractical utilization for the user (creation of investigation entity and its linkage to a site). Instead we choose to further extend the graph ARCHAEOLOGICAL HERITAGE (SITE).E27 by an investigation type (E55). This way a user can directly indicate the investigation type to a site.

Arches manages controlled vocabularies that provide the different entity values in so called authority files. Globally defined vocabularies, for instance from the Getty Conservation Institute, might be too general or too detailed for a specific use case. Defining proper vocabularies can make sense in this case. The implemented system uses

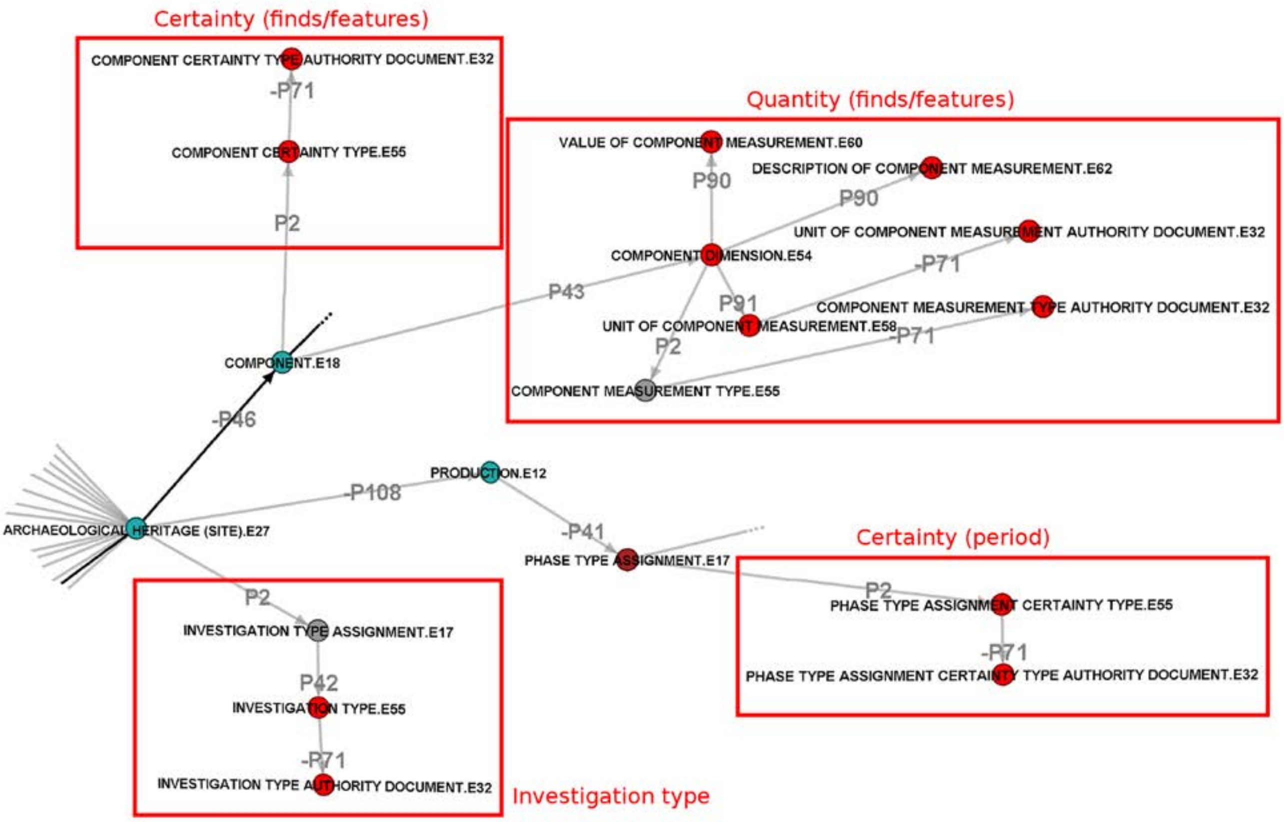


Figure1: Changes to the Arches graph 'ARCHAEOLOGICAL HERITAGE (SITE).E27'

vocabularies that arise from existent data and terms that will be needed for data capture in the future.

After defining the data structure of Arches and the terms the system uses the GUI needs to be adapted to these modifications. The modular code structure of Arches facilitates this task. While views and controllers are extended in JavaScript, a single entry in a configuration file extends search capabilities of Arches by new entities.

6.2. A generic approach

Arches is designed modularly to serve as a generic framework that allows integration of heterogeneous datasets. In version 2.0 it still takes large effort to import legacy data and to implement according extensions. To facilitate data import three aspects are relevant:

- Standard data input format
- Mapping configuration file (including editing tool)
- Generic transformation script

A workflow (Figure 2) following that approach was implemented prototypically with the mapping script. The next step would be a generic script that uses a configuration file to transform the legacy data from the standard input format to Arches' import format. At the same time the script could generate a base vocabulary for the different entities by accumulating values form the legacy data and creating the according authority files.

6.3. Current status and future developments (towards a virtual research environment)

The first phase of our project dealt with the evaluation of different projects and the integration of data into Arches. The available datasets have been mapped to the customized Arches graphs so that the data is available in its full complexity. It is possible to edit and query the data in many ways already. Basic user and rights management

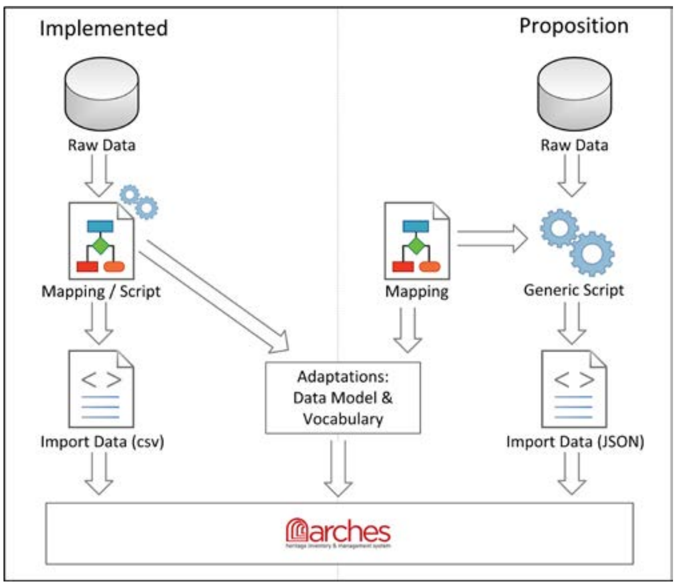


Figure 2: Proposed generic approach to data integration

are realized as well. In the next step we will introduce a GeoServer instance to DASIS, so the advantages of standardized OGC services will become available. The goal is to attach processing services over variable database views to tackle the mentioned research questions.

A main advantage of the Arches project is its strong effort in community building around the framework, which makes it the most sustainable approach. As it is 'just' a technical framework, the integration into institutional environments does not apply to Arches, but one of the key factors for discontinued technological projects is still unanswered as of now – temporary limited project funding.

7. Summary

The evaluation of selected client-server frameworks for the setup of a web-based archaeological information system yielded inconclusive results. Based on a prototypical scenario with complex legacy data the major challenges in developing a sustainable solution became apparent.

Above all, the organizational setting within a limited funding scheme and the tedious integration into institutional environments offering at least a midterm availability of the web-interface are crucial challenges. Multiple problematic issues arise in balancing necessary measures to assure interoperability and a consistent data policy with the preservation of intrinsic information stored in the legacy data.

Although several open, mature and community driven standards, models and technological solution for representing archaeological and spatial (meta-) data exist, it still is a proprietary process to interweave these approaches. When mapping existing data to such a set of models, expressing the uncertainty of values makes it necessary to enforce the adoption of an ontological approach. The selection of frameworks on the one hand proves the availability of sophisticated products but on the other hand it highlights the complexity of fulfilling specific archaeological issues. The decision to choose the Arches-Project is specific to the circumstances of the prototypical scenario and cannot claim universal validity. Nevertheless the main argument leading to the decision was the open, standard-driven and generic character of the framework and the corresponding technical implementation as well as active and lasting endeavors for establishing a broad and international community of users and developers.e a single name, a single scientific review and a single conference, well accepted and opened to all the specialists.

Acknowledgements

The authors would like to thank Krister Kowalski for providing a bibliographic study on earlier research as well as creating the controlled vocabularies and Florian Thiery for support on creating the scripted data import. For valuable discussions about conceptual aspects of the semantic graph structure, we express our thanks to Philip Carlisle from English Heritage.

References

BERG, E. (2001). National Registries of Sites and Monuments in Norway – Developing GIS-based Databases. In: Stancic, Z. and Veljanovski, T. (eds.). *Computing Archaeology for Understanding the Past. CAA 2000. Computer Applications and Quantitative Methods in Archaeology. Proceedings of the 28th Conference*, Ljubljana, April 2000 (BAR International Series 931), Oxford: Archaeopress. pp. 133-138.

BERG, E. (2007). Using a GIS-based Database as a Platform for Cultural Heritage Management of Sites and Monuments in Norway. In: Clark, J.T. and Hagemeister, E.M. (eds.) *Digital Discovery. Exploring New Frontiers in Human Heritage. CAA2006. Computer Applications and Quantitative Methods in Archaeology. Proceedings of the 34th Conference, Fargo, United States, April 2006*, Budapest: Archaeolingua. p. CD-ROM 345-351.

BINDING, C. (2008). Semantic Interoperability in Archaeological Datasets: Data Mapping and Extraction via the CIDOC CRM. In: Christensen-Dalsgaard, B., Castelli, D., Ammitzboll Jurik, B. and Lippincott, J. (eds.) Research and Advanced Technology for Digital Libraries. *12th European Conference, ECDL 2008*, Berlin Heidelberg: Springer. pp. 280-290.

DAM, C., AUSTIN, T. and KENNY, J. (2010). Breaking Down National Barriers: ARENA – A Portal to European Heritage Information. In: Nicolucci, F. and Hermon, S. (eds.). *Beyond the Artifact. Digital Interpretation of the Past. Proceedings of CAA2004, Prato 13–17 April 2004*, Budapest: Archaeolingua. pp. 94-98.

DOERR, M., KRITZOTAKI, A., ROUSAKIS, Y., HIEBEL, G., THEODORIDOU, M. (2014). CRMsci: the Scientific Observation Model, An Extension of CIDOC-CRM to support scientific observation. Produced by FORTH and collaborators. Version 1.2.1 (draft) April 2014. [Online] Available from: http://www.ics.forth.gr/isl/CRMext/CRMsci/docs/CRMsci1.2.1.pdf. [Accessed 14th July 2014].

GeoNode.org (2013) GeoNodes`s Documentation. [Online] Available from: http://docs.geonode.org/en/2.0/ [Accessed: 14th July 2014].

HASKIYA, D. (2002). Developing an Information System for Archaeological Sites and Monuments – Data Model and Construction. In: Burenhult, G. and Arvidsson, J. (eds.). *Archaeological Informatics: Pushing The Envelope. CAA2001. Computer Applications and Quantitative Methods in Archaeology. Proceedings of the 29th Conference, Gotland, April 2001* (BAR International Series 1016). Oxford: Archaeopress, pp. 49-52.

HEIPKE, C. (2002). Requirements for Modern Geographic Information Systems. Integrated System for Spatial Data Production, Custodian and Decision Support.

ISPRS Commission II, Symposium 2002, August 20-23 2002, Xi'an, China, pp. 173-184.

ISO, International Organization for Standardization (2014), ISO/IEC Standard 25000:2014(en). [Online] https://www.iso.org/obp/ui/#iso:std:iso-iec:25000:ed-2:v1:en [Accessed: 14th July 2014]

JEFFREY, S., KILBRIDE, W., WALLER, S. and RICHARDS, J. (2008). Thinking Outside the Search Box: The Common Information Environment and Archaeobrowser, In: Posluschny, A., Lambers, K. and Herzog, I. (eds.). *Layers of Perception. Proceedings of the 35th International Conference on Computer Applications and Quantitative Methods in Archaeology (CAA), Berlin, Germany, April 2–6, 2007* (Kolloquien zur Vor- und Frühgeschichte, Vol. 10), Bonn: Habelt, pp. 200-205.

JOHNSON, I. (2011). Heurist: a Web 2.0 Approach to Integrating Research, Teaching and Web Publishing, In: Jerem, E., Redo, F. and Szeverenyi, V. (eds.). *On the Road to Reconstructing the Past. Computer Applications and Quantitative Methods in Archaeology (CAA). Proceedings of the 36th International Conference. Budapest, April 2-6, 2008.* Budapest: Archeaeolingua, pp. 93-99 (CD-ROM 291-297).

KINTIGH, K. (2006). The Promise and Challenge of Archaeological Data Integration. *American Antiquity,* Vol. 71, No. 3 (July 2006), pp. 567-578.

MANTEGARI, G., CATTANI, M., DE MARINIS, R.D. and VIZARRI, G. (2007). Towards a Web-based Environment for Italian Prehistory and Protohistory, In: Clark, J.T. and Hagemeister, E.M. (eds.). *Digital Discovery. Exploring New Frontiers in Human Heritage. CAA2006. Computer Applications and Quantitative Methods in Archaeology. Proceedings of the 34th Conference, Fargo, United States, April 2006*, Budapest: Archaeolingua. CD-ROM pp. 374-381.

MANTEGARI, G., MOSCA, A. and DE SALVO, M. (2011). Free and Open Source WebGIS Solutions for the PO-BASyN Project, In: Jerem, E., Redo, F. and Szeverenyi, V. (eds.). *On the Road to Reconstructing the Past. Computer Applications and Quantitative Methods in Archaeology (CAA). Proceedings of the 36th International Conference. Budapest, April 2-6, 2008*. Budapest: Archeaeolingua, pp. 249-255 (CD-ROM 380-386).

MEYER, E., GRUSSENMEYER, P., PERRIN, J.P., DURAND, A. and DRAP, P. (2008) A Virtual Research Environment for Archaeological Data Management, Visualization and Documentation, In: POSLUSCHNY, A., LAMBERS, K. and HERZOG, I. (eds.) *Layers of Perception. Proceedings of the 35th International Conference on Computer Applications and Quantitative Methods in Archaeology (CAA), Berlin, Germany, April 2–6, 2007*, Bonn: Habelt, pp. 173-178.

MK Projekt (n.d.). Die Anfänge sozialer Komplexität: Erdwerke, Rohstoffnutzung und Territorialität im Neolithikum. Deutsch-französische Forschungen zur Michelsberger Kultur, [Online] Available from: http://www.anr-mk-projekt.fr/spip.php?lang=en. [Accessed: 14th July 2014].

MYERS, D., DALGITY, A. (2012). *The Middle Eastern Geodatabase for Antiquities (MEGA): An Open Source GIS-Based Heritage Site Inventory and Management System. Change Over Time*, Vol.2 (1), Spring 2012. p.32-57.

MYERS, D., DALGITY, A., AVRAMIDES, I. and WUTHRICH, D. (2012). Arches: An Open Source GIS for the Inventory and Management of Immovable Cultural Heritage. In: Ioannides, M., Fritsch, D., Leissner, J., Davies, R., Remondino, F. and Caffo, R. (eds.). *Progress in Cultural Heritage Preservation. 4th International Conference, EuroMed 2012, Limassol, Cyprus, October 29 – November 3 2012. Proceedings*. Berlin, Heidelberg: Springer. pp.817-824.

nlDynamics (n.d.) Studien zur Dynamik neolithischer Gesellschaften (nlDynamics), [Online] Available from: <http://web.rgzm.de/forschung/schwerpunkte-und-projekte/a/article/studien-zur-dynamik-neolithischer-gesellschaften-nldynamics/>.[Accessed 16th July 2014].

RAJA, U. and TRETTER, M.J. (2011) Predicting OSS Development Success; A Data Mining Approach. *International Journal of Information System Modelling and Design* (IJISMD). 2 (4). pp.27-48.

SALEH, F. and BAHGAT, R. (2002) The Archaeological Map of Egypt (Archaeological Heritage Resource Management System), In: Burenhult, G. and Arvidsson, J. (eds.). *Archaeological Informatics: Pushing The Envelope. CAA2001. Computer Applications and Quantitative Methods in Archaeology. Proceedings of the 29th Conference, Gotland, April 2001* (BAR International Series 1016), Oxford: Archaeopress. pp. 253-260.

SARRIS, A., TRIGKAS, V., PAPADAKIS, G., PAPAZOGLOU, M., PERAKI, E., CHETZOYIANNAKI, N., ELVANIDOU, M., KARIMALI, E., KOURIATI, K., KATIFORI, M., KAKOULAKI, G., KAPPA, E., ARTHANASAKI, K. and PAPADOPOULOS, N. (2008) A Web-GIS Approach to Cultural Resources Management in Crete: the Digital Archaeological Atlas of Crete, In: Posluschny, A., Lambers, K. and Herzog, I. (eds.). *Layers of Perception. Proceedings of the 35th International Conference on Computer Applications and Quantitative Methods in Archaeology (CAA), Berlin, Germany, April 2–6, 2007* (Kolloquien zur Vor- und Frühgeschichte, Vol. 10), Bonn: Habelt. pp. 241.

Enhanced 3D-GIS: Documenting Insula V 1 in Pompeii

Giacomo LANDESCHI,[1] Nicolò DELL'UNTO,[1]
Daniele FERDANI,[2] Stefan LINDGREN,[3] Anne-Marie LEANDER TOUATI[1]

[1] Department of Archaeology, University of Lund, Sweden
[2] Istituto per le Tecnologie Applicate ai Beni Culturali, CNR Rome (CNR ITABC), Italy
[3] Humanities Lab, University of Lund, Sweden
giacomo.landeschi@ark.lu.se, nicolo.dell_unto@klass.lu.se, daniele.ferdani@itabc.cnr.it, stefan.lindgren@humlab.lu.se, anne-marie.leander_touati@klass.lu.se

Abstract

This project was recently carried out as a part of The Swedish Pompeii Project, which was started in 2000 as a research and fieldwork activity initiated by the Swedish Institute in Rome. The aim was to record and analyze an entire Pompeian city block, Insula V 1. Since autumn 2011 a new branch of advanced digital archaeology, involving 3D reconstructions and documentation methods, was added to the project agenda. The insula was completely digitized using laser scanner technology and the raw data were employed to develop different research activities in the area of digital visualization and analysis. In this context, a newly developed research line was setup with the purpose of implementing the above mentioned dataset into a 3D-GIS platform (ESRI ArcGIS 10 suite). The system was designed with the goal of (i) increasing the connection among the different typologies of data recorded in the last 10 years, (ii) retrieving a larger number of information, and (iii) analyzing data from a fully functioning geodatabase made up of 3D models. First step of the work consisted of the GIS integration of the 3D models (previously acquired with laser scanner technology) of the south house of Caecilius Iucundus, which was used as a case study to make a general assessment of the project feasibility. Each 3D model was transformed into a multipatch feature, and it was associated to an attribute table along with its related information. In this way it was possible to link each feature to the documentation provided by the Swedish Pompeii Project website (http://www.pompejiprojektet.se/). Furthermore, the editing extension available allowed digitizing 3D features straight into GIS, providing with the opportunity to setup even very basic modelling functions. This project allowed us to use all the potential provided by a GIS platform to explore in a GIS the complexity of an ancient building geometry along with its architectural details and peculiarities. Furthermore, a virtual reconstruction of the Caecilius Iucundus house was super imposed alongside the acquired site and used to make comparisons between bottom up and top down representations. The development of this aspect will allow using GIS analytic tools for investigating aspects connected to the cognitive process in an ancient building, with the possibility of generating new hypotheses about the spatial organization of the Roman house based on the notion of visibility of specific elements.

Keywords*:* GIS, 3D, Pompeii, Building Archaeology, Virtual Reality

Introduction

The use of the third dimension as an additional field of analysis in the study of ancient buildings can dramatically contribute to explore new ways of data visualization, leading to a real breakthrough in the overall documentation and interpretation process. In this regard some experiments have been recently conducted with the purpose of investigating the use of 3D models in fieldwork activity. These were mainly aimed at assessing how 3D affected the interpretation process of an archaeological excavation (Callieri *et al.,* 2011; Dellepiane *et al.*, 2013; Forte *et al.*, 2012; Opitz, 2012). In addition, the potential of 3D in support of documentation methods has been widely discussed (Dell'Unto, 2014; De Reu et al. 2014; Doneus and Neubauer 2005; Doneus *et al.,* 2011; Forte *et al.*, 2012; Losier *et al.*, 2007) along with a set of related technical issues (Allen *et al.,* 2004; Barcelò *et al.*, 2003; Frischer, 2008; Gillings and Goodrick, 1996). Aim of the paper is thus to define an effective and solid work pipeline to integrate even geometrically complex 3D surface texturized models, in a GIS and to take advantage of the three-dimension to improve the quality of data analysis in an fully-functioning 3D space. These models, mainly deriving from laser scanner acquisition and Image Based 3D modelling, have been hence analyzed and studied by means of spatial-analysis techniques. This research activity was developed in the framework of the Swedish Pompeii project, and allowed defining limitations and potentials of such a methodological approach.

1. The Swedish Pompeii Project: previous work

Since fall 2000 the Swedish Institute in Rome started a field investigation campaign with the aim of documenting and studying a full Pompeian city block, Insula V 1. From the very beginning, different types of documentation techniques were tested in order to define an efficient investigation method, which would comprise the different aspects that characterize the ancient buildings. Goal of this work was procurring an infrastructure capable to visualize the structures of the insula not as separate entities, but as part of a total. This approach highlighted the importance of focusing on the relation between the different typologies of elements that characterize Pompeian domestic architecture (Leander Touati 2010; Staub 2009). In the frame of this work, a digital research platform was realized. This infrastructure proved to be capable to manage the data detected in the field and to visualize and highlight the relations between different typologies of data. All the information collected during the investigation campaigns are organized and published on a website (www.pompejiprojektet.se/insula.php), that allows rapid access to different levels of data, going from general information toward more detailed documentation, such as photographs, ortho-mosaics, ground or section plans and excavation reports. Since autumn 2011, Lund University (Institute of Archaeology and Ancient History together with the Humanistic Laboratory), in collaboration with the National Research Council of Italy (Institute of information Technology and science 'A. Faedo'), started a project of digital acquisition by means of integrated spatial technologies, such as Phase Shift Laser Scanners and Image Based 3D Modeling of the Insula V 1. (Dell'Unto et al., 2013). Goal of this work was to acquire in high-resolution the Pompeian city block to document in three dimensions the structures that characterize the insula, and increasing the possibilities, for studying the relationships between the different parts that characterize Pompeian domestic architecture. The digital material developed in the framework of this new project was used to design a data model capable to manage -in three dimensions- most of the information detected during the investigation of the insula V 1, combining into a 3D-GIS platform (i) 3D surfaced models of the ancient structures, (ii) the 3D interpretation (virtual reconstruction) of the structures, (iii) the previous documentation realized in the frame of the 3D Swedish Pompeii Project such as: ground and section plans. Moreover, once processed, the 3D models of the actual structures of the insula have been made available on line using WebGL to visualize the high resolution models of the insula directly through web browsers (Fig. 1). The development and the experimentation of this tool allowed connecting this work with the classic documentation disseminated during these years through the internet.

The development of such web access to visualize the 3D data would allow anyone interested in studying insula V1 a direct access to the information elaborated by the project team. On the other hand, the set up of a 3D-GIS was tested on the house of Caecilius Iucundus as a case study and was aimed at (i) developing an effective documentation pipeline to be extended to the rest of the insula (ii) making an assessment of GIS as a tool for monitoring and quantifying the conservation status of ancient structures (iii) delivering new solutions to obtain more accurate information about geometrical relations of walls shared between adjacent rooms. In addition, the spatial analysis tools which are available in GISs, have been tested to make a quantitative assessment of the spatial configuration of the virtual buildings. Such an analysis was intended to provide some insights into the visual properties of specific objects, like wall inscriptions or paintings, placed at their actual location inside the 3D environment. Archaeologists can thus obtain some clues about the cognitive framework and the visual impact that particular categories of objects could have exerted on the ancient inhabitants of the building.

Figure 1: This image shows the high resolution model of the atrium of Casa del Torello di Bronzo, visualized through the webGL system developed by The Visual Computing Lab, CNR, Pisa. Once post processed, an high resolute version of the 3D models of the entire insula have been implemented in the Swedish Pompeii project website. The 3D models can be use interactivelly through Mozilla or Google Chrome.

2. 3D-GIS implementation

2.1 Dataset framework

One of the primary goal in this research was to improve current site documentation strategies by implementing geometrically complex texturized 3D models in a GIS platform. In this respect, it was crucial to develop an effective 3D Geographic Database Management System (GDBMS) to collect and store most of the archaeological documentation gathered in the context of the Swedish Pompeii Project. A GIS system was thus set up with the purpose of interconnecting different categories of data (3D models, raster, vector, images etc.). Although the potentialities offered by GIS in documenting the archaeological record do not need to be mentioned as they are well known, using the three-dimension as a further informative layer can dramatically enhance the analytic performance of the system. Due to the strong limitations occurring in traditional bi-dimensional GISs (Llobera 2003), archaeologists are not usually allowed to effectively investigate and fully understand the spatial configuration of the ancient buildings. This research was partly focused on developing an effective methodology to overcome these limitations. The strategy was to setup a fully-functioning GIS platform that could be used to manage, store and analyze those geometrically complex 3D models, derived both from laser scanner acquisition and image-based 3D modelling. In this regard, ESRI ArcGIS software has been chosen to conduct this research. This choice was due to the significant technical improvements produced on 3D Analyst extension, that gave us the possibility of testing an enhanced data visualization experience, a rendering speed-up along with better performing allocation settings (ESRI 2010). In addition, the geodatabase structure, a native data format for ArcGIS, constitutes a de facto standard for the geo-spatial dataset management, providing a lot of solutions for different categories of users (Zeiler, 2010). Here, specific functions can be defined in order to maintain data integrity, to define topological rules and to set relations between features. Another relevant improvement is given by the editing functions that enable users to digitize in 3D. Unlike bi-dimensional GISs, in which 3D-related information is stored in the form of a feature's attribute, here the height information is stored directly inside vector feature's geometry so that each shape is defined by three different coordinates: x, y, z (ESRI 2012). Such an aspect enables users to take advantage of a 'real' three-dimension to perform advanced analysis on the 3D models imported within the GIS environment. Moreover, a general ease-of-use of software, made ArcGIS the ideal platform to be adopted in this project. In the context of the Swedish Pompeii Project archaeologists are indeed encouraged to use the tools available in GIS to perform different kinds of analysis on the 3D models without receiving any technical support. In brief, the process of knowledge can be increased by the joint work of scientists creating their own interpretation in the same digital environment; at the same time, different interpretations can be shared and compared, enhancing the aspects of multi-vocality and reflexivity.

2.2 The Caecilius Iucundus' South House Case Study

As a project case study the southern part of the house of Caecilius Iucundus was chosen. As afore mentioned, different datasets have been imported into the geodatabase, which was set at the local Italian coordinate system, as it is the current standard format in use by the Archaeological Superintendence of Pompeii (Foss and Dobbins, 2007). According to the scheme showed in (Fig. 2), (i) a ground plan, (ii) the 3D models and (iii) the digitized features derived from the topographical survey were integrated into the systems. A dataset of raster maps spanning in scale from 1:1000 to 1:20 was added; a general map of Pompeii's archaeological area was completed with the recently established plan drawn in the field by means of total station survey during the work of the Swedish Pompeii project, and verified by comparisons with the scanned maps, in particular, with a detailed plan of Caecilius Iucundus' South House. A Digital Elevation Model (DEM) of the house was obtained by interpolating elevation values from vector points digitized over the scanned house plan. On a broader scale, a less resolute DEM (one spot per 20 meters) provided the topography of the Pompeii area.

In regard with 3D model GIS implementation, as already mentioned, few attempts have been made thus far to implement such complex 3D models on a GIS platform. Among these, it is worth mentioning the work carried out by Koehl and Lott (2008), which illustrates an integrated approach of 3D acquisition techniques and GIS implementation. Recently, Opitz and Nowlin (2012) described the data import of image-based 3D models into a GIS. Similarly, our case study constitutes one of the few systematic attempts to integrate geometrically complex 3D data in a geo-referenced system and to analyze and interpret the data collected in a fully three-dimensional space. In the framework of the Swedish Pompeii Project, one of the purpose was actually to draw from scratch an innovative methodology to use the third dimension as an additional field of analysis. Firstly, 3D meshes were optimized and texturized using high resolution images acquired in the framework of the Swedish Pompeii Project. Then, the resulting models were scaled based on a scale factor of 0.001, according to the difference in measurement units used in data acquisition (millimeters) and GIS data visualization (meters). Subsequently, data were imported as COLLADA files into the ArcScene 3D Analyst extension, a visualization tool based on OpenGL, that supports texture mapping, complex 3D line symbology, surface creation and display of Digital Terrains models (ESRI 2013). Here the previously geo-

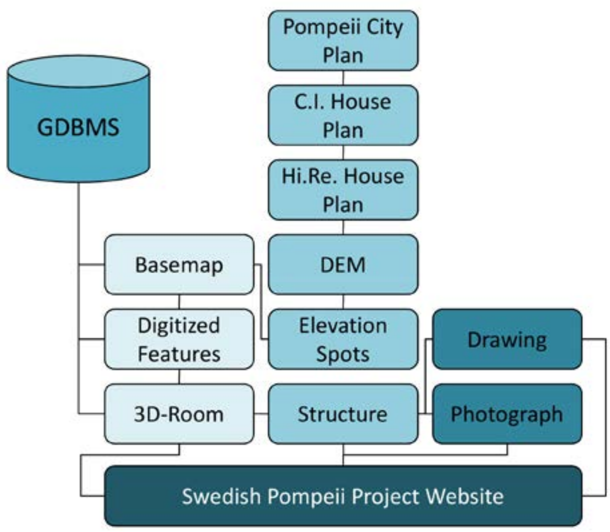


Figure 2: Geodatabase architecture featuring all the inter-connected layers. Notably, each 3D model imported as a single room of the house is connected to many structures (N wall, floor, etc.). Each structure is in turn connected to a set of photographs and drawings based on the documentation collected in the context of the Swedish Pompeii Project. Each one of these layers is hence linked to the project website where further information can be retrieved.

referenced maps were used as a reference to place the 3D models at their absolute coordinates. Each COLLADA file was thus imported in ArcScene and transformed into a multipatch file, a data format designed for the boundary representation of 3D objects (ESRI 2012). Then each room's model was set at its actual location based on the Caecilus Iucundus' house plan as a reference. In addition, an accuracy less than 1 millimeter was reached by using the snapping tool to match the models together; the final result was a very precise alignment of Caecilus Iucundus' south house at its absolute x, y coordinates. A further issue addressed was the actual alignment along the z axis; for this purpose, additional data had to be implemented based on the height information available for the house. Subsequently, in order to get the necessary reference for the 3D building alignment, a set of elevation spots was drawn over the 1:20 scale scanned map. Based on the interpolation of those vector points, a Digital Elevation Model was thus produced. Next, the DEM was set as a base level for the house plan which was moved at its absolute z coordinates. Finally, the whole multipatch model of the house was 'lifted up' to match its ground floor based on the DEM plan (Fig. 3).

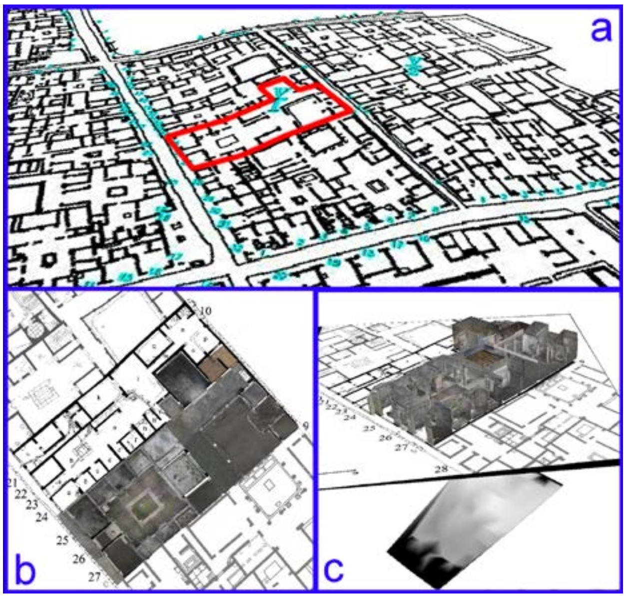


Figure 3: The image shows the south house of Caecilius Iucundus after being imported in GIS. The city plan of Pompeii (after Foss and Dobbins 2007) has been used as a general reference for the entire dataset a); then a small scale plan (after Ezequiel Pinto-Guillaume) has been added to georeference the 3D models of he house b) and a DEM has been produced based on a set of vector elevation spots derived from a total station survey, providing the models alignment along z coordinate c).

A further part of the process was made by the web connection between the 3D models and the database structure featured on the Swedish Pompeii Project website (http://www.pompejiprojektet.se/insula.php). The room was chosen as the basic database entity and subsequently, all the metadata architecture was designed based upon the original framework defined by the project website. Thus, each room entity was related both to the 3D object (one-to-one relationship) and to other entities, namely single architectural structures (N wall, floor, etc.). Each structure was in turn connected to different sets of photographs, drawings (one-to-many relationship) that were used to store more detailed information related to the single structure itself. Therefore, specific topological rules were defined so as to improve the operational performances as well as to increase the general database consistency. Additionally, each attribute table associated to each feature was provided with a specific hyperlink field; this was made to connect the selected record to its corresponding descriptive webpage available on the Swedish Pompeii Project website. As a final result, a three-dimensional

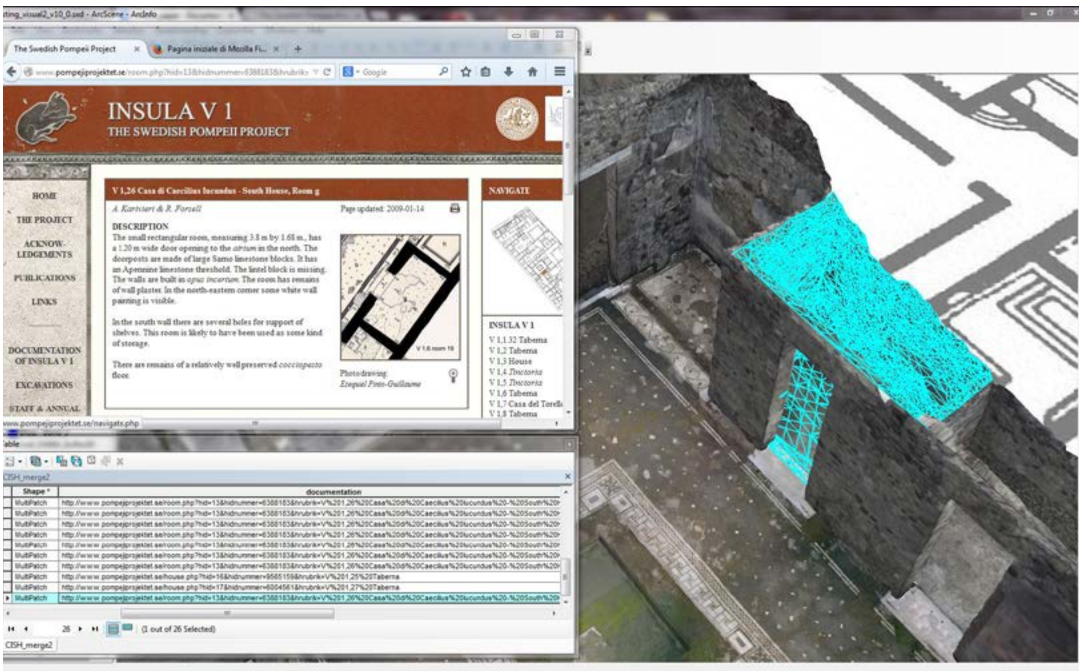


Figure 4: This image shows the possibility to query each part of the 3D models implemented in the GIS, retrieving through an hyperlink the information previously implemented in the website by the Swedish Pompeii project team.

environment directly connected to the currently available project documentation was obtained. Users are thus enabled to interrogate the objects, query the database and retrieve information from the website (Fig. 4).

2.3 Results

One of the main goals was to develop a solid data implementation pipeline to explore, analyze and measure in a geo-referenced system, all the geometric elements that characterize the structures of the house. A crucial achievement resides in the new scenarios that open for what site documentation strategies, as all of the architectural features detected in the course of the building investigation can be surveyed in 3D. The possibility to edit directly in 3D enables archaeologists to bypass analysis through bi-dimensional drawing in the form of plans and elevations. Instead they may conduct their work by annotating the results of the analysis performed in the field directly on the 3D replica of the surveyed object. By recording information directly in the field, - using the 3D-GIS as main documentation/interpretation tool - archaeologists are enabled to keep a direct visual relation with the object and make a real-time assessment of the quality of their own observations. The only limitations -in terms of accuracy- are due to the sampling choices made by the surveyor (i.e. the number of points chosen to draw the 3D polygon or polyline). In addition, the 3D-GIS provides a platform on which archaeologists can use the acquired models as a geometrical reference for data analysis, thereby significantly improving the quality of their own interpretation. This combination of 3D and GIS provides archaeologists with a direct access to the entire dataset of spatial information previously collected in the geodatabase. They also have the opportunity to benefit (already during fieldwork activity) from

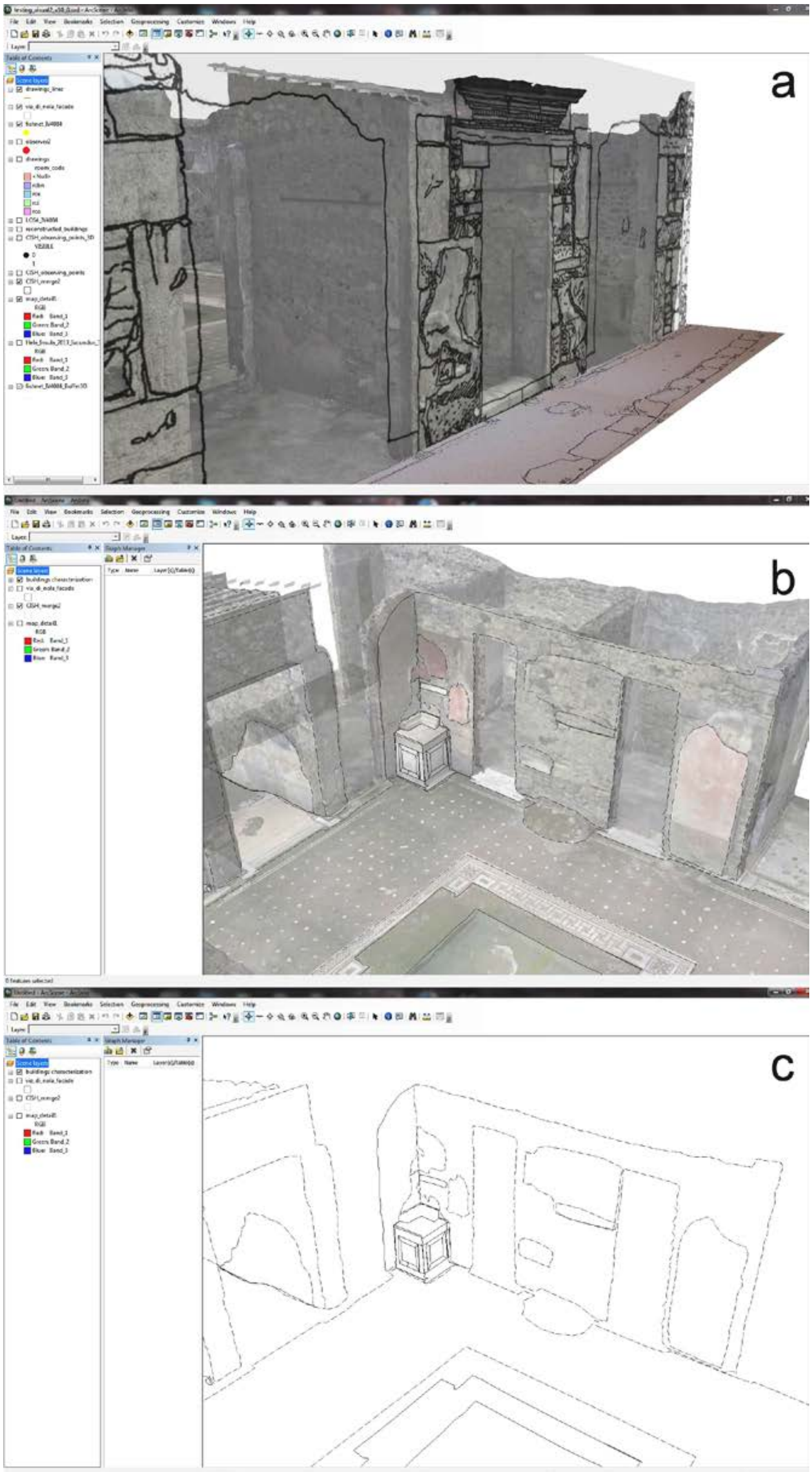


Figure 5: 3D-GIS becomes a platform where different types of documentation are put in relation and compared: image (a) shows raster format files related to the elevations and plans realized during previous field campaigns by project team members (after Ezequiel Pinto-Guillaume). Each element can thus be compared with the 3D models of the corresponding architectural elements: image (b) shows a 3D vector drawing recently made to facilitate the interpretation of the features visualized through the 3D models. The same interpretative layer can be seen as a stand-alone layer where the different contexts are better highlighted (c).

Figure 6: 3D editing tool increases the possibility of using the GIS platform to analyze the geometrical relations between opposite faces of a same wall. In this example, the perimeter of an alcove built in the southwestern wall of room p is examined a). It is interesting to know which portion of the opposite facade of the wall is affected by the orthogonal projection of the previously drawn perimeter; to know this, it is possible to use a Delta function, enabling the move of a digitized feature along x,y axes. Starting from the wall in room p b), final result is the drawing of the perimeter moved based on the thickness of the wall (0.2 m) and re-projected (orthogonally compared to the original x,y alignment) on the wall in the tablinum c).

a system that allows 'dynamic' comparison; that is, to access all data detected in the same area, notwithstanding if retrieved at different stages of the investigations, by other team members and by other teams. This potential significantly increases the ability of exploiting the multi-temporal dimension (often recalled as 4D); that is to examine chronological dimensions pertaining to the history of the studied structures, to the field investigation itself (Fig. 5) and to monitor the degree of degradation affecting architectural features over time. Among the analytical improvements due to this kind of 3D-GIS is worth to mention also the possibility to connect and measure stratigraphic units belongings to different sides of a same wall (Fig. 6).

To sum up, the obtained results advance a powerful instrument of analysis able to significantly improve the strategies of documentation in the field. The 3D-GIS favors fast comprehension of relationships between different architectural entities or buildings. The system has been set up with the purpose of managing the 3D-related information together with the existing datasets previously realized, such as plans and elevation maps. In a diachronic perspective this is an important achievement that enables archaeologists to obtain a complete status presentation and a complete picture of all information recorded in the field.

3. A tool for 3D Visual Analysis

In the frame of this research line, some preliminary experiments have been conducted with the purpose of testing 3D-GIS as a possible new way to make an assessment of the visual properties of certain categories of objects placed in the three-dimensional space of the Roman house. In the following two kinds of analyses will be presented. The use of 3D-GIS for interpretation and virtual reconstruction of the ancient environment and for defining viewing lines, that is cognitive aspects pertaining to how the ancient viewer may have apprehended this environment.

Figure 7: In the upper part of the image the virtual interpretation of the House of Caecilius Iucundus displayed with the different level of consistency adopted in the model. Each level is represented by a different color .

3.1 Virtual Reconstruction in 3D-GIS

The framework of this examination, the virtual reconstruction of the domus was performed by using several types of archeological and architectural data. The interpretation was realized with the support of experts from the Archeological Department of Lund University. As first, the documentation regarding the archaeological remains, (pictures, archaeological reports, stratigraphic interpretation, etc.) were collected and organized, as described in the earlier parts of this article. The integration of all materials was supplemented by using various range-based and image based data-gathering techniques, involving both digital photographic straightening techniques, enabling acquisition of 'rectified' images of the elevations (eliminating distortion arising from the camera lens), and 'laser scanning' techniques for 3D modelling. Subsequently, the reconstruction process was developed combining and integrating the three dimensional information with the archaeological data previously collected during the site investigation, such as previous publications, bibliographic resources, hypothetical scientific-based reconstructions, drawings, paintings, etc. Virtual hypotheses, concerning parts of the domus for which we lack archaeological evidences or other trustworthy sources, were complemented by a procedure pertaining to different levels of consistency, based on general knowledge of Pompeian construction techniques and building or decorative modules, or on comparative data identified in other, nearby Pompeian buildings. After the interpretative studies, the scanned model of the domus and the rectified photos were used as references to draw up the 3D reconstructive model (Fig. 7).

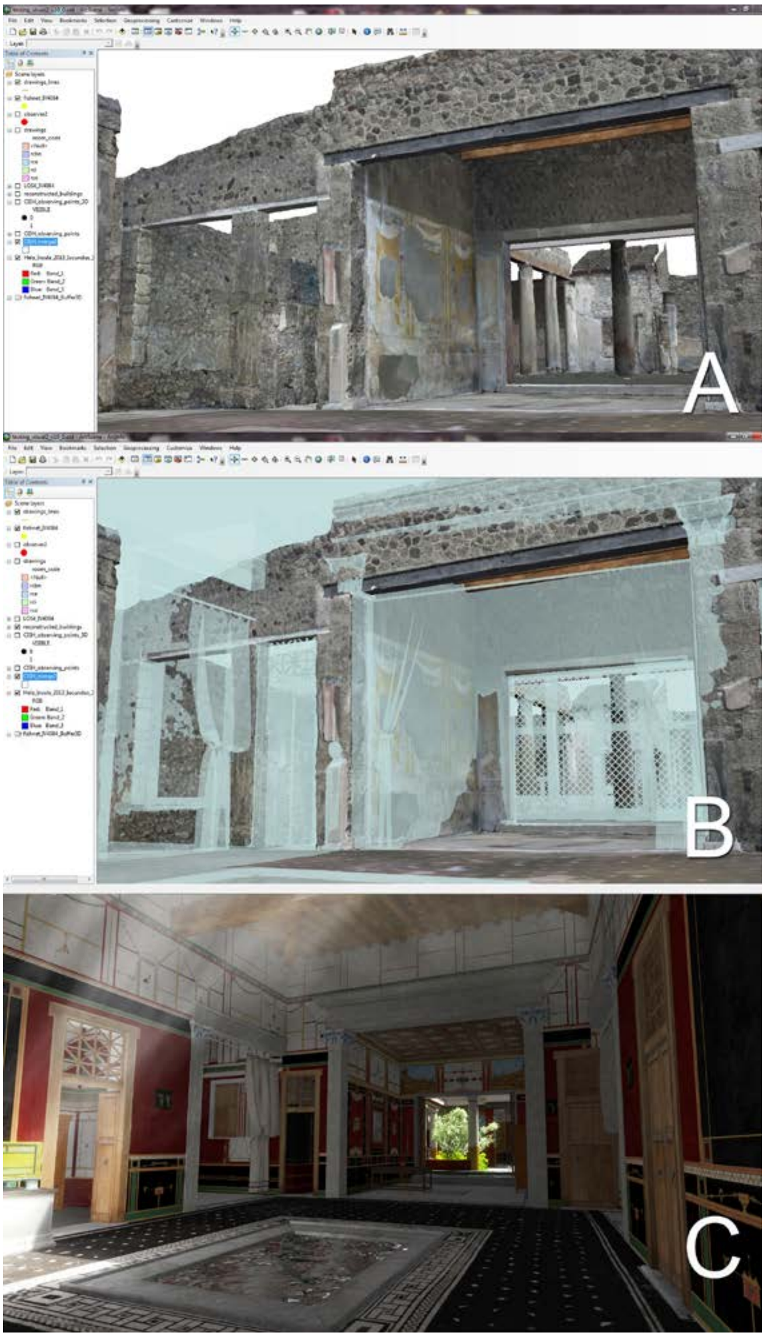


Figure 8: This image shows the result of the 3D interpretation of the house of Caecilius Iucundus and its implementation into the 3D GIS. In specific A) view of the original structures of part of the house implemented into ArcScene, B) view of the interpreted structures implemented and superimposed to the original structure, C) rendering of the house of Caecilius Iucundus fully reconstructed and visualized in Unity 3D.

The computer graphics-based reconstruction was developed using Autodesk 3DStudio Max, a 3D computer graphics program used to create 3D animations, models and images. The 3D modelling work was carried out by using the imported 3D model derived from the laser scanning survey as geometrical reference (Dell'Unto *et al.*, 2013). The volumes of the domus and the frescoes have been reconstructed following the construction lines of the scanned model. This approach allowed great accuracy and control during the modelling processes. The model was unwrapped and mapped with textures that simulate the wall decoration and the frescoes. The textures have been designed 'ad hoc', using photographic documentation for the existing frescoes still 'in situ', integrating the missing parts with information derived from archaeological evidence, historical photographs and water color representations, the 'model of Pompeii' (a 1:100 scaled model of the roman city preserved in the Archaeological museum of Neaples which shows the state of conservation at the end of the XIX century) and parallels found in the surrounding domus. The first 3D drafts were used as a basis for discussion and analysis for further interpretative decisions, in order to refine interpretation and reconstruction. This part of the workflow was crucial as it permitted to verify some hypotheses and to reject

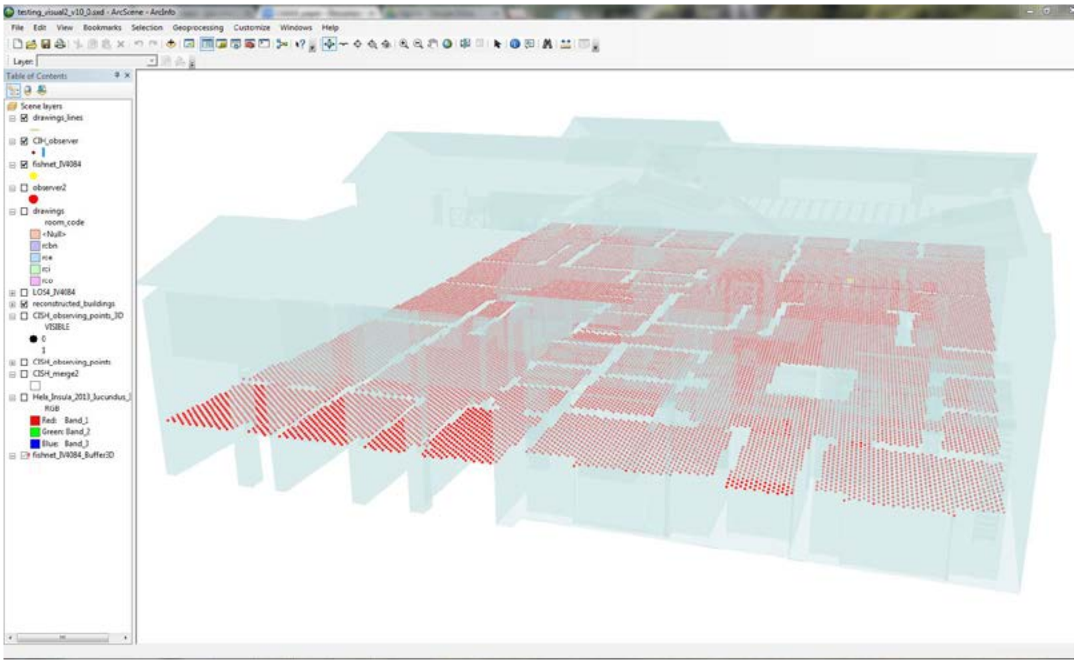


Figure 9: 3D grid of vector points (one spot per 20 cm) placed at a height of 1.60 m above the ground floor. Each one of these points, simulating the viewpoint of an original house's inhabitant, has been used as an origin point for the line-of-sight analysis targeting at the two different wall inscriptions.

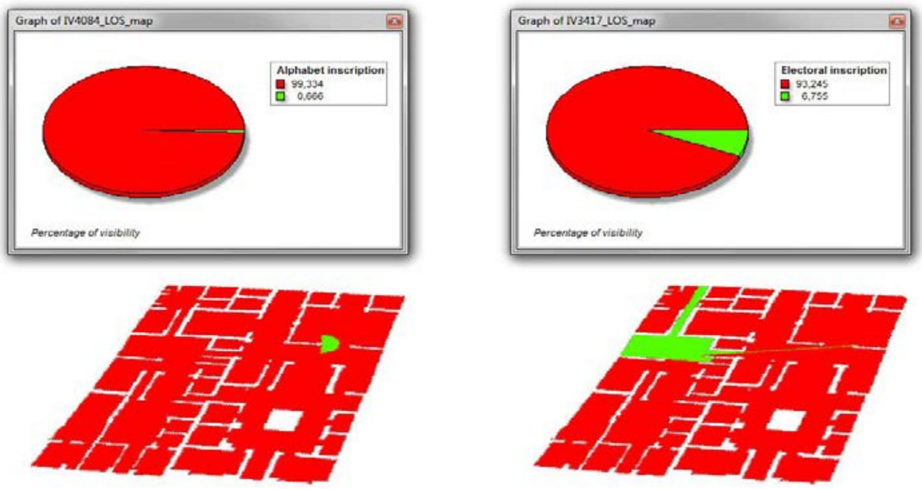


Figure 10: A quite different percentage of visibility between the alphabet and the electoral inscriptions has been detected. The original grid of observing points has been transformed in a raster map in which each cell expresses a boolean value of visibility (green visible/red not visible) compared to the selected targets. The alphabet inscription is visible to a very restricted number of viewers (0.6%), while the electoral one is definitely more exposed (6.7%).

others, proved wrong. The 3D models were not only the end result of the work, but also a scientific instrument for interrogating the architectures, and understanding their original shapes. The virtual reconstruction made of the Caecilus Iucundus house, as previously described (Dell'Unto *et al.,* 2013), has been imported into the 3D-GIS and used as a spatial reference for the visualscape analysis.

3.2 Visualscape Analysis

In recent years, some attempts have been made by integrating the use of GIS and 3D software for the analysis of the so-called 'visualscape' (Llobera 2003; Paliou 2013:1-4; Paliou and Knight 2013). The symbolic dimension of Roman domestic space constitutes an interesting case study to test 3D-GIS as an analytic tool to make a quantitative assessment of the cognitive structure related to specific objects (Lake and Woodman, 2003:694). As a test case to experiment this analytic approach, a couple of wall inscriptions, originally placed in two different rooms within the house of Caecilius Iucundus have been examined within the setting provided by the virtual reconstruction of the house. As a form of communication, wall inscriptions are among those symbolic objects that could be better ‘understood’ through a visual approach. There are major advantages to carry on this kind of analysis in fully-3D space. Firstly, the perspective of the ancient space is examined in such a way that the primary role in the analytic process is given the original inhabitants of the house and their potential viewing-lines. The 3D representation immediately reveals any visual obstacle that could have affected or impeded the view from a chosen position in the house towards the chosen target. In our case study, the choice of targets was decided by their difference and thereby in terms of what they reveal about their original audience of observers: the first target is a very small (0.9 x 10 cm) alphabet inscription, originally located on a column in the peristyle of the South House of Caecilius Iucundus. The second is a quite big electoral programma (30 x 100 cm), originally placed in a courtyard of the North House. In this simulation process a 3D vector grid (1 point per 20 cm) of sampled observing points was placed in ArcScene at a height of 1.60 meters above ground floor (Fig. 9) – a height based on the estimated average stature of ancient inhabitants of the Vesuvian area (Feemster Jashemski and Meyer 2002:455).

Next, the use of a line-of-sight algorithm provided an effective means to calculate which portions of the wall inscriptions were visible from the specified viewpoints and which visual obstacles prevented targets from being viewed. Not surprisingly, the final analysis showed us quite different percentages of visibility by comparing the alphabet inscription and the electoral programma (respectively 0.6% vs. 6.7%), reflecting their different purposes (Fig. 10).

Thanks to the combined use of line-of-sight and intervisibility algorithms, available with the latest version of ArcGIS (ESRI 2014), it is now possible to perform a fully-3D analysis inside a GIS environment. The reliability of this kind of analysis is further improved by the quality of the 3D models imported in GIS and used as basis to provide the ‘original’ spatial context in which the symbolic objects were placed. Basically, despite that this was just a case study (with just two examples on which visual analysis was tested), the results are encouraging and it would be useful to extend this methodological approach on a large scale sample (in order to detect some significant statistical patterns, it is better to have a high quantity of data to test). At least three distinct objectives could be achieved: (i) to formulate hypotheses about inscriptions’ original location; (ii) to generate insights on the most ‘suitable’ areas to view the inscriptions; (iii) to compare different categories of inscriptions in terms of their visual impact. Similarly to any other GIS project dealing with the landscape, it is crucial to identify patterns (Bevan and Connolly 2006; Chapman 2006) that could allow us to formulate some solid interpretations of the archaeological record. To reach this point it is necessary to extend the scale of analysis to a wider sample of data. In this sense, part of the ongoing research activity is targeted at identifying the original spatial distribution of wall inscriptions inside the Caecilius Iucundus’ house and its implementation in 3D-GIS. Then, most of these data will be available to be analyzed by means of visual analysis tools so as to produce new insights about any possible cognitive pattern and visual connectivity detected inside this Pompeian house.

4. Final Remarks

The case of the Pompeian house of Caecilius Iucundus clearly shows how the third dimension, used as an additional exploratory field, can dramatically increase the analytical potential of GIS. The different achievements in the documentation process outlined, increase what Gillings and Goodrick (1996) described as being one of the ultimate scopes of GIS: meeting the unique demands of archaeology problematic. As Goodchild recalls (1995), GIS systems have been used by archaeologists mainly as a mapping tool, and 3D has often been introduced in the frame of the archaeological projects with the exclusive purpose of improving the qualitative experience of a user in terms of visualization (Landeschi and Carrozzino, 2011; 2013). As Frischer still notes (2008), a sort of separation seems to characterize the domains of GIS and 3D in archaeology, with GIS users focusing on the application of tools for spatial analysis and 3D specialists more concerned with data visualization. In this regard, part of the project was aimed at overcoming this methodological divide in order to integrate the potential of GIS in conjunction with 3D technology so as to improve and enhance the overall analytic capacity. the implementation of high quality models along with a virtual interpretation provided in the form of a 3D model of the reconstructed house based on the information coming from the acquired models (Dell’Unto *et al.,* 2013) is a notable achievement in terms of data accuracy. It allows us to define a well documented and clear awareness of the degree of uncertainty in the reconstruction process which is the essential basis for further research lines, such as cognitive and visual analysis. According to Lake and Connolly (2006:8-10), the integrated use of GIS and 3D can provide a ‘localized experience of past material conditions’. Remarkably, in the framework of the case study presented above, the 3D experience has been contextualized so as to provide not only a mimetic representation of reality but also an effective means for depicting the dynamic complexity of a past social landscape (Gillings and Goodrick, 1996). This kind of approach can generate new insights into the phenomenological study of the ancient space, by considering it not just as a neutral backdrop of action (Tilley 1994:7-11) which would have been hardly recognized by the original inhabitants (Connolly and Lake 2006:8) but the means through which events and activities actually took place. The high level of precision of the acquired 3D models implemented in GIS provides the opportunity to enhance the reliability of the reconstructed ancient space (in our case the Roman house) reconstructed based on this dataset of information. In turn, this reconstruction is an essential premise to develop a work methodology that allows archaeologists to explore the cognitive dimension of the ancient space, a dimension that always needs to be based on a careful analysis and examination of the archaeological record (Renfrew 1993:259). In this respect Merlo (2004) stresses the importance of considering the concept of 'contemporary mind' where any possible bias in the process of understanding the ancient perception of space is due to the a priori interpretative paradigms of the archaeologists, which are strongly affected by the way digital technology has been used in the analysis of the research context. In this regard it is crucial to be very clear about the sources employed in the virtual reconstruction process, so as to state different degrees of reliability connected to each single element inside the virtual space. Accordingly, the Caecilius Iucundus residence, made at a former stage of the project in the GIS, was reconstructed with this specific purpose in mind (Dell'Unto *et al.,* 2013) (Fig 11). Remarkably, this research line opens up to innovative research paths in the frame of computer-based simulation within an ancient space. The technological developments previously mentioned could partially satisfy the demand for a more 'sensual' and 'phenomenological' approach (Brück, 2005:51-64; Shanks and Tilley, 1992:103-115). Considering the social significance intrinsically embedded in the ancient space of the Roman buildings (Allison 1997; Foss, 1997; Grahame, 1997; George, 1999), the use of advanced visual analysis tools can significantly improve our understanding of any symbolic content connected to specific categories of objects. These are remains with unquestionable symbolic value, which still may be associated with their original spatial location and which may be examined in new cognitive depth thanks to the use of 3D-GIS analytic tools.

Acknowledgements

This research activity was funded by the Swedish Research Council Grant (340-2012-5751) Archaeological information in the digital society (ARKDIS), by the Birgit ans Sven Håkan Ohlssons Foundation, by Fondazione Famiglia Rausing, by C.M. Lerici Foundation and by Marcus and Amalia Wallenberg Foundation. The authors would like to thank Karin Lundqvist and Renée Forsell for their precious consultancy

References

ALLEN, P., FEINER, S., TROCCOLI, A., BENKO, H., ISHAK, E., AND SMITH, B. (2004). Seeing into the past: creating a 3D modeling pipeline for archaeological visualization. In: *Proceedings of the 3D data Processing, Visualization, and Transmission, 2nd International Symposium on 3D Data Processing, Visualization, and Transmission (3DPVT'04).* Thessalonika: IEEE Computer Society, pp. 751-758.

ALLISON, P.M., (1997). Artefact distribution and spatial function in Pompeian houses. In: B. Rawson and P. Weaver, (Eds.). *The Roman family in Italy: Status, sentiment and space*. Oxford: Clarendon Press, pp. 321-354.

BARCELÓ, J.A., DE CASTRO, O., TRAVER, D., AND VICENTE, O. (2003). A 3D model of an archaeological excavation. In: M. Doerr, and A. Sarris, (Eds.). *The Digital Heritage of Archaeology. Computer Applications and Quantitative Methods in Archaeology, Proceedings of the 30th Conference, Heraklion,Crete, April 2002*. Greece: Hellenic Ministry of Culture, pp. 85-87.

BEVAN, A. AND CONOLLY, J. (2006). Multi-scalar approaches to settlement pattern analysis. In: G. Lock, and B. Molyneaux, (eds.). *Confronting Scale in Archaeology: Issues of Theory and Practice*. New York: Springer, pp. 217-234.

BRÜCK, J. (2005). Experiencing the past? The development of a phenomenological archaeology in British prehistory. *Archaeological Dialogues* 12(1), pp. 45–72.

CALLIERI, M., DELL'UNTO, N., DELLEPIANE, M., SCOPIGNO, R., SODERBERG, B. AND LARSSON, L. (2011). Documentation and Interpretation of an Archeological Excavation: an Experience with Dense Stereo econstruction Tools. In: *VAST The 11th International Symposium on Virtual Reality Archaeology and Cultural Heritage.* Eurographics, pp. 33-40.

CHAPMAN, H. (2006). *Landscape Archaeology and GIS.* Stroud: Tempus.

CONOLLY, J., AND LAKE, M. (2006). *Geographical Information Systems in Archaeology.* Cambridge Manuals in Archaeology. Cambridge: Cambridge University Press.

DELLEPIANE, M., DELL'UNTO, N., CALLIERI, M., LINDGREN, S. & SCOPIGNO, R. (2013). Archeological excavation monitoring using dense stereo matching techniques. *Journal of Cultural Heritage* 14, pp. 201-210.

DELL'UNTO, N. (2014). The Use of 3D Models for Intra-Site Investigation in Archaeology. In S. Campana, and F. Remondino, (Eds.). *3D recording and modeling in archaeology and cultural heritage. Theory and best practices.* BAR international series Oxford: Archaeopress.

DELL'UNTO, N., LANDESCHI, G., FERDANI, D., DELLEPIANE, M., CALLIERI, M., LINDGREN, S., LEANDER TOUATI A.-M. (2015). Experiencing Ancient Buildings from a 3D GIS Perspective: A Case Drawn from the Swedish Pompeii Project. *Journal of Archaeological Method and Theory*. 28 Jan 2015 DOI 10.1007/s10816-014-9226-7

DELL'UNTO, N., FERDANI, D., LEANDER TOUATI, A., DELLEPIANE, M., CALLIERI, M. AND LINDGREN, S. (2013). Digital reconstruction and visualization in archaeology Case-study drawn from the work of the Swedish Pompeii Project. In: A. Addison, G. Guidi, L. De Luca, and S. Pescarin, (eds.). *Proceedings of 2013 Digital Heritage International Congress*, UNESCO/ IEEE/EG, pp. 621-628.

DONEUS, M. AND NEUBAUER, W. (2005). Laser scanners for 3D documentation of stratigraphic excavations. In: E. Baltsavias *et al.* (ed.) *Recording, Modeling and Visualization of Cultural Heritage. Proceedings of the International Workshop,* Ascona, May 22-27, London: Taylor and Francis.

DONEUS, M., G.J.J. VERHOEVEN, M. FERA, C. BRIESE, M. KUCERA & W. NEUBAUER. (2011). From deposit to point cloud: a study of low-cost computer vision approaches for the straightforward documentation of archaeological excavations. In: *Geoinformatics 6 XXIIIrd International CIPA Symposium*, Prague: Faculty of Civil Engineering, Czech Technical University, pp. 81–88. Available from: http://geoinformatics.fsv.cvut.cz/pdf/geoinformatics-fce-ctu-2011-06.pdf [Accessed 27 July 2014].

ESRI (2010). *What's new in ArcGIS 3D Analyst 10.* [Online]. Available at: http://help.arcgis.com/en/arcgisdesktop/10.0/help/index.html#//00qp0000000z000000.htm [Accessed 27 July 2014].

ESRI (2012). *Introduction to editing in 3D* [Online]. Available at: http://help.arcgis.com/en/arcgisdesktop/10.0/help/index.html#//00q80000003n000000. [Accessed 27 July 2014].

ESRI (2012). T*he Multipatch Geometry Type - An ESRI White Paper December 2008.* [Online]. Available at: http://www.esri.com/library/whitepapers/pdfs/multipatch-geometry-type.pdf. [Accessed 27 July 2014].

ESRI (2013). *Working with ArcGlobe and ArcScene.* [Online]. Available at: http://resources.arcgis.com/en/help/main/10.1/index.html#//00q8000000sv000000. [Accessed 27 July 2014].

ESRI (2014). *Intervisibility (3D Analyst)* [Online]. Available at: http://resources.arcgis.com/en/help/main/10.2/index.html#//00q9000000nn000000 [Accessed 27 July 2014].

FEEMSTER JASHEMSKI, W. AND MEYER, F.G. (2002). *The natural history of Pompeii.* Cambridge; New York: Cambridge University Press, 2002.

FORTE, M., DELL'UNTO, N., ISSAVI, J., ONSUREZ, L. AND LERCARI, N. (2012). 3D Archaeology at Çatalhöyük. *International Journal of Heritage in the Digital Era*, 1(3), pp. 351-377.

FOSS, P. (1997). Watchful Lares: Roman Household Organization and the Rituals of Cooking and Dining. In: R. Laurence and A. Wallace-Hadrill (eds.). *Domestic Space in the Roman World: Pompeii and Beyond.* Portsmouth, Journal of Roman Archaeology Supplementary Series 22, pp. 196-218.

FOSS, P.W. AND DOBBINS, J.J. (2007). *The World of Pompeii*. London: Routledge.

FRISCHER, B. (2008). From Digital Illustration to Digital Heuristics. In B. Frischerand A. Dakouri-Hild (eds.). *Beyond Illustration. 2D and 3D Digital Technologies as Tools for Discovery in Archaeology*. Oxford: Archaeopress, BAR International Series 1805, pp. 5-24.

GEORGE, M. (1999). Repopulating the Roman House. In B. Rawson and P. Weaver (eds.), The Roman Family in Italy: Status, Sentiment, Space. Oxford University Press, pp. 299-319.

GILLINGS, M. and GOODRICK, G.T. (1996). Sensuous and reflexive GIS: exploring visualization and VRML. *Internet Archaeology* 1. [Online]. Available at: http://intarch.ac.uk/journal/issue1/gillings_index.html. [Accessed 27 July 2014].

GOODCHILD, M.F. (1995). Geographic Information Systems and geographic research. In J. Pickels, (Ed.). *Ground truth*. The Guilford Press: New York, pp. 31-50.

GRAHAME, M. (1997). Public and private in the Roman house: the spatial order of the Casa del Fauno in R. Laurence and A. Wallace-Hadrill (Eds.). *Domestic Space in the Roman World and Beyond.* Journal of Roman Archaeology, Supplementary Series, 22, pp. 137-164.

KOEHL, M. AND LOTT, C. (2008). *De la maquette 3D au SIG 3D pour la gestion du patrimoine archéologique. SIG 2008.* [Online]. Available at: http://www.esrifrance.fr/sig2008/insa_strasbourg.htm. [Accessed 27 July 2014].

LAKE, M.W., AND WOODMAN, P.E. (2003). Visibility studies in archaeology: a review and case study. *Environment and Planning B: Planning and Design* 30(5), pp. 689 – 707.

LANDESCHI, G. AND CARROZZINO, M. (2013). A 3D model for detecting and communicating the archaeological risk. In: F. Contreras, M. Farjas, F.J. Melero, (Eds.). *Proceedings of the 38th Annual Conference on Computer Applications and Quantitative Methods in Archaeology – Granada.* Oxford: Archeopress, pp. 295-297.

LANDESCHI, G., AND CARROZZINO, M. (2011). Communicating Archaeological Risk with Web-Based Virtual Reality: A Case Study. Internet Archaeology 31. [Online]. Available at: http://intarch.ac.uk/journal/issue31/landeschi-carrozzino_index.html. [Accessed 27 July 2014].

LEANDER TOUATI, A.M. (2010). Water, well-being and social complexity in Insula V 1. A Pompeian city block revisited. *Opuscula* 3, pp. 105-161.

LLOBERA, M. (2003). Extending GIS-based visual analysis: the concept of visualscapes. International *Journal of Geographical Information Science* 17 (1), 25-48.

LOSIER, L.,M., POULIOT, J., AND FORTIN, M., (2007). 3D geometrical modeling of excavation units at the archaeological site of Tell 'Acharneh (Syria). *Journal of Archaeological Scienc*e 34, 272-288.

MERLO, S. (2004). The 'contemporary mind'. 3D GIS as a challenge in excavation practice. In: K.F. Ausserer, W. Brner, M. Goriany, and L. Karlhuber-Vckl, (Eds). *Enter the Past. The E-way into the four Dimensions of Cultural Heritage. CAA 2003, Computer Applications and Quantitative Methods in Archaeology*. BAR International Series 1227, Oxford: Archaeopress, pp. 276-279.

OPITZ, R., AND NOWLIN, J. (2012). Photogrammetric modeling + GIS. Better methods for working with mesh data. ArcUser. [Online]. Available at: http://www.esri.com/news/arcuser/0312/files/archaeology-inventory.pdf. . [Accessed 27 July 2014].

PALIOU, E., AND KNIGHT, D.J. (2013). Mapping the senses: Perceptual and Social aspects of Late Antique Liturgy in San Vitale, Ravenna. In: F. Contreras, M. Farjas and F.J. Melero (eds.). *Proceedings of the 38th Annula Conference on Computer Applications and Quantitative methods in Archaeology*, *Granada Spain, April 2010,* Oxford: Archaeopress, BAR International Series 2494, pp. 229-236.

PALIOU., E. (2013). Reconsidering the concept of visualscapes: Recent advances in three-dimensional visibility analysis. In A. Bevan, and M. Lake, (Eds). *Computational Approaches to Archaeological Spaces.* Walnut Creek: Left Coast Press, pp. 243-263.

RENFREW, C. (1993). Cognitive archaeology: some thoughts on the archaeology of thought. *Cambridge Archaeological Journal* 3(2), pp. 248-250.

STAUB, T. (2009). Decorative effects and room functions. Evidence of thresholds studied in the residential quarters of Insula V 1, Pompeii. *Opuscula* 2, pp. 205-218.

TILLEY, C. (1994). *A Phenomenology of Landscape: Places, Paths and Monuments*. Oxford: Berg.

ZEILER, M. (2010). *Modeling Our World: The ESRI Guide to Geodatabase Concepts*, ESRI Press.

MAPPA Open Data Metadata. The Importance of Archaeological Background

Francesca ANICHINI, Gabriele GATTIGLIA

MAPPA Lab - Dipartimento di Civiltà e Forme del Sapere - University of Pisa (MAPPA), Italy
f.anichini@arch.unipi.it, g.gattiglia@arch.unipi.it

Abstract

The MOD (MAPPA Open Data), the first Italian repository of open archaeological data, was conceived by the MAPPA project for making archaeological data easily accessible to everyone and for any need. The MOD allows users to download the raw documentation and the grey literature of archaeological interventions.

The MOD Metadata schema is based on Dublin Core and on ISO 19115 schemas. Each archaeological intervention is described according to a schema that defines the history of the archaeological intervention, the sources used for creating the dataset, the method and the structure of the data and the physical data relations. A particular relevance in the schema is given to the description of the methodological background. This part of the metadata schema is fundamental to translate to future generation of archaeologists the subjective part of the archaeological record, because only the comprehension of the methodological background permit a real semantic interoperability.

Keywords: Open Data, Metadata, Dublin Core, Methodological background, MAPPA Project

1. Introduction - Opening up Italian archaeological data: a cultural problem

The MAPPA project set itself a highly innovative aim (Anichini et al., 2012; Anichini et al. 2013), quite revolutionary for Italian archaeology: lifting out of the archives the documents containing the data of archaeological investigations, both the raw documentation and the grey literature, and making this information easily accessible to everyone and for any need: from research to protection, from urban planning to quality tourism, or even just for simple learned curiosity. This is how the MOD (MAPPA Open Data, http://www.mappaproject.org/mod), the first Italian repository of open archaeological data was conceived in keeping with European directives regarding easy access to Public Sector data, and to Research data.

Originally conceived to collect the documentation of the excavations carried out in Pisa, the MOD considerably grew as the months went by (Anichini and Gattiglia, 2012; Anichini, Ciurcina and Noti, 2013). Also in consideration of the results of a survey promoted by the MAPPA project on 'Open data and Italian archaeology' (Anichini, 2013), which showed how the need to share data is strongly felt by the majority of the archaeological community, the MOD, a little at a time, has become the open data repository of Italian archaeology, where archaeologists can publish excavation data, and in 2013 it entered in the list of recommended repositories of the Journal of Open Archaeological Data. In fact entering archaeological documentation in the MOD is a publication to all effects, whose authorship is protected by a DOI (Digital Object Identifier) code and a CC BY or CC BY SA license.

At first it was decided to create a 1 star repository and to transform it step by step to a 5 star one. The rationale beneath this choice was the fact that data and the possibility to circulate and spread them are the key infrastructure of the world of archaeology. Well aware that part of the (interpretative) information underlying data is often connected to the know-how of each single researcher, the MOD stemmed exactly from this precondition as well as from the certainty that data are the structure upon which historical-archaeological interpretations are based and from the essential need for data to freely and rapidly circulate in the archaeological community. The guiding philosophy was based the ecological cycle of the data produced every day regardless of their ultimate use. We are talking about the large amount of 'raw' data that make up the documentation of every archaeological intervention: context sheets, catalogues, maps, photographs, inventories, reports, etc., which are produced and deposited in ministerial archives, regardless of whether the investigation is subsequently the subject or not of a publication. Many data, therefore, are often never published or reused for other research. In the ecological cycle of data, instead, it is good practice to use the data as much as possible. Alongside 'scientific' publications, where the interpreted results of an investigation are reported, many new and often unexpected applications may be created, using the same data with different purposes or to reach different interpretative hypotheses. The turning point is essentially of a cultural rather than technological nature. Data sharing is one of the few paths to be taken today that can allow knowledge to progress without having to sustain huge costs. On the contrary, it is possible to optimise the large amount of data that are produced every day and are underused. ‘Opening’ also means ‘protecting’, so that widespread sharing, in a community that points towards the idea of historical archaeological heritage as a common value, can gradually ensure control, quality and preservation. To make this happen another fundamental cultural step needs to be taken: recognising intellectual authorship and the related copyrights of the person who produced the data. At the same time, it is also of extreme importance to overcome the concept of ownership, because archaeological data are not owned (as in the case of private property) by the archaeologist who produces them but are part of our collective heritage: of course, we are speaking about the raw data, not the subjective interpretations of those data, which obviously belong to their author. Without this awareness, it would be very difficult to overcome the mistrust and fear that archaeologists have of being deprived of the result of their work.

2. Basic choices

Defining open data is quite simple: their essential features are coded in Tim Berners Lee’s five stars and in the indications provided by the Open Knowledge Foundation. Open data must be:

- ***complete***, so that they can be exported and used both online and offline, reporting the specifications adopted;
- ***primary***, i.e. in a raw processing state, so that they can be integrated and aggregated with other digital resources;
- ***timely*** and ***accessible***, so that users can access the data quickly, using Internet protocols without the need for subscriptions, payments or registrations, and can transmit and exchange them directly via the web;
- ***machine–readable***, i.e. automatically processed by the computer, so that users are not forced to use proprietary softwares, applications or interfaces to carry out these operations;
- ***searchable***, fully ***reusable*** and that can be ***integrated*** to create new resources, applications, softwares and services, also for commercial purposes.

All these features must be ***permanent*** for the entire lifecycle on the web (Anichini and Gattiglia, 2012). The data must also be issued with licences that respect the above features and do not limit their reuse in any way. They can be associated with licences that require recognition of authorship, but this must be free of charge.

The first decision we took was to enter the data in the repository exactly as they were presented to us and so quickly place them on the web. Our decision was based on the fact that real battle to be fought in Italy was clearly to change archaeologists’ mentality, usually reluctant to share data. It was important, therefore, to make primary data begin to circulate as quickly as possible, regardless of their type. We decided to arrange the repository for any format of data, as long as they could be entirely downloaded. We did not impose a guideline for publication standards in the MOD but, on the contrary, left it to the authors to decide what to publish and how: a lot, a little or everything, in a more or less open format. Every archaeologist, therefore, could decide how to take part in the transformation process towards open Archaeology. Detailed indications were provided instead on quotation rules, types of licences used and legal requisites regarding the compliance of privacy and copyright laws (Ciurcina, 2013). Specific attention was devoted to providing user-friendly access and use of the repository

3. Inside the MOD

The repository was built by taking the Archaeological Data Service of York University as example. Starting from this model, we chose a very simple user interface which makes it easy to search, view and download the documentation available. Another consideration needs to be carefully considered. The documentation of an archaeological excavation represents the only trace of an operation that, given its nature, destroys the stratification it is investigating: the MOD guarantees that those documents and, therefore, the data contained in them have potentially endless digital life – i.e. the life of the web.

The MOD, hosted on GNU/Linux of the Centro Interdipartimentale di Servizi Informatici per l’Area Umanistica (CISIAU) of Pisa University, was designed by Valerio Noti on an Open Source LAMP technological platform using an Apache HTTP Server, PHP 5.x scripting language and MySQL Open Source relational database. This platform, fully proven worldwide for the development of IT repositories, guarantees adequate stability, security and performances for the project purposes. The database structure focuses on the single archaeological investigation, defined by a set of information (title, author, DOI, region, location, year, main contact, introduction, overview) and by connection to diverse data sources such as text documents, images, multimedia objects, database files, geographical data, etc. A repository of files, downloadable by users, is associated with the repository, which authorised operators can enter via upload functions. From an operating viewpoint, the application – which can be consulted using any browser compatible with W3C standards – is made up of two distinct segments: an administrative section with confidential access and a public section. The administrative section is used by operators to enter and change the database contents. Categories (e.g. Chronology) and sub-categories (e.g. Medieval, Roman, etc.) can be managed, and an evolved HTML editor can be used for editing the Introduction and Overview sections and for uploading the files in the repository. The public segment can be directly accessed from the MAPPA project website. It can be used to consult the repository, view the record of each single dataset, carry out free full-text or categorised searches and download the files from the repository. An advanced query section is also available where users can query different fields of the database (chronology, topics, year, author, title, type of file in the repository) and carry out searches by geographical area (at the moment by Region). Particular importance was given to the scalability of the application in view of the growth in size and the issues relating to the repository's digital lifecycle. The MOD was designed to ensure easy migration to other hardware or software platforms and to highly reduce future technological adjustment activities (Anichini, Ciurcina and Noti, 2013).

4. Metadata

So the new step was to provide the MOD with a standardised management of the metadata referred to the single dataset, defining a minimum set of information so as to guarantee correct use of the data. Every archaeological intervention is associated, through a metadata schema, with all the information regarding the intervention itself, the archaeographic production, and the structure and format of the digital data, following a pattern that describes the history of the intervention, the sources used, the method and the relationship with the physical data. Archiving entire archaeographic and archaeological datasets and making them readily available on the web is a novelty for Italian archaeology, which until now has dealt almost exclusively with the metadating of summary archaeological data (e.g. the repositories of CulturaItalia or SigecWeb). The MOD Metadata schema is based on Dublin Core and the ISO 19115 schemas, it uses both thesauri realised by the MAPPA project itself and thesauri from ICCD (Istituto Centrale per il Catalogo e la Documentazione – National Institute for Catalogue and Documentation). A particular relevance in the schema is given to the description of the methodological background of the archaeological intervention: who direct the intervention, in which year, with what kind of method and so on. This part of the metadata schema is fundamental to translate to future generation of archaeologists the subjective part of the archaeological record, because only the comprehension of the methodological background permit a real semantic interoperability.

The current version of the MOD proposes a schema comprising [the underlined elements are from the ISO 19115 core (M = Mandatory), (C = Conditional), (O = Optional); in italics the equivalent ICCD thesaurus]:

1. History of the investigation:

- 1.1 Dataset title (M) (title of investigation/dataset, free text),
- 001.2 Purpose of investigation (brief description of purpose and main results of the investigation),
- 1.3 Method (MAPPA thesaurus),
- 1.4Type of documentation: drawn, photographic, written, video, multimedia,
- 1.5 Geographical location:
  - 1.5.1 Region (*PVCR*),
  - 1.5.2 Province (*PVCP*),
  - 1.5.3 Municipality (*PVCC*),
  - 1.5.4.Address (*PVCI*),
- 1.6 Additional extent information for the dataset (vertical and temporal) (O):
  - 1.6.1 Chronological range, MAPPA thesaurus for all identified chronologies),
  - 1.6.2 Chronological period,
- 1.7 Principal Investigator/team,
- 1.8 Year (interval or single year of investigation).

2. Sources used to create the data:

- 2.1 Archives queried,
- 2.2 Cartography used for georeferencing,
- 2.3 Previous investigations in the investigated area.

3. Method and structure of data:

- 3.1 Dataset reference date (M) (Date of creation of dataset),
- 3.2 Dataset topic category (M),
  - 3.3 Data georeferencing (GAT):
  - 3.3.1Geographic location of the dataset (by four coordinates or by geographic identifier) (C),
  - 0 3.3.2 Spatial resolution of the dataset (O),
  - 0000000 3.3.3 Reference system (O),
- 3.4 Abstract describing the dataset (M),
- 3.5 List of files present and their content (name of file with extension, Distribution format (O), software used for creating the file, version, description, relations,
- 3.6 List of assigned identifiers,
- 3.7 List of codes used,
- 3.8 Thesauri,
- 3.9 Description of any conversion to other formats,
- 3.10 Staff (all staff components with their tasks in the production of the paper or digital dataset),
- 3.11 Dataset authorship (curatorship of the DOI),
- 3.11.1 Licence,
- 3.12 Dataset language (M).

4. Reports:

- 4.1 Bibliography,
  - 4.1.1 From,
  - 4.1.2 To,
- 4.2 Place of preservation of archaeographic documentation (*ICCD S+ region number; C+ municipality number*),
- 4.3 Place of preservation of finds (ICCD S+ region number; C+ municipality number),
- 4.4 On-line resource (O) (URL Dataset),
- 4.5 Metadata:
  - 4.5.1 Metadata language (C),
  - 4.5.2 Metadata character set (C),
  - 4.5.3 Metadata point of contact (M),
  - 4.5.4 Metadata date stamp (M).

5. Conclusions: the importance of archaeological background

We must be aware that data collection by archaeologists is partly subjective (although limited by the use of standardised procedures) and that only a full and accurate account of the methodological and scientific procedures used by archaeologists when constructing their data representation model will allow easier data integration and reuse. This means that the data collected by various researchers can be compared only by taking into account their intellectual history and individual background (Terrenato, 2006, p.19), and that it would be better to make them available timely, without seeking perfection, when the scientific community is in greater methodological harmony with whoever has produced the data (Gattiglia, 2009, p.56). From a semantic viewpoint, this means that the concept of absolute objectivity needs to be abandoned in an attempt to make subjectivity objective. Nevertheless, this requires the codification of data through scientifically shared procedures that ensure their future use. Objectivising subjectivity means explaining the intellectual and methodological background and the expertise used for digitally codifying the data. This process, called 'semantic interoperability', does not provide a basis for the 'technical' grouping of data and creates illusory superstandards that group existing standards (D'Andrea, 2006, p. 120). On the contrary, since it does not alter the formalisation of data adopted by each single researcher, it ensures the codification of information on more abstract and general formal models, capable of capturing the semantics inherent in the stored data. To allow semantic interoperability, it is necessary to record the motivations and circumstances regarding the creation of a digital source, the details of its origin, content, structure and of the terms and conditions applicable to its use, both in terms of a complex source (an entire dataset) and digital object (single file).These aspects are recorded (recording allows extensive and continuous use of data by the scientific community) by creating metadata which are used for recording how the data were formed, thus making information freely and correctly accessible, even across time and space, and simplifying search, localisation, selection and semantic interoperability operations.

References

ANICHINI, F. (2013). MAPPA Survey: gli Open Data nell'archeologia italiana. In: Anichini, F., Dubbini, N., Fabiani, F., Gattiglia, G. and Gualandi, M.L *MAPPA. Metodologie Applicate alla Predittività del Potenziale Archaeologico*. 2. Roma: Edizioni Nuova Cultura.

ANICHINI, F., CIURCINA, M. and NOTI, V. (2013). Il MOD: l'archivio Open Data dell'archeologia italiana. In: ANICHINI, F., DUBBINI, N., FABIANI, F., GATTIGLIA, G. and GUALANDI, M.L. *MAPPA. Metodologie Applicate alla Predittività del Potenziale Archaeologico*. 2. Roma: Edizioni Nuova Cultura.

ANICHINI, F., FABIANI, F., GATTIGLIA, G. and GUALANDI, M.L (2012). *MAPPA. Methodology Applied to Archaeological potentialPredictivity*. 1. Roma: Edizioni Nuova Cultura.

ANICHINI, F., DUBBINI, N., FABIANI, F., GATTIGLIA, G. and GUALANDI, M.L. (2013). *MAPPA. Metodologie Applicate alla Predittività del Potenziale Archaeologico*. 2. Roma: Edizioni Nuova Cultura.

ANICHINI, F. and GATTIGLIA, G. (2012). #MappaOpenData. From web to society. Archaeological open data testing. *MapPapers*.2/2012. pp. 54-56.

CIURCINA, M. (2013). *P*arere legale sul portale MAPPA Open Data. In: Anichini, F., Bini, M., Dubbini, N., Fabiani, F., Gattiglia, G., Ghizzani Marcia, F. And Gualandi, M.L. (eds.) *MapPapers* 4/2013. pp.87-106.

D'ANDREA, A. (2006). *Documentazione archeologica standard e trattemento informatico*. Budapest: Archeolingua.

GATTIGLIA, G. (2009). Open digital archives in archaeology. A good practice. *Archeologia e Calcolatori* 20 (2). pp. 49-63.

TERRENATO, N. (2006). Le misure del campione contano! Il paradosso dei fenomeni globali e delle ricognizioni locali. In: Mancassola, N. and Saggioro, F. *Medioevo, Paesaggi e Metodi*. Mantova: SAP.

A Simple Way to Formalize the Dating of Stratigraphic Units

Bruno DESACHY

UMR 7041 équipe archéologies environnementales
Ministère de la Culture et de la Communication, France
bruno.desachy@gmail.com

Abstract

Building the stratigraphic relative chronology from field observations is at present a quite clear intellectual task, thanks to Edward Harris (Harris 1979). However, the following step – placing stratigraphic units in the absolute time by means of dated materials, historical indications, and stratigraphic observations – is far from being always totally explicit. The frequently used notions of terminus post quem (TPQ) and terminus ante quem (TAQ) do not define all the aspects of this dating reasoning. In particular, the formation durations of the stratigraphic units are rarely taken into account in a quantified way. So, even the basic notion of 'dating a layer' may be ambiguous.

This paper briefly presents some developments (from previous works about formalized stratigraphic data process and tool - Desachy 2010, 2012) in order to formalize and to computerize this basic but no so clear process of absolute dating of the stratigraphic units.

Keywords: stratigraphy, chronology, absolute time.

1. From stratigraphic time to quantified time: intervals of inaccuracy

A stratigraphic unit is defined at first in a relative time. It necessary exists also in the absolute time: it has a beginning of formation, an end of formation, and a duration between these two endpoints. The formation duration may be seen as a segment on the absolute time line; the beginning and the end are just moments, i.e. points on this line. Another duration must be taken into account: the duration 'masked' by a stratigraphic relationship when there is a discontinuity of absolute time between two successive stratigraphic units; more exactly, between the end of the formation of the previous unit and the beginning of the formation of the next unit.

These two moments (beginning and end of the formation of a unit) and these two durations (formation duration of a unit, and duration between two successive units) are the absolute dating framework of a stratigraphic sequence. Most frequently, the archaeologists do not have direct indication about those moments and durations. However, it is always possible to define an interval of inaccuracy for each of these moments and durations.

The [TPQ, TAQ] interval is traditionally used by the archaeologists to date the deposits. According to the most common definition of the TPQ (as the date provided by the later dated find in the deposit), this time slot is actually an inaccuracy interval for the end of the formation of the considered unit (so that this end time cannot be older than the TPQ nor more recent than the TAQ), but not for the duration of this formation. Indeed, it is worth emphasizing that a TPQ provides an earliest limit for the end, and not for the beginning of the duration formation.

The same principle of interval of inaccuracy may be applied to the whole dating framework of a stratigraphic sequence. So, besides the [TPQ, TAQ] interval, we can consider 3 others intervals of inaccuracy: one interval [earliest date, latest date] including the beginning of the unit formation; and two intervals [shortest duration, longest duration] to include the unit formation duration, and to include the duration between two successive units. Those four intervals involve height limits in the absolute (quantified) time (figure 1):

– Be(x): earliest date for the formation beginning of a unit x;
– Bl(x): latest date for the formation beginning of a unit x;
– Ee(x): earliest date for the formation end of a unit x (corresponding to the common notion of terminus post quem);
– El(x): latest date for the formation end of a unit x (corresponding to the common notion of terminus ante quem);
– Ds(x): shortest possible duration for the formation of a unit x;
– Dl(x): longest possible duration for the formation of a unit x;
– Ds(x, y): shortest possible duration assignable to the order relation between units x and y;
– Dl(x, y): longest possible duration assignable to the order relation between units x and y;

Of course, all these quantified values must use the same time measuring unit (usually, this unit is the year).

2. Taking into account the diversity of the dating clues

As every archaeologist knows, different kinds of dating elements may be available to localize the stratigraphic units

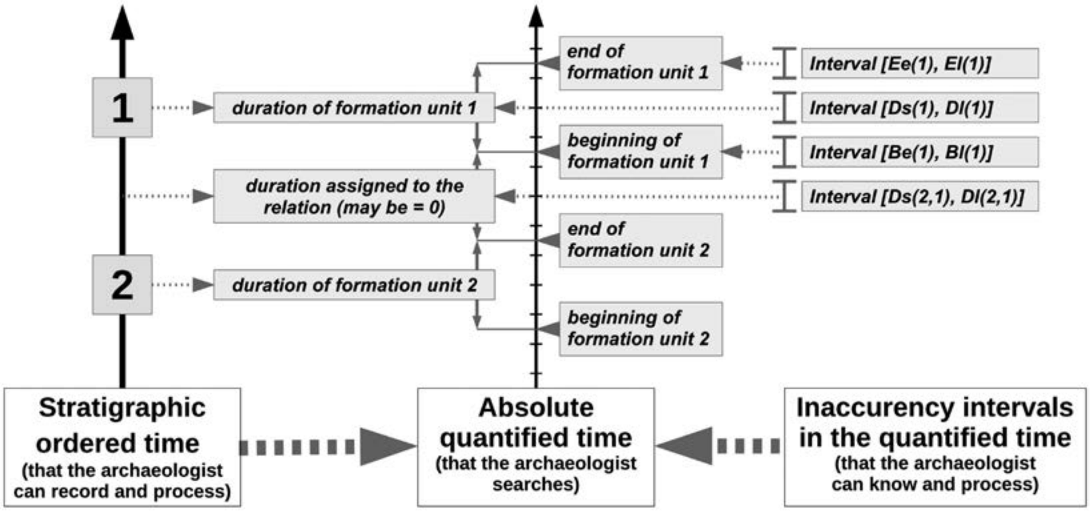


FIGURE. 1: INTERVALS AND ABSOLUTE QUANTIFIED TIME FOR STRATIGRAPHICAL UNITS.

in the absolute time. These diverse dating elements do not inform the same intervals and limits, according to their nature and to their position related to the considered unit (typological or analytical dating of finds contained in this unit, absolute time data about the unit formation provided by geophysical or geoarchaeological studies, extrinsic historical sources...).

Concerning specifically the artifacts found in archaeological context, there is an abundant literature about their chronological characteristics and their use to dating. Let us only remind that the moment when an artifact goes into the archaeological context (where the archaeologist will find it later) is the end time of a whole previous chronological path including production, use, sometimes reuse, discard or deposit, and sometimes (very often in the urban contexts for instance) transfers from a stratigraphic unit to a new unit disrupting the previous one. This trajectory must be taken in account if an artifact is used as absolute dating element (what is possible also by the means of inaccuracy intervals applied to the artifact chronological path). Of course, in all cases, it is necessary to eliminate the intrusive elements, which actually belong to later stratigraphic events.

More generally, archaeologists record carefully a lot of chronological evidences from the contexts they study, but these evidences are not always fully used in the absolute dating reasoning. The set of intervals exposed above allows to value all the dating clues available in order to date an archaeological context, whatever their nature; then, the question is 'what interval and what limit(s) does this clue inform?' For instance, field observations, possibly resorting to geoarchaeology or taphonomy, may provide evidences about the formation duration of a unit. Then, it may be necessary to translate a qualitative assessment (e.g. 'not very long') into a quantified order-of-magnitude (e.g. 'from one to ten years') usable as an interval [shortest possible duration, longest possible duration].

Sometimes the beginning, or the end, or the duration of the formation of a unit is directly known and dated (for instance if this unit is related to a known historical event). In this case the inaccuracy interval has a nil length and its two limits receive this same known chronological value.

3. Processing the inequations system

The intervals and the stratigraphic relations are the components of a system of inequations which can be automatically processed by the means of simple – but boring to execute manually – logical and algebra calculation. The system has the following form:

for each unit x,

$$Ee(x) - Dl(x) \le Be(x) \le Ee(x) - Ds(x)$$
$$El(x) - Dl(x) \le Bl(x) \le El(x) - Ds(x)$$
$$Be(x) + Ds(x) \le Ee(x) \le Be(x) + Dl(x)$$
$$Bl(x) + Ds(x) \le El(x) \le Bl(x) + Dl(x)$$
$$Ee(x) - Bl(x) \le Ds(x) \le Dl(x) \le El(x) - Be(x)$$

for each relation x before y,

$$Be(y) - Dl(x, y) \le Ee(x) \le Be(y) - Ds(x, y)$$
$$Bl(y) - Dl(x, y) \le El(x) \le Bl(y) - Ds(x, y)$$
$$Ee(x) + Ds(x, y) \le Be(y) \le Ee(x) + Dl(x, y)$$
$$El(x) + Ds(x, y) \le Bl(y) \le El(x) + Dl(x, y)$$
$$Be(y) - El(x) \le Ds(x, y) \le Dl(x, y) \le Bl(y) - Ee(x)$$

For each unit and relation, the unknown values (limits not informed by dating elements) may be completed: an 'absolute beginning' and an 'absolute end', chosen by the archaeologist so as to include the whole studied occupation, are the default values for the unknown limits at the earliest and at the latest; the difference between them gives the total maximal duration, default value for the unknown longest possible durations; as for the unknown shortest possible durations, the logical default value is zero.

Let us take an example: archaeologists found remains of wall, with a coin dated from 1600 BC lost in the masonry. The analysis of this masonry leads to conclude that its

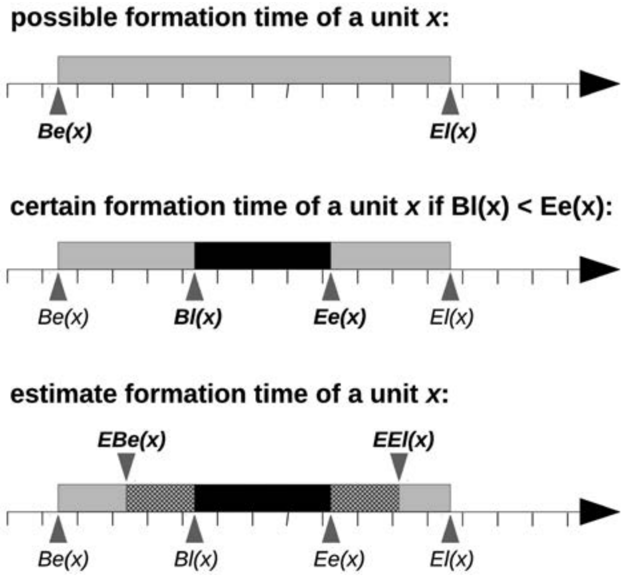


Figure 2: Uncertainity in the formation time of a stratigraphical unit.

building was quite fast, certainly less than one year. This masonry was set in a foundation trench, itself excavated and filled in a short time (deduced, for instance, from vertical edges in a loose ground), certainly also less than one year. Furthermore, an old photo and later documents indicate that in 1860 BC, this wall was no more visible, thus it was already destroyed. If we consider the two stratigraphic units: 'x' for the foundation trench, 'y' for the masonry of the wall, and the relation x before (filled by) y, we can write these limits (with year as time unit):

Dl(x) = 1
Dl(x, y) = 1
Ee(y) = 1600
El(y) = 1860
Dl(y) = 1

The aim of the process is to reduce the initial intervals by deducing increased earliest limits and shortest duration, and reduced latest limits and longest durations. So we can use only the corresponding parts of the inequations exposed above to deduce, as far as possible, the new more accurate limits:

Ee(y) – Dl(y) = Be(y) thus Be(y) = 1600 – 1 = 1599 (beginning at the earliest for the formation of y)
Bl(y) = El(y) – Ds(y) thus Bl(y) = 1860 – 0 = 1860 (beginning at the latest for the formation of y)
El(x) = Bl(y) – Ds(x, y) thus El(x) = 1860 – 0 = 1860 (end at the latest for the formation of x)
Be(y) – Dl(x, y) = Ee(x) thus Ee(x) = 1599 – 1 = 1598 (end at the earliest for the formation of x)
Bl(x) = El(x) – Ds(x) thus Bl(x) = 1860 – 0 = 1860 (beginning at the latest for the formation of x)
Ee(x) – Dl(x) = Be(x) thus Be(x) = 1598 – 1 = 1597 (beginning at the earlier for the formation of x)

The new deduced values provide absolute time intervals for each unit with the best (or the least bad) accuracy according to the whole available dating elements and stratigraphic relations. This deducing process uses the whole stratigraphic partial order, including transitively deductible relations, so that the deduced duration interval of each relation gives the critical path duration between the two units. Logical faults due to contradictory dating elements are detectable by the deduction of negative intervals of time.

4. Results: the times of a stratigraphic sequence and their display

The deduced intervals provide the units with two kinds of localization in the absolute time: a possible time zone, and/or a certain time zone (figure 2).

If a unit has inaccuracy intervals (with a length more than zero), this inaccuracy involves logically a time zone of possibility (and not certainty) of formation of this unit, between its oldest possible beginning date (Be) and its latest possible end date (El).

The earliest date for the formation beginning of a unit is necessary prior to the latest date for this beginning, but the latest for the beginning (Bl) may be later than the earliest date for the end (Ee). In this case, the inaccuracy is so strong that the unit has no certain time into its possible time of formation. Indeed, at each moment of its possible time, its formation may be not yet begun, or may be already ended. It is the case of the example above, and it is generally the case when only dated finds giving TPQ are available as dating elements.

If Bl(x) is earlier than Ee(x), the unit x has a partial time zone of certainty between these two limits. If its inaccuracy intervals have a nil length, the unit has a totally certain time of formation.

Displaying these intervals and time zones in the absolute time needs a specific graph. The Harris Matrix (and its variants) is the best and well-known way to display synthetically a stratigraphic sequence. However, it is actually a graph of the partially ordered set composed of the stratigraphic units of this sequence, what involves that the units are displayed as nodes, by only one point in the time. It is possible to show on such a graph either the TPQ or the TAQ (as horizontal thresholds, and assuming that the displayed point in time for each unit is the end of its formation – Desachy 2010), but this kind of graph can't practically show absolute durations. So, a complementary kind of graph, on a quantified time scale, is necessary to display the absolute time (figure 3). This sort of quantified time scale graph may be more or less detailed: showing all the inaccuracy intervals and dating elements for a single unit graph, or only the certain/possible time zones of each unit for a more synthetic view of the whole sequence.

If a shortest and/or a longest possible duration of formation (smaller than the possible time zone of the unit) is known, it may be displayed at the earliest or at the latest, showing

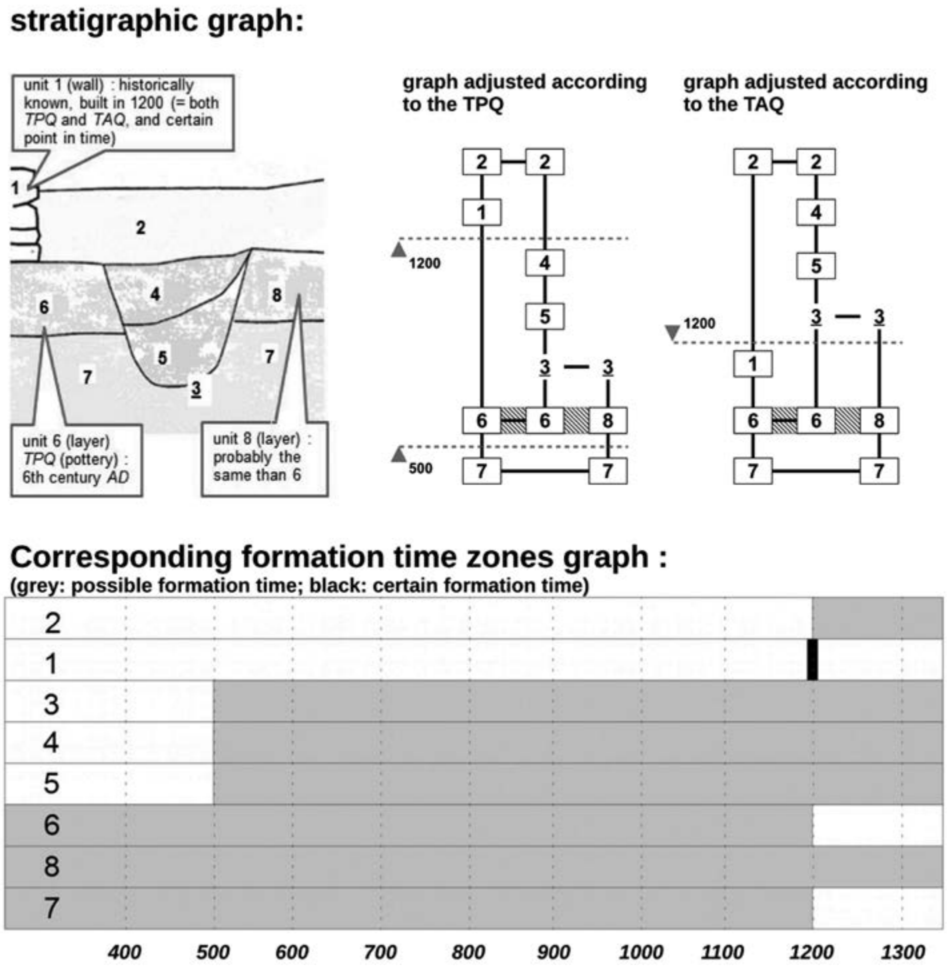


Figure 3: Quantified time scale graph for the formation time of a stratigraphical unit.

the two extreme possible chronological hypothesis. It must be adjusted according to the certain time zone, if this one exists (figure 4).

5. Uncertainty and estimate time

Inaccuracy is not the only kind of imperfection which affects the data: some stratigraphic relations may be uncertain, or less certain than others, when field observations provide significant clues but not totally decisive evidence. It was proposed (Desachy 2010) to take in account this difference of status through a modality of uncertainty assignable to the stratigraphic relations, so that a relation may be recorded as 'estimated' (not certain). This 'two-speeds' system of statements is only a very simple form of modal logic, but it allows to deal more flexibly with the qualitative heterogeneousness of the data.

Transforming this modal stratigraphic time into quantified time involves a double set of intervals and limits in the absolute time: the set of certain limits exposed above, and, possibly, a second set of 'estimated' limits and intervals, more accurate but less certain than the first one. These estimated intervals of inaccuracy may be used for less certain dating elements, or for chronological hypothesis testing.

If this double set of limits is used, it may be an estimated earliest date for the formation beginning of a unit x – EBe(x) – and an estimated latest date for the formation end of x – EEl(x), defining a third kind of time zone: an estimated formation time zone, contained into the possible time zone and including the certain time zone (if it exists) (figure 2)

6. Towards a more explicit chronological reasoning: limits and prospects

The formalization discussed above in terms of inaccuracy intervals and inequations system relates only to the basic – but fundamental – chronological reasoning from stratigraphic data and dating elements, that archaeologists practise every day to produce statements such as 'this feature dates from the second half of the 1st century BC'. Our formalization is an attempt to find the simplest way to make sounder and more explicit the production of this sort of statements. Indeed, a more explicit view of the contexts in the absolute time (in terms of certain, possible or estimate time) is useful, in particular to typochronological researches using the archaeological contexts to date artifacts. The aim is also, of course, a computerizable process, in order to save time and to get easily a chronological synthesis of available data as an archaeological excavation progresses.

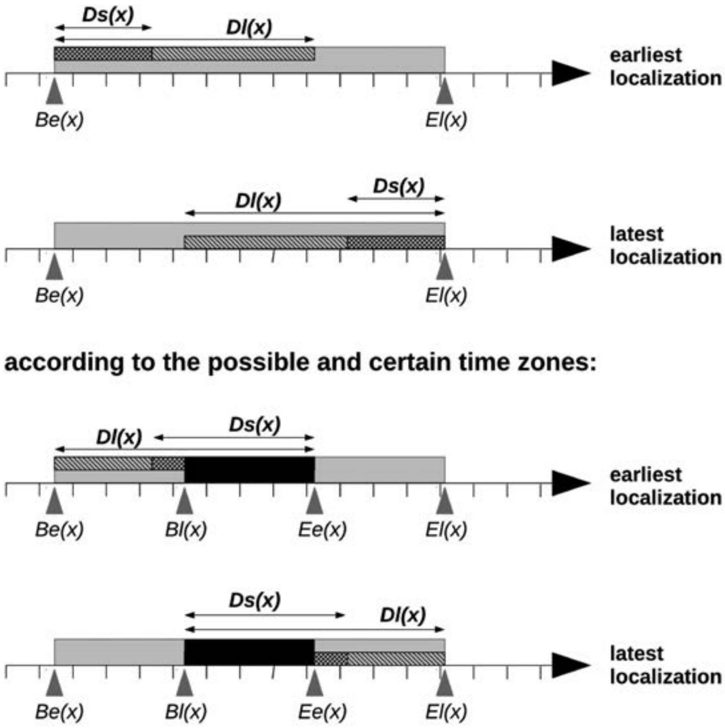


Figure 4: Chronological hypotheses for a stratigraphical unit, with the shortest and the longest formation durations.

Practically, the immediate prospects include the integration of this process into the next version of the computerized stratigraphic data processing tool *Le Stratifiant* (Desachy 2010).

This simplicity involves conceptual limits: into the 'possible time' defined by the inaccuracy intervals, even if chronological hypotheses may be displayed (by the way of 'estimated' time zones), there is no way to decide if a hypothesis is better than another one. The answer to this problem is a further step, of real chronological modelling, using more advanced mathematical tools (probability theory, fuzzy logic...).

Another limit, which is also a way to development, is that the process exposed above is applied only to the stratigraphic chronology, i.e. the succession of the archaeological ground formation steps. The same simple way of inaccuracy intervals process is applicable to the more synthetic and elaborate scale of the historic entities (who can, or not, arise from archaeological observations and in this case from one or several stratigraphic units) whose chronology is imperfectly known; but then, it will be necessary to move from the stratigraphic notion of formation duration of a unit, to a wider historical and functional notion of cultural existence duration (for instance, in this wider notion, an archaeological feature may have times of use and reuse after its time of formation); and to move from the strict notion of stratigraphic relationship to a wider notion of constraint of succession.

References

HARRIS, E.C. (1979). *Principles of Archaeological Stratigraphy*. London: Academic Press

DESACHY, B. (2010). Le Stratifiant: a simple tool for processing stratigraphic data. In: Börner, W., Uhlirz, S. (eds.). *Proceedings of the 14th International Congress 'Cultural Heritage and New Technologies' held in Vienna, Austria November 2009.* Vienna: Museen der Stadt Wien – Stadtarchäologie. p. 687-699. Available from: http://www.chnt.at/chnt-14-2009-proceedings/ [Accessed: 25th July 2014].

DESACHY, B. (2012). Formaliser le raisonnement chronologique et son incertitude en archéologie de terrain. *Cybergeo : European Journal of Geography*. [OnLine] Systèmes, Modélisation, Géostatistique, article 597. Available from: http://cybergeo.revues.org/25233 [Accessed: 25th July 2014].

Recognizing Temporalities in Urban Units from a Functional Approach: Three Case Studies

Julie GRAVIER

UMR 8504 - CNRS, Géographie-cités P.A.R.I.S
University of Paris 1, Panthéon-Sorbonne
ANR TransMonDyn, France
julie.gravier@parisgeo.cnrs.fr

Abstract

The city is a complex object and its descriptions by urban actors have evolved over time. Therefore many ways to investigate it have emerged. In this paper, we propose to study cities through the combination of units of analysis, called 'urban entity' and defined by a function, a form and a date on a continuous timeline. However, while the dating of items is often done in absolute time, the way it is determined is not always clearly formalized by archaeologists and historians. We present a formalization of the dates of urban entities, which take into account certainty and estimation. Each entity is defined by numeric dates so it is possible to explore the statistical distribution of urban entities through function and chronological periods, based here on 50 years (units). Then, the analysis and comparison of change and sustainability of three towns from the 1st to the 18th century may be systematized.

Keywords: Urban archaeology, multivariate statistics, long-term dynamics, absolute chronology.

1. Introduction

'Chronology is one of the first task of historian, which often serves as a framework for overall analysis'[1] (Offenstadt, 2006: 23). However, while the dating of items – in a broad sense – is often done in absolute time, the way it is determined is not always clearly formalized by archaeologists and historians. Consequently studies of change and sustainability of systems over the long term are frequently based on conventional time periods. One of the main questions of this paper is the analysis and synthesis of temporalities of urban spatial systems. It is established on the functions of space through a diachronic vision. The aim is to observe functional changes, paces and continuities, and it relies on three case studies. This subset of three towns: Beauvais, Noyon and Saint-Quentin located in Picardy (France), implies a comparative approach and allows the exploration of specificities and recurrences of urban trajectories (fig. 1).

Furthermore, the city is a complex object and its definitions by urban actors have evolved through periods of time. Therefore there are many ways to study it. For urban archaeology – which emerged during the 1960s in England – cities could be observed over long periods of time thanks to functions describing the physical urban space (Biddle, Hudson, Heighway, 1973). This approach is based on functional intra-urban units of analysis, such as abbeys, roads, handicraft activities, rivers etc. Then the combination of these entities allows to recompose and to analyze the occupation of a city and its changes.

[1] French quotation : 'La chronologie est un des premiers travaux de l'historien, qui sert souvent de cadre aux analyses d'ensemble.'

In France during the 1970s, this empirical process was applied by H. Galinié on the city of Tours (Galinié, Randoin, 1979). Since the 1980s, it has been afterwards systematized with monograph series titled Documents of Archaeological Evaluation of French Towns. Each monograph has been written by local archaeologists in collaboration with the National Center of Urban Archaeology, which has published the books.

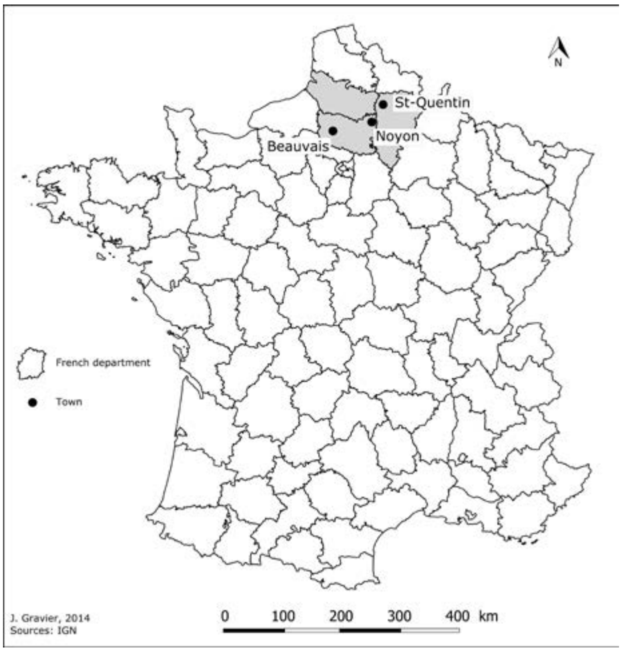


Figure 1: Location map of the three case studies: Beauvais, Noyon and Saint-Quentin.

Our approach inherits from these researches, and like these, it is based on the assumption that cities could be studied through units of analysis. In this paper, these units are named ‘urban entities’ and each one is defined by a function, a location (or a spatial form) and a date on a continuous timeline. Although our methodology is inspired by these works, we decided to add a dynamic aspect with the elaboration of an Archaeological Information System, as it has been done by others like the TOTOPI system (Galinié, Rodier, 2002). Thanks to computer science, the analysis of evolutions of urban spatial systems could be realized with a statistical approach and not only in an empirical way.

First we present the construction of urban entities. More specifically, it is an explanation of the processes of creating these entities from archaeological and historical documentation. Given that the study is diachronic, it involves difficulties in examining heterogeneous documentation. Next, we propose a way to systematize the examination of ‘source effects’ over time. Lastly, we focus on a statistical approach toward the repartition of urban entities through function and time, using correspondence analysis and hierarchical cluster analysis. These analyses provide a strong basis to identify functional profiles of cities from the 1st century AD – origins of the three towns – to the 18th century.

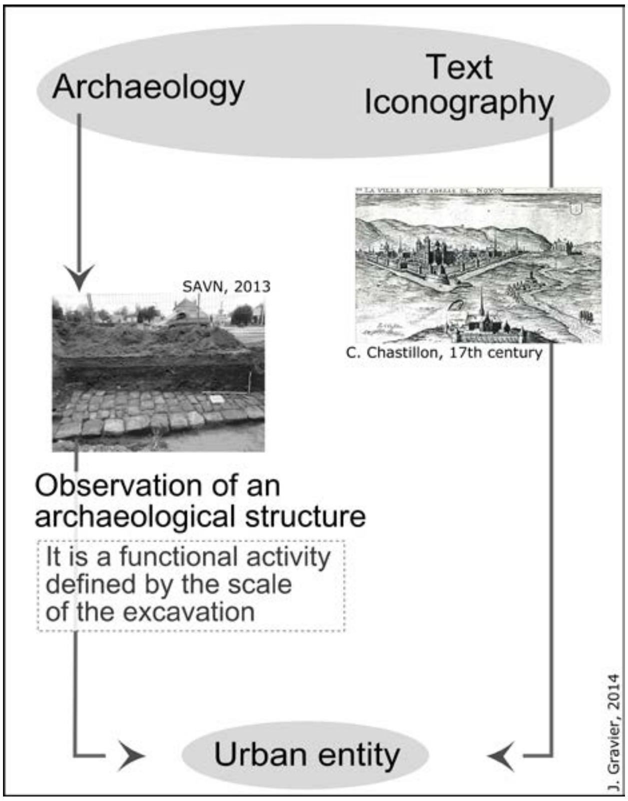


Figure 2: Simplified schematic of the data processing from archaeological remains, texts and iconography to urban entities.

2. What exactly is an urban entity? Methodological aspects of data processing

First, we propose to explain our interpretation of archaeological and historical documentation and what methodological choices were made for the processing of such data in the project. Each urban entity is based on archaeological remains, texts and iconography. However this documentation is in itself very heterogeneous, but also according to time periods and authors (Galinié, 2000). Therefore, the heterogeneity implies incomplete data and it is essential to take its uncertainty into account.

2.1. Archaeological remains, texts and iconography: from documentation to urban entity

Archaeology, text and iconography are studied as complementary documentation to fill the data gaps existing between each one. Nevertheless, the data processing of archaeological remains is distinguished from text and iconography, because the first are more precise compared to the latter since they are based on a material point of view (fig. 2).

Remains are first processed as an ‘observation of an archaeological structure’ (OAS), which represents a

Sections	Functions
road	**F1**. Road and free space
free space	
bank and river planning	**F2**. Bank, river, relief planning and crossing system
relief planning	
crossing system	
water supply	**F3**. Water supply and evacuation system
collecting, evacuation system	
urban defensive system	**F4**. Defensive and military system
fortify structure	
garrison, barracks	
public space	**F5**. Civilian and religious public spaces
civil authorities, justice	
spectacle, sport	
public bath	
education, culture	**F6**. Host establishment (health, education)
health	
private housing	**F7**. Habitat
pagan worship	**F8**. Religious establishment
catholic place of worship	
conventual or monastic building	
ecclesiastical building	
worship other than catholic	
funerary	**F9**. Funerary
trade, exchange	**F10**. Production, trade and craft
handicraft	
agriculture, breeding	
industry	
extraction	
natural formation	**F11**. Natural formation

Table 1: Thesaurus of the functions used to describe the units of analysis.

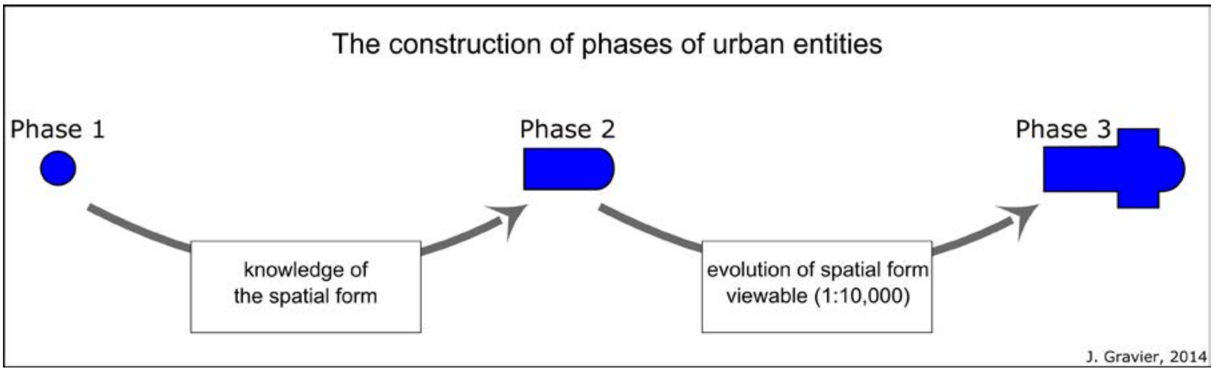


Figure 3: The construction of phases of urban entities: a typical example of the spatial change of a church thought time.

functional activity defined by the scale of the excavation. Each observation is described by a function, called a section, inherited from the thesaurus of the National Center of Urban Archaeology (tab. 1). This stage suggests a first inference from an observable fact – the archaeological remains – to the analysis of an activity of a past society.

In a second phase, the OAS can be used to create urban entities by cross-checking the information. As an example, a rescue excavation was done in 1980 in the city center of Noyon by J.-P. Angot, wherein he found remains of pillars and a semi-circular structure (Ben Redjeb et al. 1992). Eight years later, another excavation was done by M. Talon, located further northeast. Some important pillars, about 0.70 meters high, and structures linked to an hypocaust were discovered. Finally, two OAS in data processing were recorded, described by a bath section. However on a city-wide scale, these observations were located near a main road, next to a Roman market and other public buildings. Therefore we can consider these OAS as the same part of a construction, which were probably some public thermal baths. Therefore they have been registered as a single urban entity, and thus, this step represents a second inference in the data processing.

On the contrary, text and iconography are directly processed as urban entities because the observable fact of these documents is not the same as archaeology. Indeed, a mention in a text or a representation in an iconography present direct activities of past societies.

Each urban entity is defined by one of the functions proposed (tab. 1). This thesaurus is mostly inherited by the National Center of Urban Archaeology but reviewed for this project (Borderie et al. 2014). In the end, only eleven functions are taken into account since a city-wide scale study involves a generalization of the information. However some archaeological findings, despite being punctual knowledge, can be registered as urban entities if they are considered essential for the understanding of the town. What matters is the categorization and selection of the information that allows to grasp urban dynamics.

2.2. Spatial processing of urban entity

Besides a function, an urban entity is composed of a location or a spatial form. Sometimes an urban entity can be represented by a punctual location if it is not visible at the urban scale, considered for this purpose on a scale of 1:10,000; but it is also seen as a punctual location if the form of an urban entity is unknown. Moreover the form of an entity can evolve through time. Therefore in the data processing of the spatial form, we have decided that it is appropriate to translate these specificities as 'phases' of urban entities (fig. 3).

This figure presents a theoretical example of a church to understand the formalization of these possible phases. This church is first mentioned in a text in 1240. At this date (phase 1, fig. 3), its existence and location are known, but not its form. However, by means of a rescue excavation the spatial form was recognized during the 14th century (phase 2, fig. 3). The knowledge of the form leads to a modification of phase of urban entity. A transept was added to the church during the 15th century (phase 3, fig. 3). In the process of managing change, this new phase of the urban entity represents an evolution of the spatial form visible on urban scale.

2.3. Processing of time of urban entity

Each entity exists on a timeline. In fact, it has a beginning, a duration of existence, and an ending (A, fig. 4). However due to the lack of archaeological or textual data, the exact moment of the start date or of the end date are often not known. In most cases, they are discerned with intervals of time (B, fig. 4).

- On one hand, the dates indicate the certainty of existence of the urban entity (B, fig. 4). For example, the first mention of a hospital in a text provides this date for the beginning of the entity. At this exact moment the hospital does exist, but it has existed previously.
- On the other hand, they reveal the certainty of inexistence of the entity (B, fig. 4). For example, the ending of the existence of a street at Noyon is known by maps. On a first map, dated on 1889, the street is represented. At this point, it is sure that the entity continued to exist. However on another map dated on 1929 the street is not visible: it did not exist anymore at this time. Consequently, the ending of this urban entity is between 1889 and 1929.

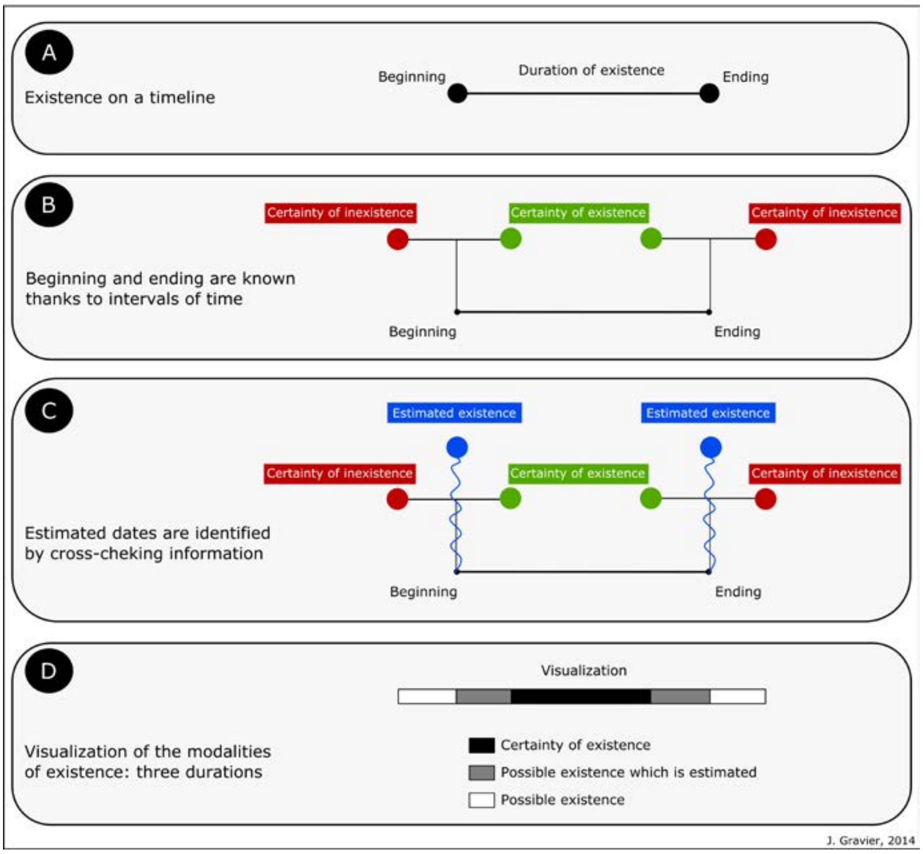


FIGURE 4: SCHEMA OF THE FORMALIZATION OF THE DATES OF URBAN ENTITIES.

- To address one of the main purpose of archaeologists and historians, which is chronology of events, we propose the notion of estimated dates (C, fig. 4). In the example above, we know that the district where the street was had been destroyed by the French as they took control of the city during the First World War. So at this stage, the ending of the street can be estimated to be 1917. Therefore, the relations between urban entities and the general analysis of the documentation allow the estimation of dates inside intervals of certainty.

Moreover, it is interesting to note that the dates of certainty of existence often are revealed by texts and iconography, while those of certainty of inexistence are mostly indicated by archaeological remains. In fact, mentions and representations provide the proof of the existence of an urban entity.[2] But frequently, the dates known from remains come from analysis of the archaeological material discovered in later layers. This is the case for layers coming from postholes, storage silos etc. Therefore archaeological dates often signal the moment when the activity is over and, thus, reveal the certainty of inexistence of an urban entity.

Finally, managing uncertainty occurs within the data processing thanks to six different dates and this formalization of dates of urban entities can be viewed with three modalities of duration (D, fig. 4).

[2] Obviously, this statement should be qualified since a text or an iconography could include false reality. This is especially true in the case of ancient map, hagiography... and a first historical criticism of sources is necessary.

3. Managing uncertainty and source effects over time: an automated tool for stratigraphic units applied to urban entities

The visualization proposed in this project (D, fig. 4) is inherited from the ongoing work of B. Desachy (Desachy, 2014). His research focuses on the creation of a beta version of ‘Le Stratifiant’, which is an automated data processing system for the management of stratigraphic units (Desachy, 2007). The automated tool provides answers to frequently asked questions about urban stratigraphy. It is based on layers recording, layers dates (when known) and the relations between stratigraphic units. It then allows us to:

- produce a graph, representing the relations of the units, in a similar fashion to the Harris matrix (Harris, 1979). In particular, it consists of assessing the uncertainty of the relations and offers various visualization possibilities (different phasing, choice of colors according to the type of layers etc.);
- produce an image of all the stratigraphic units on a timeline, which depends on the dates and relationships of layers. It is an automated process to see relative chronology, inspired by the common Gantt chart.

3.1. Stratigraphic units and urban entities: a possible analogy

In this project, we assume that the automated tool can be applied to urban entities. This assumption is based on the idea that dating and relations between urban entities may be formalized as stratigraphic units. Indeed three possible

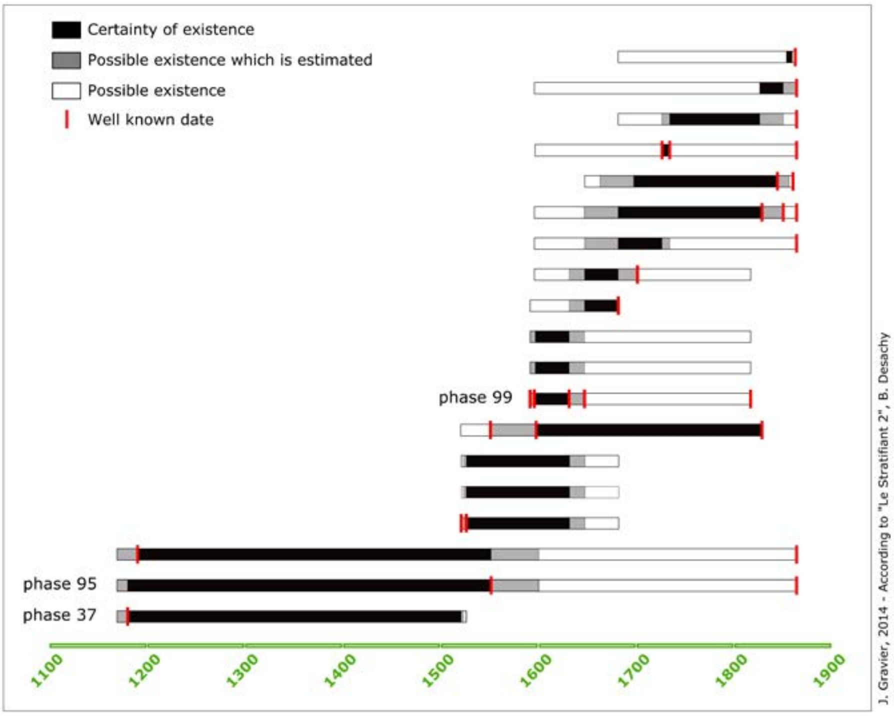


Figure 5: Flow chart of the phases of urban entities placed on a timeline. The defensive system of Noyon.

relationships between entities do exist: an urban entity is prior, posterior or synchronous to another.

However contrary to stratigraphic units, the nature and relations of the urban entities are constructed. As an example, we know that city walls are built in the same time as ditches. In this case we induce that the two urban entities are synchronous. Therefore, layers and relationships between them are observable facts during an excavation, while urban entities are complex objects and their relationships are noted by inference during the analysis of a city.

3.2. An automated tool for visualization of source effects: the example of the defensive urban system of Noyon

What comes next is an example on the subject of the phases of urban entities related to the evolution of the defensive urban system of Noyon. This is done in order to show the usefulness of this tool for visualization of source effects over the long-term.

At the end of the 12th century, city walls were built in accordance with a Municipality and a bishop-count decision. During the middle of the 16th century, city walls were strengthened and fortified constructions raised to protect the town against new artillery techniques. Then at the end of the century, King Henri IV ordered the construction of a citadel, which was rapidly destroyed. At the end, city walls were dismantled during the 19th century. This brief summary of the evolution of the defensive urban system turns out to be registered with nineteen phases of urban entities in the database.

- First, we process the relations of the phases of urban entities using the graph (Harris matrix). The resulting representation is an interesting way of checking data entry. In fact, although there are only nineteen phases in this example, we notice that some relations during the first data processing had been forgotten.
- Second, once the graph is established, we process data to produce a visualization of all the phases of urban entities trough time (fig. 5). Each duration of phase of urban entity is placed on a timeline where markers represent a specific well known date for the phase. For example, city walls (phase 37) are specifically mentioned in a text in 1180. However, the beginning of the urban ditches isn't dated (phase 95). Since these two phases are synchronous, so the start date of ditches is the same as city walls.

We strongly believe that this tool assists in ensuring the view of source effects. As an example, the citadel (phase 99) is described by five well known dates and its relationships with other urban entities provide lots of deduced dates, up to eleven. In this case, the changes of the urban defensive system which were observed at the end of the 16th and the beginning of the 17th centuries are linked to the knowledge of the citadel dates.

In conclusion, this example reveals that a common dating system may be applied to different units of analysis and at separate scales. Moreover, this dating system is a useful tool for the visualization of source effects and it is an essential prerequisite to start an analysis of urban temporalities over the long-term.

4. Functional approach of cities: Beauvais, Noyon and Saint-Quentin

Beauvais, Noyon and Saint-Quentin were important nodes for networks of cities in the north of France from Antiquity

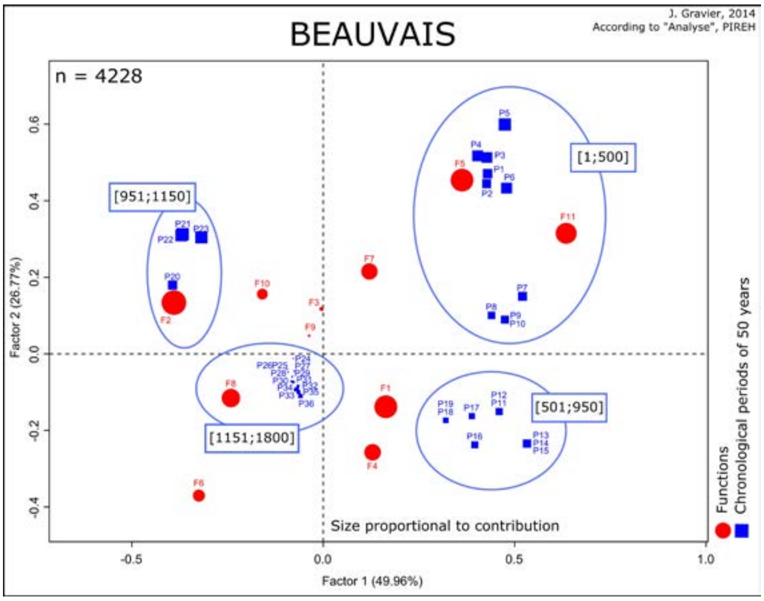


Figure 6: Beauvais: factorial correspondence analysis of the urban entities through function and time.

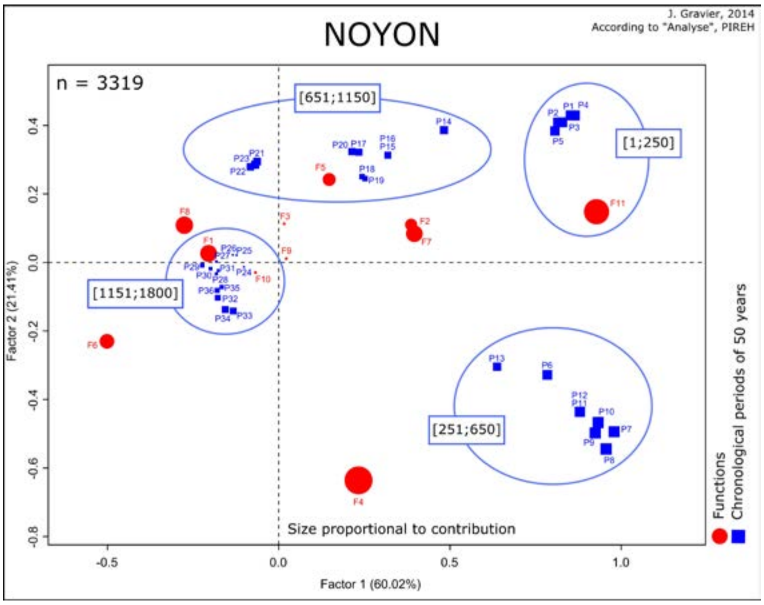


Figure 7: Noyon: factorial correspondence analysis of the urban entities through function and time.

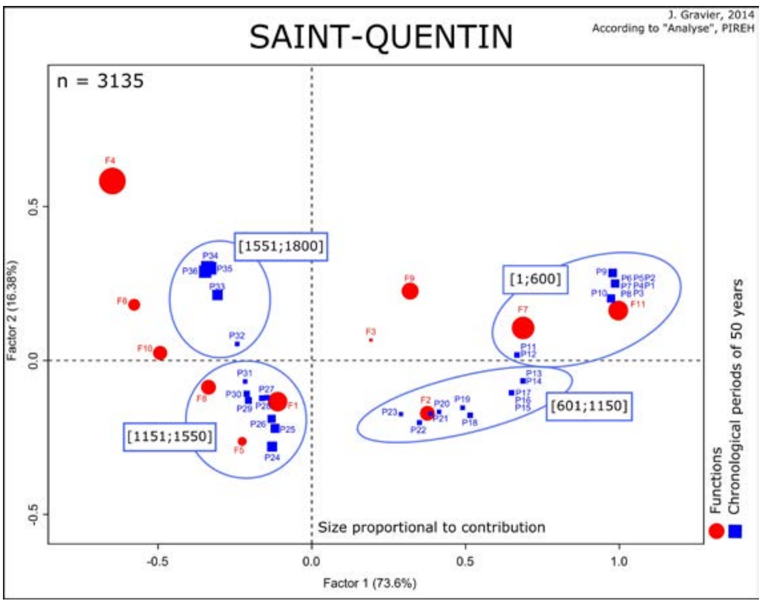


Figure 8: Saint-Quentin: factorial correspondence analysis of the urban entities through function and time.

to the end of the modern age. Indeed, it appears that they were part of two systems of cities.[3] The first one was composed by Beauvais and the cities of the ancient civitas, then, of the diocese. Those towns had special political and religious relations with Beauvais, at least during Antiquity and the Middle Ages. The second network was partly composed by Noyon and St-Quentin. In this case, St-Quentin was the capital of the roman civitas, and then Noyon was the seat of the diocese from the 7th century (Collart, 1984). The two cities were then in direct conflict with each other during the medieval and modern periods, and neither of those towns succeeded to be an economic and a political pole of relative stability.

4.1. The use of statistics in understanding change and sustainability of urban systems

In this section, urban temporalities are examined through multivariate statistics, which provide a systematic way to compare the trajectories of towns. Given that each urban entity is defined by numeric dates, it is possible to study the statistical distribution of urban entities through function (variables) and through chronological period (units), based on 50 years. Factorial correspondence analysis (FCA) for each town has been established[4] and is viewed in this paper based on the first two factors.

Four major periods seem to emerge from the analysis of the Beauvais data set (fig. 6). The first periods are mostly associated with the functions of civilian and religious public spaces, and natural formations (see the functions on tab. 1). The periods between 501 and 950 AD are linked with roads and the defensive and military system. It is probably a visualization of the importance of the city walls during Late Antiquity and the early Middle Ages, in comparison with other functions. The next periods are characterized by the function 2. The latter have a significant contribution on the factor 1 (38.8%). According to the knowledge of the city, this association can be understood as the statistical visualization of the relief and river management, which has been achieved for sheet production. Indeed, Beauvais was one of the main cities for such production during the Middle Ages and it has belonged to the Hanseatic League – a commercial confederation of cities probably created at the end of the 12th and known from the 13th century (Laurent, 1935, Carolus-Barré, 1965). The last periods are associated with the religious and host establishments.

[3] This current study is part of a PhD, focusing on system of cities over long periods of time.

[4] The statistical distribution depends on the estimated dates of urban entities. Beauvais is known from the 1st to the 18th century thanks to 294 urban entities, Noyon due to 253 and St-Quentin according to 281 entities.

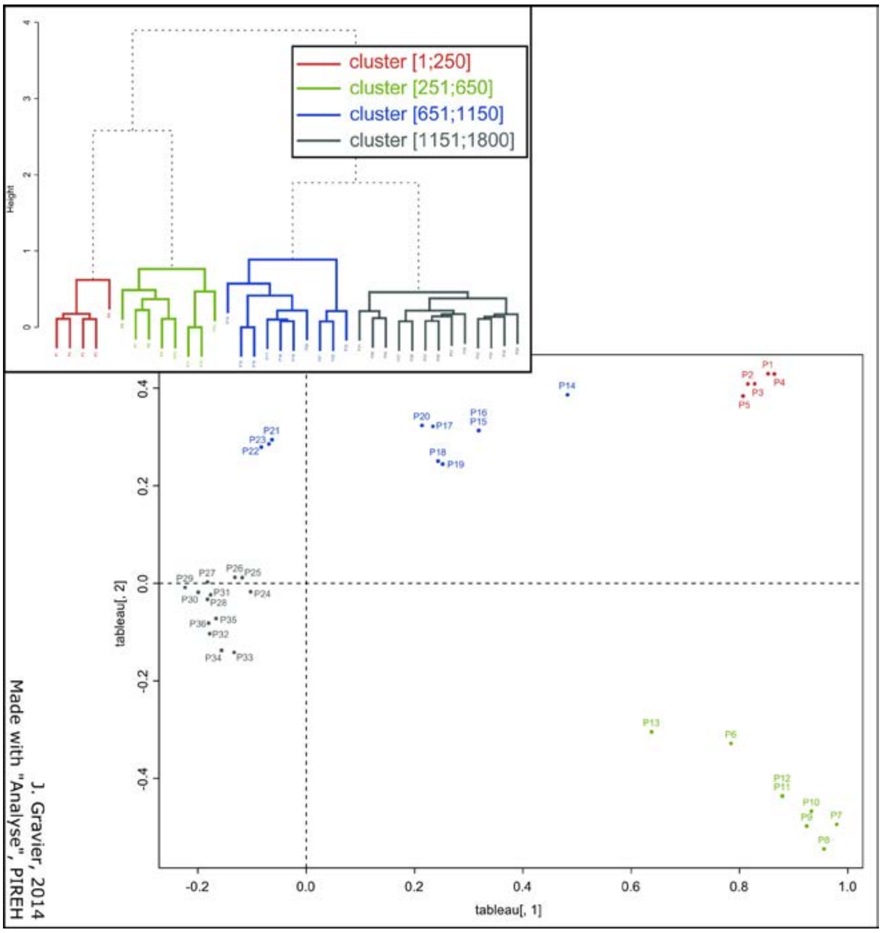


Figure 9: Noyon: hierarchical cluster analysis on correspondence analysis coordinates.

Those chronological periods have a low statistical weight on both factors. Furthermore they are especially linked with each other since these units are very similar.

Four major epochs can be distinguished from the analysis of Noyon, as in the case of the city of Beauvais (fig. 7). However, the changes do not emerge at the same time. The periods between 1 and 250 AD are associated with the natural formations and the habitat, while subsequent periods are strongly linked with the defensive and military system. This function does not have an important statistical weight on factor 1 but really does contribute to factor 2 (78.4%). Then a set of periods is portrayed by the function 5. Nevertheless, two subgroups can be observed: the periods between 651 and 1000 are more associated with functions 2 and 7 on factor 1, whereas the periods between 1001 and 1151 are further determined by function 8. The last periods make up a coherent group linked with the roads and the religious and host establishments.

As well as the previous towns that I just discussed in this paper, the trajectory of St-Quentin is characterized by four major ages (fig. 8). The periods between 1 and 600 AD are linked with the habitat, the natural formations and the funerary function. Unlike the cities of Beauvais and Noyon, the latter contributes to the factors. This is probably a resurgence of source effects because the knowledge of cities during the early Middle Ages in Northern Europe is incomplete due to the complexity of the stratigraphy, pictured by the 'dark earth'.[5] Therefore, those centuries are often known only by the funerary function, as for the city of St-Quentin. Whereas for Beauvais and Noyon some others functions can be discerned, recognizing by the analysis of dark earth (Borderie, 2011). The next periods are associated with the same functions, but also with the second function on factor 2. Moreover, the period from [1151; 1550] differs from the others periods by the importance of the roads, the civilian and religious public spaces, and religious establishments. Finally the periods between 1551 and 1800 are highly associated with the defensive system of the city. The latter does contribute on both factors: approximately 15% on factor 1 and 54% on factor 2. These last associations reveal the building of many fortified structures in the mid-16th century.

In order to complement and to specify the FCA, a hierarchical cluster analysis of the coordinates has been performed. Indeed, it seems necessary to explore chronological periods clusters with this technique, since it involves the data processing of all factors. For example, the process has been applied to the city of Noyon (fig. 9). The associations of chronological periods which have emerged are the same as for the correspondence analysis, probably because the cumulative contribution of the first two factors is 81.4%. However, this process strengthens the assumptions adopted in the previous statistical approach, and thus, provides a basis to the establishment of a comparative approach of the towns.

4.2. A comparative approach of three towns: specificities and recurrences of urban trajectories

The multivariate statistical analyses imply a generalization of information, and consequently, facilitate a comparative approach, specifying the recurrences and the particularities of urban trajectories. As we have been able to note, each town is characterized by functional changes during the middle of the 12th century (fig. 6, 7 & 8). After that date, the periods are strongly associated, excepted for the periods between 1551 and 1800 in the city of St-Quentin. However, the statistical pattern of Beauvais and Noyon during the chronological interval [1151; 1800] is really similar: the periods have low statistical weight and they are linked with the host and religious establishments. This functional profile probably demonstrates that the urban landscape did not vary much from the 12th to the 18th century. Even though the periods between 1151 and 1550 of St-Quentin are analogous to the other two cities, an important change is visible during the mid-16th century. The most recent periods are associated with the defensive and military system because the town had become highly fortified. Indeed, St-Quentin was a border city and a strategic node at that time.

Furthermore, for the three cities, the oldest periods are characterized by different time of changes. This situation probably depends on source effects since the documentation is truly incomplete during that epoch. However, it could also be a picture of the different trajectories of the cities over time, well known in the scientific literature. Beauvais is the capital of an ancient civitas and the seat of a diocese (1st-18th century), while St-Quentin is only known as the capital of a civitas (1st-5th century), and Noyon as the capital city of a diocese (7th-18th century). Therefore, the particularities of each town could affect the intra-urban functional profiles; but the only way to disentangle these two assumptions is to conduct additional analyses on others cities.

The research should also be supplemented by a detailed analysis of networks of cities over long periods of time, where Beauvais, Noyon and Saint-Quentin have existed. This could lead to a second point of view of the towns on a different scale and would allow an interesting prospect of the results.

Moreover, we would like to conduct a study of the three towns from Antiquity to nowadays, which is already the case for the city of Noyon, since this would enable an overview of the changes of urban systems.

[5] Dark earth in archaeology represents layers that frequently separate Roman from Medieval stratigraphy (Macphail, Galinié, Verhaeghe, 2003). Those appeared to be very uniform and have been for a long time considered as gardens or abandoned zones, related to the Barbarian Invasions and a gradual collapse of cities. But current researches consider that the layers actually reflect a variety of the uses of the urban spaces.

References

BEN REDJEB, T. *et al.* (1992). Une agglomération secondaire des Viromanduens : Noyon (Oise). *Revue archéologique de Picardie* 1(1), pp. 37-74.

BIDDLE, M., HUDSON, D. M. and HEIGHWAY, C. M. (1973) *The future of London's past: a survey of the archaeological implications of planning and development in the nation's capital,* Worcester: Rescue.

BORDERIE, Q. (2011). *L'espace urbain entre Antiquité et Moyen Age: analyse géoarchéologique des terres noires.* Thesis, Archaeology. University of Paris 1 - Panthéon-Sorbonne.

BORDERIE, Q., DELAHAYE, J., DESACHY, B., GRAVIER, J. and PINHÈDE, A. (2014). Les synthèses archéologiques urbaines, un projet en cours. In: Lorans, É. & Rodier, X. (eds.). *Archéologie de l'espace urbain.* Paris/Tours: Cths/PUFR, pp. 263-276.

COLLART, J.-L. (1984). Le déplacement du chef lieu des Viromandui au Bas-Empire, de Saint-Quentin à Vermand. *Revue archéologique de Picardie*. 3(1), pp. 245-258.

CAROLUS-BARRE, L. (1965). Les XVII villes, une hanse vouée au grand commerce de la draperie. *Comptes-rendus des séances de l'Académie des Inscriptions et Belles-Lettres* 109(1), pp. 20-30.

DESACHY, B. (2007). *De la formalisation du traitement de données stratigraphiques en archéologie de terrain.* Thesis, Archaeology. University of Paris 1 - Panthéon-Sorbonne.

DESACHY, B. (2014). Manage stratigraphic data with Le Stratifiant. Poster presentation at: CAA2014, Computers Applications and Quantitative Methods in Archaeology, April 22-25, 2014, Paris.

GALINIE, H. (2000). *Ville, espace urbain et archéologie: essai.* Tours: Maison des sciences de la ville de l'urbanisme et des paysages.

GALINIE, H. and RANDOIN, B. (1979). *Les Archives du sol à Tours: survie et avenir de l'archéologie de la ville.* Tours: L.A.U.

GALINIE, H. and RODIER, X. (2002). TOTOPI, 'Topographie de TOurs Pré-Industriel, un outil d'analyse urbaine'. *Les Petits Cahiers d'Anatole*, 11. Available from: http://citeres.univ-tours.fr/doc/lat/pecada/F2_11.pdf [Accessed: 17th December 2010]

HARRIS, E. C. (1979). *Principles of archaeological stratigraphy*, London/New-York: Academic Press.

LAURENT, H. (1935) Un comptoir de vente international au Moyen Age. Nouvelles recherches sur la Hanse des XVII Villes. In: *Le Moyen Age*, 6, Paris: Albert, p. 3-16.

MACPHAIL, R. I., GALINIÉ and H., VERHAEGE, F. (2003). A future for Dark Earth ? *Antiquity : a quarterly review of archaeology* 77(296), pp. 349-358.

OFFENDSTADT, N. (ed.) (2006). *Les mots de l'historien,* Toulouse: PU du Mirail.

OH_FET: A Computer Application for Analysing Urban Dynamics Over Long Time Spans

Laure SALIGNY,[1, 4] Ludovic GRANJON,[1, 4] Thomas HUET,[1, 4]
Gaël SIMON,[2, 4] Xavier RODIER,[2, 4] Bastien LEFEBVRE[3, 4]

[1]Maison des Sciences de l'Homme de Dijon, USR 3516, CNRS, Université de Bourgogne, France
[2]UMR 7324 CITERES Laboratoire Archéologie et Territoires, CNRS, Université François-Rabelais, Tours, France
[3]TRACES, équipe Terrae, UMR 5608, CNRS, Université Toulouse II - Le Mirail, France
[4]Réseau ISA : Information Spatiale et Archéologie – http://isa.univ-tours.fr, France
GDR 3359 MoDyS : Modélisation des dynamiques spatiales – http://modys.univ-tours.fr
laure.saligny@u-bourgogne.fr, ludovic.granjon@u-bourgogne.fr, thomas.huet@u-bourgogne.fr,
gael.simon@etu.univ-tours.fr, xavier.rodier@univ-tours.fr, bastien.lefebvre@univ-tlse2.fr

Abstract

OH_FET formalism has been designed to analyse urban fabric over long time spans (Lefebvre et al., 2008; Rodier and Saligny, 2010; Rodier et al., 2010). The conceptual model is founded on the idea of historical object (OH – objet historique) which is defined as the Cartesian product of three elementary dimensions: social use (what), space (where) and time (when). Processes are performed in Python programming language and data are stored in various kinds of spatial database. With a minimum of required data from archaeologists' observations, the application allows users to automate the calculation of standardised indicators about urban transformations and to generate graphical and cartographic representations. As an example, it is possible to query the database about functional transformations over time and across space either quantitatively (How many social uses for a given date in a given place?) or qualitatively (What changes occurred in social uses over time for a particular place?).

Keywords: Urban Dynamics, Space-Time Modelling, Software Development, Temporal Analyses

Urban fabric (Galinié, 2000) designates the interactions between a space, an environment and social and human activities. The city is the outcome of 'interactive dynamics between the evolving structure of social groups and the development of space' (Galinié *et al.*, 2007). Based on documentation describing chronological and evidential surveys, the study examines the formation and the transformations of urbanised space over long time periods, from the origins of the urban settlement to the present day. The snapshot character of the chronological states described is a stumbling block that prevents the perception and restitution of changes between states and examination of the transformation processes.

After describing the OH_FET model (Rodier *et al.*, 2010), we explain how the model is implemented and the data structured for software development. Finally, we present the analysis module grounded specifically on the time dimension using the the city of Vendôme as an example.

1. The OH_FET conceptual model

1.1 Historical Objects (OH) as the result of functions (F), space (E) and time (T)

The idea of historical object (OH) is the crux of the model. It is defined as the basic unit of analysis, unequivocally distinct from others by the same criteria as the geographical object 'relative to a scale, time-frame and material character of data' brought together in 'the idea of space-time granularity' (Langlois, 2005: 311; Saint-Gérand, 2005). It may be, say, a church, a market or a workshop. The historical object is determined by its interpretation or its social use, its location and spatial extent, its date and its duration. A new historical object is created whenever its function changes or moves in space or whenever its spatial extent evolves.

Accordingly, the model is based on interpreted knowledge and is not a solution for a documentary system of comparison of historical sources.

The second specificity of the model is that it deals with urban space as a system composed of three functional (social use), spatial (location, extent and morphology) and temporal (date and chronology) dimensions meaning that the historical object is the Cartesian product of the Function, Space and Time sets (Fig. 1). This principle is consistent with the logic of the three Ws (What, Where, When) of Peuquet (1994: 447-451) which is often used (Egenhofer *et al.*, 1998; Lardon *et al.*,1999 ; Thériault *et al.*, 1999; Ott *et al.*, 2001).

The combination of the three sets implies that each process is determined by the other two as explained by Donna Peuquet (1994: 448):

> *The triad framework permits the user to pose three basic kinds of questions:*

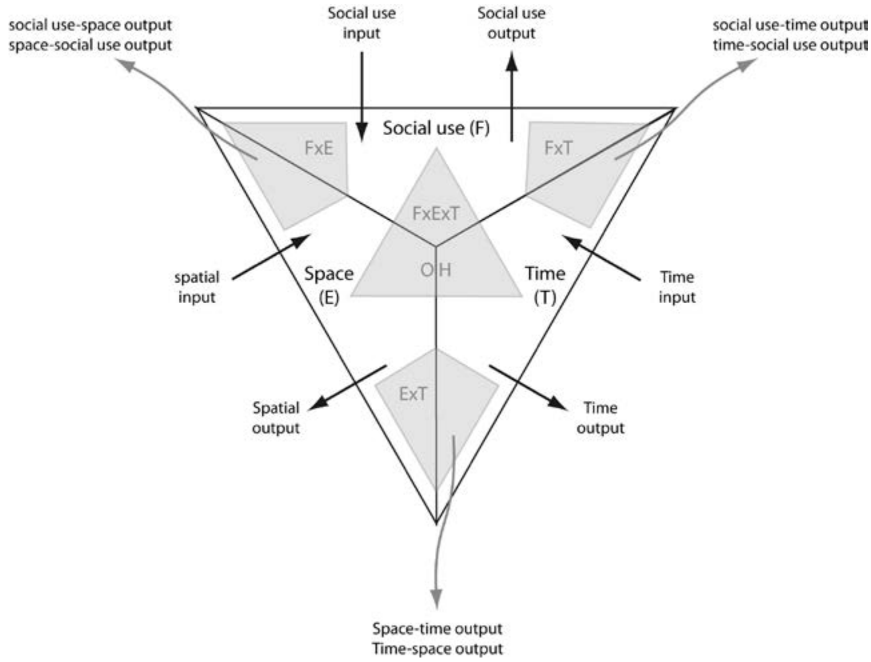


Figure 1: Conceptual model: historical model included in the three dimensions (Rodier and Saligny, 2010).

(1) when + where -> what: Describe the objects or set of objects (what) that are present at a given location or set of locations (where) at a given time or set of times (when).

(2) when +what -> where: Describe the location or set of locations (where) occupied by a given object or set of objects (what) at a given at a given time or set of times (when).

(3) where + what -> when: Describe the times or set of times (when) that a given object or set of objects (what) occupied a given location or set of locations (where).

- The functional interpretation of a historical object is made by selecting a social use from a thesaurus. The dating of the historical object (or rather its inclusion within a time-frame) and its location (or rather its inclusion within a space) directly influence this choice. Some items in the thesaurus are functions determined by a particular space (canonical cloister, burial area, etc.) and/or chronology (*domus*, parish churches).
- The location and shape of a historical object are determined by the social use (necropolises, entertainment edifices) and chronology (necropolises, defensive systems). In addition, the subdivision of space into spatial entities is determined by the successive construction of historical objects depending on their temporal and functional definition (there is no prior matrix subdivision of the space under study).
- The dating of a historical object is characterised by its dates of appearance and disappearance. Even when the temporal continuity of a function is ensured, a change of place (movement) or a significant change of shape imply the transition from one historical object to another.

1.2 Social Use

Functions identifying a social use of space over a certain span of time. To be efficient for their study, functions must be selected from a thesaurus. It is essential to use a common thesaurus to standardise function interpretations so as to make comparison possible. We use the thesaurus drawn up and tested by the *Centre national d'archéologie urbaine* (CNAU) of the French Ministry of Culture which since 1990 has proved its worth for the processing of topographic data of pre-industrial cities (fig. 2).[1]

The social use in the model is defined by the functional entity (EF) corresponding to a term from the thesaurus.

1.3 Space

The spatial modelling is based on the principle of non redundancy of entities. We postulate that space is continuous, bounded by the definition of a study zone and that in the absence of a spatial entity (ES) it contains voids, that is, unoccupied spaces. There can be only one spatial entity in any one place. Accordingly a historical object can be composed of one or more spatial entities and a spatial entity may serve one or more historical objects.

We obtain the spatial entity by subdividing the space arising from the superimposition of archaeological observations that provide input for the model. The shape of the spatial entity has no functional or interpretative meaning; it results solely from spatial changes observed in a place.

Modelling therefore consists in deconstructing the information, even if this means countering the synthetic perception we have of a place (Fig. 3).

[1] http://www2.culture.gouv.fr/culture/cnau/fr/doc_3.html

1. Roads, development
11.thoroughfares, streets
12. unoccupied spaces
13. riverbank works
14. landscapes works
15. crossing points
16. water supply systems
17. sewers/ drains
18. monuments, vestiges
19. unspecified monuments
2. Defence and military structures
21. urban defence systems
22. fortified structures
23. garrisons, barrack buildings
3. Civil constructions
31. public spaces
32. civil authorities, justice
33. education, culture
34. health
35. entertainment, sports
36. baths, thermal baths
37. private homes
4. Religious buildings
41. pagan worship
42. Buildings for Catholic worship
43. convent or monastery buildings
44. ecclesiastical buildings
45. worship other than catholic
46. parish churches
5. Burial places
51. burial area
52. parish cemetery
53. special burial place
6. Trade, crafts, production places
61. trade, markets, shops
62.crafts, workshops
63. agriculture, livestock farms
64. manufacture, industrial premises
65. extraction, quarries
7. Natural formations
71. coast lines
72. rivers (alluvions)
73. marshes
74. colluvial events
8. Other
81. unspecified
82. no confirmed occupation
83. abandoned
9. Non-urban
91. complex settlement of a non-urban character
92. peripheral structure

Figure 2: Urban values (1 to 9) and use values (11 to 92) according to the CNAU thesaurus.

For example, in the illustration above, Observation 1 is composed by ES_0, ES_2 and ES_4.

1.4 Time

Time is modelled by analogy with space so that it is not confined to the role of an attribute. So that it can be mobilised and queried globally, it is necessary to define a temporal entity (ET) that is neutral and defined by the smallest unit of time that is meaningful in dating the phenomenon under study.

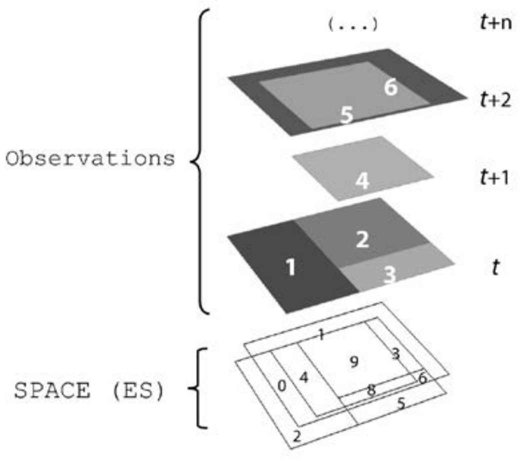


Figure 3: Example of calculation of ES resulting from the spatial subdivision of various Observations.

To this end, the modelling is based on the formalisation of topological relations among time intervals in artificial intelligence by J.F. Allen (Fig. 4: Allen, 1984).

Relations between two entities (A and B) can be written as an equation. For example, the relation A *before* B respects the equation $A^+ < B^-$, where X^- is the start date and X^+ is the end date of the duration.[2] A duration (an interval) is strictly positive ($A^- < A^+$). Equivalence can be found in the modelling of spatial topological relations.[3]

We eliminate all forms of intersection between two intervals so that our temporal entities have non-redundant relations only. All notions of duration (centuries, years, periods) can then be recomposed from the model. Like spatial entities, temporal entities are disconnected from the functional and spatial interpretation. The duration and number of temporal entities for a time period determine a frequency.

Time, like space, is continuous, bounded by the chronological boundaries of the study object. The temporal resolution selected for time entities determines the dating of historical objects. The creation of time entities depends on the prior definition of observations. There is no a priori subdivision of time. At any given time, there can be one and only one time entity but it may serve as many historical objects as necessary. It is the subdivision of time by the accumulation of historical objects and with time that defines time entities.

There are two types of time entity: dates and durations. For example, in Fig. 5, dates are: 150 (ET_1), 200 (ET_3), 250 (ET_5), 350 (ET_7) and durations: [150–200] (ET_2), [200–250] (ET_4), [250–350] (ET_6).

After being applied to the study of the cities of Tours and Vendôme (Lefebvre, 2012; Rodier *et al.*, 2010; Simon, 2012) this model has proved effective in analysing urban change. But the exploratory works also showed that the constraints of data integration, structuring and manipulation imposed by the model were too stringent with the software available for it to be diffused and used in the archaeological community. This is why we undertook the design and development of ad hoc software.

[2] By transitivity: $A^- < B^-$; $A^+ < B^+$ and $A^- < B^+$.
[3] Where the temporal *before* or *after* is equivalent to the spatial *disjoint*; *meets* or *met-by* is equivalent to *touches*; *overlaps* or *overlapped-by* is equivalent to *intersects*; *starts* or *ends* or *during* or *contains* or *finishes* or *finished-by* is equivalent to *contains*; *equals* is equivalent to *equals*.

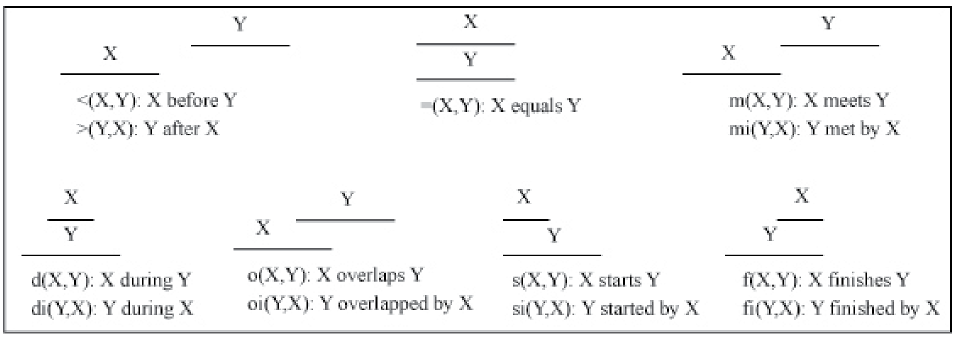


Figure 4: Allen's time relationships.

2. OH_FET computer application

2.1 Principles and technical criteria for software development

Two fundamental principles have been applied. First, the development concerns only the processing and analysis implementations that are new or that could not be performed automatically with the tools available. This is why all aspects of the display, management and interrogation of semantic and spatial data is done on existing software. Second, the OH_FET application must be composed of independent modules that can be used separately.

Four modules have been developed (Fig. 6). The first is for integration (cf. section 2.2), the second for deconstructing data into the three dimensions (cf. section 2.3), the third is for storage in DBMS (cf. section 2.4) and the fourth handles analyses and graph production over the time dimension (cf. section 3).

The principle is that the user may integrate a set of vector data in the *shapefile* format together with a file describing the projection system (*SRID*) and a tabular file (*csv*) containing the values and levels of the hierarchical thesaurus. From these input files, the modules integrate rules to check the compatibility of the data sets and then rules for constructing the functional (EF), spatial (ES) and temporal (ET) entity classes and their relations. These are stored in a relational and spatial data base system. Lastly, the display module proposes to restore them in the form of graphs and diagrams and as output files in *shapefile* and *comma-separated values* from the processing of the data.

We opted to develop software independently of any platform because the archaeological community uses several GIS software packages and we chose to develop it in Python, which is a simple, expressive, multi-paradigm and multi-platform programming language widely used by the scientific community. Python offers a large range of libraries for passing from one application to another. For example, the Python module pyshp can read *shapefiles* as if they were simple lists (both for column names and series of values); the Python psycopg module can execute SQL queries on a spatial database (such as *PostgreSQL/PostGIS* or *SpatiaLite*) from a Python script. Python is also used for programming into leading GIS applications: *QGIS* and more recently *ArcGIS*.

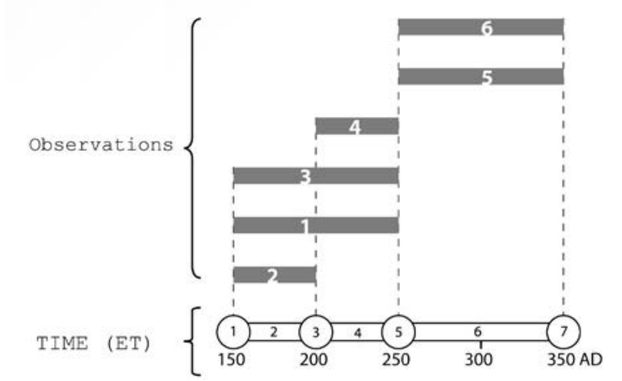


Figure 5: Example of calculation of ET resulting from the temporal subdivision of various Observations.

For the database, we chose *PostgreSQL/PostGIS* and *SpatiaLite.* Both were selected because they are fully compatible with current GIS, they comply with international standards such as the WKT format[4] and OGC recommendations,[5] they have features for exporting/importing shapefiles (the common format of geographical data) and they can be interrogated with procedural language such as PL/SQL or PL/Python.

2.2 The graphical user interface and integration module

The software is installed from an executable file that launches a graphic interface designed to be as simple as can be. It prompts the user to open a shapefile containing all the data of the urban space under study. These data are termed 'observations'. It is mandatory for these observations to have a function among the terms of the hierarchical thesaurus, plus a start and end date. Accordingly, the shapefile must have at least four fields that cannot be left blank: ID, Function, Start date, End date (fig. 7). Fields names are

[4] *Well-Known Text* (WKT) is a text markup language (human readable) representation of geometry, spatial reference systems, spatial transformation, etc. It is also used by various Python geometric modules such as Shapely.

[5] *Open Geospatial Consortium* (OGC) is an international consortium that makes recommendations (formats WKT, WKB, WMS, etc.) to ensure interoperability of contents and services in geomatic and spatial information fields.

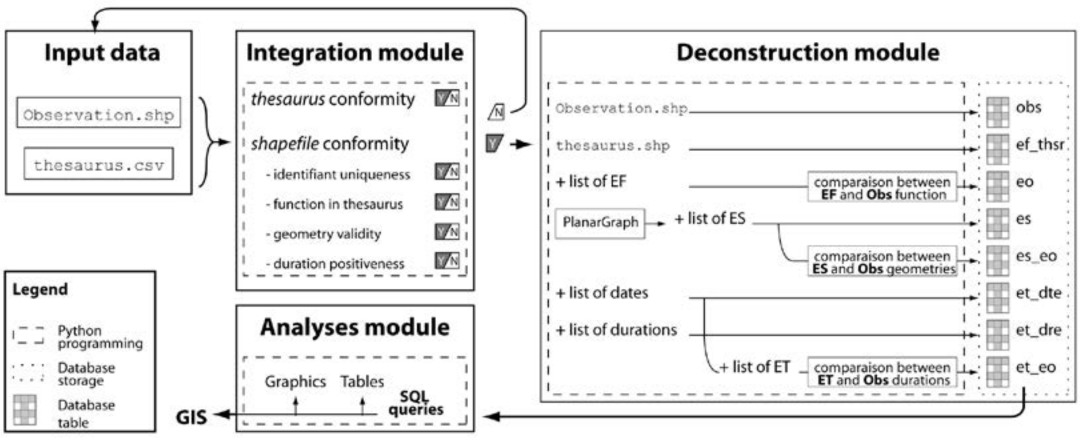


Figure 6: OH_FET application diagram.

1. Selection of the shapefile and its SRID
2. Matching of shapefile's fields with OH-FET required fields (identifiant, function, start date, end date)
3. Choice of the thesaurus (.csv)
4. Selection of the results folder, where the SpatiaLite database and log-files will be saved
5. Storage interface: check boxes to export in a PostGRE/GIS database (connection parameters)
6. Analyses interface: check boxes to calculate different graphical/spatial outputs (chronographs)

Figure 8: OH_FET graphical user interface.

not important, although an explicit label is recommended. Besides these four mandatory fields, the shapefile may have any number of fields for ‘attribute data’. The ID field must be a unique integer, the Function field a function value that must be referenced in the thesaurus, and the duration must be strictly positive (Start date < End date).

The user also uploads a thesaurus and may specify the projection system. The user also indicates the output folder where the results are to be stored. The results are created by default in a *SpatiaLite* data base and possibly also in *PostgreSQL/PostGIS* DBMS. The default base, *SpatiaLite*, is included in the executable file whereas *PostgreSQL/PostGIS* must be pre-installed locally or on a server with PostGis activated (Fig. 8).

The graphical user interface also proposes the types of graphical output the user may choose. These outputs are those executed in the display module. Currently, the executable file proposes just a single output, the chronograph for two types of object classes (functional entities and historical objects, cf. section 3).

The integration module is the interface between the user and the deconstruction module. Its purpose is to control the coherence of values and types of data when the user clicks the ‘OK’ button. For example, a procedure of the integration module will verify that all OH durations are strictly positive (Start date < End date, where Start date and End date are respectively the lower limit and the upper limit of a duration) (cf. Fig. 6).

2.3 Deconstruction module

The deconstruction module groups all processes starting from the data source and results in the recording in the

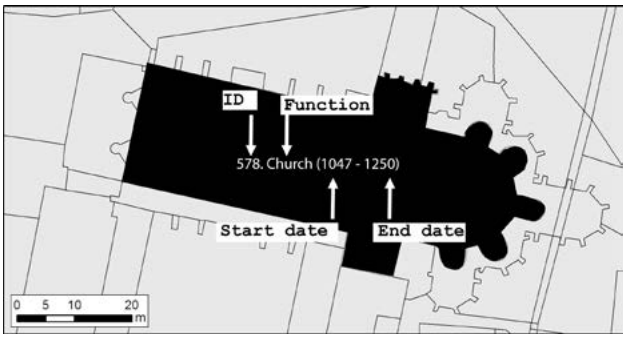


Figure 7: Example of an Observation with the mandatory fields.

Figure 9: Logical constraint for Observation recording.

spatial database. The result of on-the-fly deconstruction is the creation of four classes in the OH_FET data base: the Function class of functional entities (EF), the geometry of the spatial class with spatial entities (ES) and the time dimensions classes (ET), date and duration.

Deconstruction is performed with just one logical constraint (Fig. 9) which can be formulated: ‘if A.duration crosses B.duration then A.location does not cross B.location’. The reciprocal is true: ‘if A.duration does not cross B.duration then A.location may cross B.location’.

a. Social use, EF

The OH_FET model has been designed on the basis of the CNAU thesaurus but the application can import and use any thesaurus.

The identifier of a functional entity is the concatenation of its path from the highest level of generalisation to its own definition. For example, the key of the EF ‘unoccupied spaces’ is ‘12’ where ‘1’ refers to ‘Roads, development’ (highest level of generalisation) and ‘2’ is its own key. So, the identifier of the ‘unoccupied spaces’ function is eF_12.

b. Space, ES

As said, after the deconstruction operation, each Observation is divided into its most elementary part for either the geographical (Fig. 10) or the temporal dimension.

We can observe from the illustration below that Observation 578 is composed of two spatial entities: ES_3226 and ES_3365. Relations between Observation and spatial entity are recorded in the database.

This first step is performed with the PlanarGraph Python module.[6] The PlanarGraph module can be used to obtain ES, the vertices and arcs of ES, but without recovering the initial attributes of observations. Attributes are recovered

[6] Developed by Pascal Leroux https://github.com/paleroux/PlanarGraph

FIGURE 10: SPACE DECONSTRUCTION (ES)

by spatial queries and stored in the data base to link a spatial entity and a historical object (cf. Fig. 6).

The initial data tested often exhibit problems with geometrical capture, seldom obeying strict topological rules. The polygons which should have common boundaries are not always properly captured. This results in overlapping of polygons.

If these data sets are processed directly, the deconstruction engenders far too many spatial entities and so skews the analyses. Several approaches have been contemplated to overcome this.

First, so as to work with existing data, two data cleaning protocols have been proposed (one under *Arcgis* based on topological coverage and the other on *QGIS* by using the topological properties of *GRASS GIS*). These two protocols are based on a system of tolerances, the parameters of which are to be defined. Unfortunately, this raises other problems. Specifying too low a threshold entails redundant spatial entities; specifying too high a threshold eliminates spatial entities that go to constitute historical objects.

We plan eventually to propose a good practice guide so as to avoid these errors at the time of data capture. After the subdivision, no geometric entities should overlap.

c. Time, ET

Deconstruction of time consists in coding the creations of temporal entities in a Python script so that the 4 out of 13 Allen's relations used here are validated (Fig. 11).

The duplicates are eliminated and the list is ordered from the most ancient to the most recent date. The process then consists in parsing this list and identifying whether a date lies between two durations or just touches one duration. In the latter case, the date is a *terminus*. Respecting predicates of the OH_FET model, the list of temporal entities is constructed (Fig. 11).

The list of ETs starts/ends necessarily with this sequence:	[■],[■$^-$,■$^+$],[■]
A duration has necessarily two dates touching it	[■],[■$^-$,■$^+$],[■]
A date has, at least, one duration touching it	[■$^-$,■$^+$],[■] or [■],[■$^-$,■$^+$]
There cannot be two consecutive durations	~~[■$^-$,■$^+$],[■$^-$,■$^+$]~~
There cannot be three consecutive dates	~~[■],[■],[■]~~

FIGURE 11: PREDICATES FOR CONSTRUCTING THE ET LIST

The following sequence of tests ('a' to 'g') can be used to calculate a variable for each ET/Observation. Initially the variable value is 0 (Fig. 12).

All tests, 'a' to 'g', are performed for each Observation and ET. After 'g', if the index >= 1 then the ET is recorded in the list of ETs. In other words, this ET is one component of the Observation duration.

All ETs composing an OH must necessarily follow each other without interruption (e.g.: …, ET_6, ET_7, Et_8,…)

Remember that temporal entities result from a topological deconstruction in the time dimension and there is no overlap between two durations. In other words, for a given duration there is no creation and no destruction of Observations.

2.4 Storage (spatial DB)

For storage, two kinds of spatial database have been chosen to export the result of the deconstruction process: *SpatiaLite* and *PostgreSQL* (with its spatial extension *PostGIS*). These two databases (DB) serve two different purposes. *SpatiaLite* is a simple, easily transportable DB file. *PostgreSQL/PostGIS* has been designed for DB administrators. It enables more complex management of data and supports multi-user access. Furthermore, *PostgreSQL/PostGIS* has a topological extension with topology whereas topology is not yet completely implemented in *SpatiaLite*.

	Relations between Obs durations and ET	Logical test	variable
a	eT *contains* Obs	$Obs^- >= eT^-$ and $Obs^+ <= eT^+$	variable =variable +1
b	Obs *contains* eT	$Obs^- <= eT^-$ and $Obs^+ >= eT^+$	variable =variable +1
c	Obs (older) *overlaps* eT	$Obs^- <= eT^-$ and $Obs^+ <= eT^+$ and $Obs^+ >= eT^-$	variable =variable +1
d	eT (older) *overlaps* Obs	$Obs^- >= eT^-$ and $Obs^+ >= eT^+$ and $Obs^- <= eT^+$	variable =variable +1
e	Obs *meets* eT	$Obs^+ = eT^-$ and $Obs^- < eT^+$	variable =0
f	eT *meets* Obs	$Obs^- = eT^+$ and $Obs^+ > eT^-$	variable =0
g	eT *equals* Obs	$Obs^- = eT^-$ and $Obs^+ = eT^+$	variable =variable +1

FIGURE 12: SEQUENCE OF TESTS TO MATCH OBSERVATIONS AND ETS.

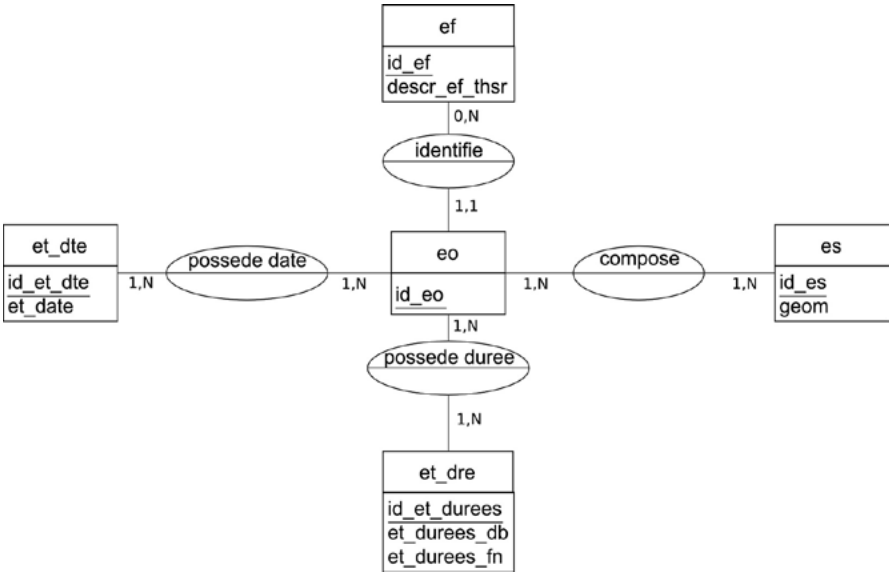


Figure 13: Data base conceptual model.

Data are integrated within data bases by the conceptual model below (MERISE formalism) (Fig. 13). A table stores the durations to improve performances, but they could be calcualted on-the-fly.

Physically, tables are created and then populated with the data from the deconstruction module.

The historical objects are the crossover of the ES, EF, ET_dte tables grouped by the identifier of the initial observations. To reconstruct the historical objects, we recover the function of the historical object, the 'lowest' date of the historical object as the date of appearance and the 'highest' date of the historical object as the date of disappearance. The geometry of the historical object may take on two forms, either by merging spatial entities into a simple geometry or by aggregating spatial entities into a multipolygon.

3. Analysis modules

The principle of the OH_FET model is to provide output from which to observe on the one hand the distributions of function, space or time and on the other hand functiono-spatial or spatio-functional (F x E), functiono-temporal or temporo-functional (F x T), spatio-temporal or temporo-spatial (E x T) variabilities (See Fig. 1, section 1.1). The crossover of the three dimensions at the core of the model corresponds to the state of the historical objects at a given time. The aim of this approach (Rodier and Saligny, 2010; Lefebvre *et al.*, 2012) is to:

- restore all possible states, that is, all maps at all possible dates or intervals;
- analyse and observe all possible state changes by the difference between two states;
- understand transformations, that is, the state-change process by questioning the dimensions two-by-two in order to observe the factors affecting the change and to estimate the role or preponderance of one relative to the other.

The OH_FET model is designed to produce new elements of analysis for observing:

- the functional distributions (F): the number of times each functional entity is used to form historical objects;
- the spatial distribution or solicitation of space (E): the number of times each spatial entity is used to form historical objects;
- the temporal distribution or solicitation of time (T); the number of times each temporal entity is employed to form historical objects;
- the functional variability in space (FxE) based on the frequency of demand for functions by spatial entity and the number of different functions per spatial entity;
- the functional variability in time (FxT) based on demand for functions by temporal entity and the number of different functions per temporal entity;
- the spatial variability in time (ExT) based on the demand for temporal entity by spatial entity and the temporal variability in space (ExT) based on the demand for spatial entities per temporal entity.

Each result represents an aspect involved in understanding the whole by supplying various complementary information for comprehending the dynamics of the system.

3.1 Development

In order that the various modules can eventually be included in a single software package, the analysis module is also developed in Python language.

The analysis module connects to the data base created by the integration module from a Python, SQLAlchemy module.[7] This module can be used to connect to various types of DBMS and to query them with the same code. SQLAlchemy is an Object Relational Mapper and will allow us to process all of the tables of the relational base

[7] http://www.sqlalchemy.org/

as classes of computer objects. A Python (OH_FET) class is created with which to connect to the data base. As many methods are created as queries are needed for the analyses. The result is formatted to the corresponding output. We obtain either a graph (through the *matplotlib* module[8]) that can be safeguarded, or a csv file containing the data recovered which can then be used in GIS software to produce further analyses (cf. Fig. 6).

Only analyses bearing on the time dimension can be restored in graph form. The OH_FET application is currently oriented mainly to processing the time dimension so as to automate analyses and restorations that are difficult to implement in other software packages.

The DBMS structure can be used to recover the crude data required for the analyses implemented by Bastien Lefebvre in his thesis (Lefebvre, 2008). The analyses are currently in the form of Python scripts and shall soon be integrated in the OH_FET graphic interface so that users can produce them automatically, as is already the case for chronographs (cf. below).

3.2 Specimen analyses

The analyses proposed below are from data in a PhD thesis currently being written on the town of Vendôme (France) which was a castle-city that developed in the Middle Ages (1000 years occupation). Only one district of the town is studied, the abbey district (Simon, 2012).

a - The demand for time, the distribution of historical objects over time

It is a question on the one hand of the distribution of historical objects over time and on the other hand of the number of times each temporal entity is solicited by the formation of historical objects. This indicator can be used to evaluate source effects in part and to observe urban transformations (Lefebvre *et al.*, 2012).

For the abbey district of Vendôme, the graph (Fig. 14) displays five documentary lines corresponding to the founding of the abbey in the mid 11th century (SD1), the increase in the number of medieval charters in the 13th century (SD2), the multiplication of texts and the improved conservation of the built environment between the end of the 15th and the 16th centuries (SD3), the Maurist plans and engravings at the very end of the 17th century (SD4) and lastly the old land registry in 1811 (SD5).

The OH_FET application proposes to supplement this indicator with a graph compiling the distribution of historical objects on a time line. This chronograph shows the number and duration of historical objects over time (Fig. 16).

A query made of the class of temporal entities can identify the proportion of historical objects appearing and disappearing on the time line, highlighting the period of urban transformation and renewal but also the source effect (Fig. 16).

[8] http://matplotlib.org/

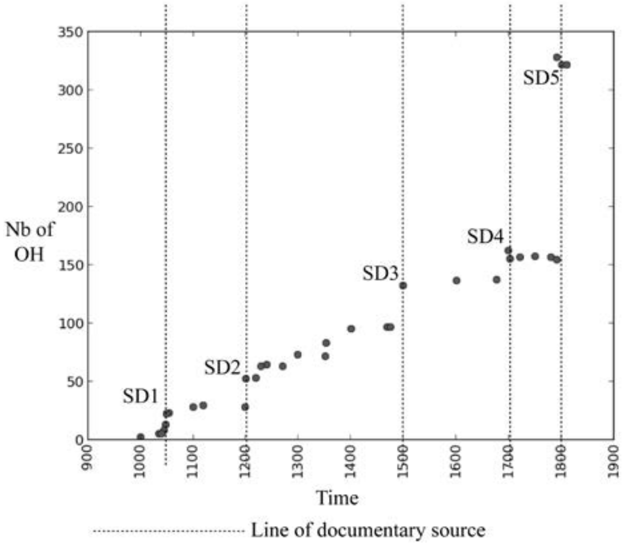


FIGURE 14: DISTRIBUTION OF OH OVER TIME

Again in the case of Vendôme, the appearance curve shows the five source effects observed before (Fig. 15) while the disappearance curve shows a strong propensity for historical objects to be conserved since on average never more than ten or so disappear at the same time and appearances greatly outnumber disappearances. The two peaks of disappearance around 1050 and 1700 nuance the effects of sources SD1 and SD 5 and reflect real transformations in the urban space.

b - The distribution of functions over time

We have implemented a series of representations based on combined queries of functional entities and temporal entities.

This time plot of function represents the times at which functions were present (fig. 17). It shows that some were permanent, others lasted for part of the total duration and other functions were discontinuous, appearing, disappearing then reappearing.

The plot below (Fig. 18) shows the percentage of functions present over time.

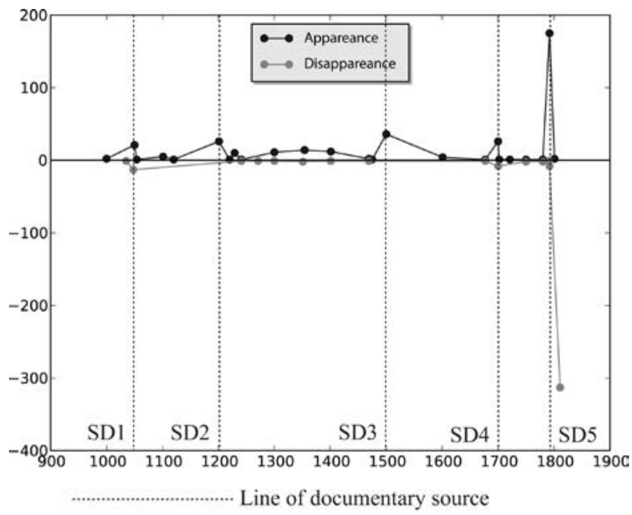


FIGURE 16: NUMBER OF HISTORICAL OBJECTS PER TEMPORAL ENTITY ACCORDING TO THEIR APPEARANCE AND DISAPPEARANCE

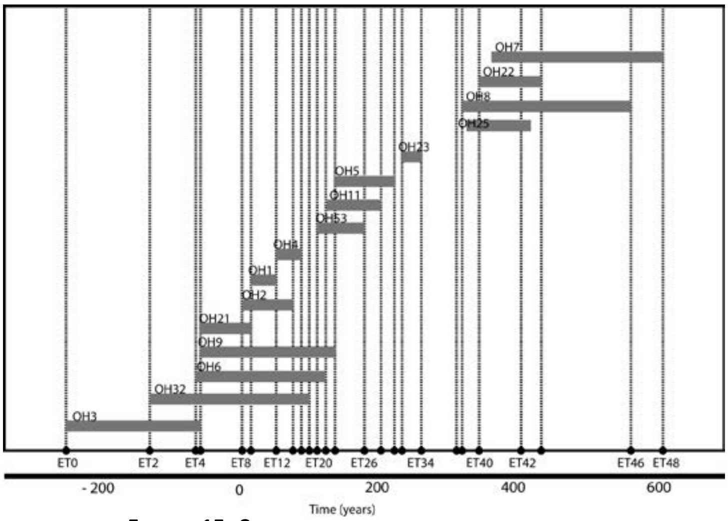


Figure 15: Chronograph of historical objects.

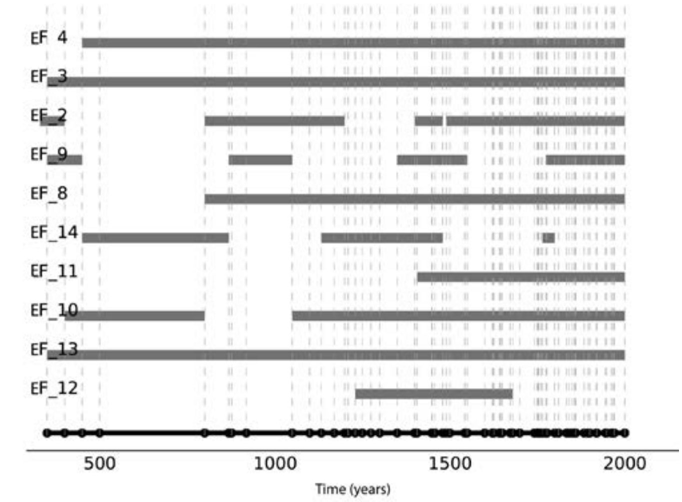


Figure 17: Chronograph of social use.

In this example for the town of Vendôme, the plot shows the variations in the functional fractionation of historical objects over time. This can only be interpreted by bearing in mind the substantial source effect. Two periods must be isolated in the documentation, the hinge point of which lies between the mid 14th century and the mid 15th century. The first period is characterised by sporadic documents which, by chance effect of the texts conserved of archaeological operations conducted, provide information about one or more historical objects but never over a long stretch. The second period is that of serial sources (land tax records and plans, land registry) which by nature provide information in one go about much or all of the space. These two stages of documentation are particularly apparent in this figure in which the general shape of the curves for dwellings, gardens and courtyards that must have made up most of the urban space surged during the 15th century because of the conservation of several land tax registers describing the town plot by plot. Although the appearances result mostly from the source effect, disappearances can be more readily interpreted. The number of defensive features declined markedly in the late 18th century as did the number of religious buildings, graveyards and water courses. These disappearances or marked reductions are the first signs of the transition from the pre-industrial to the industrial town under the influence, among other things, of concern for hygiene that sought to push the dead out of cities, to backfill open sewers that were sources of disease and to open up the city by tearing down the surrounding walls.

c - The demand for space over time

Figure 19 is a bar chart showing the number of spatial entities per temporal entity. It would be an indicator of spatial subdivision and its development over the course of time if it ware paired with the same plot showing not the number of spatial entities but their surface areas. This plot alone shows the steady increase in the number of spatial entities over time, except for two dips in the mid 11th and late 13th centuries, indicating a slight decline in the extent for which information is available for the space in question.

d- Temporal synthesis

The temporal synthesis proposed takes up the results of several plots and can be used to describe the architecture of time that B. Lefebvre referred to as a temporal 'map' (Lefebvre, 2008).

The rhythm of appearances and disappearances is represented by a bar indicating for each event-type temporal

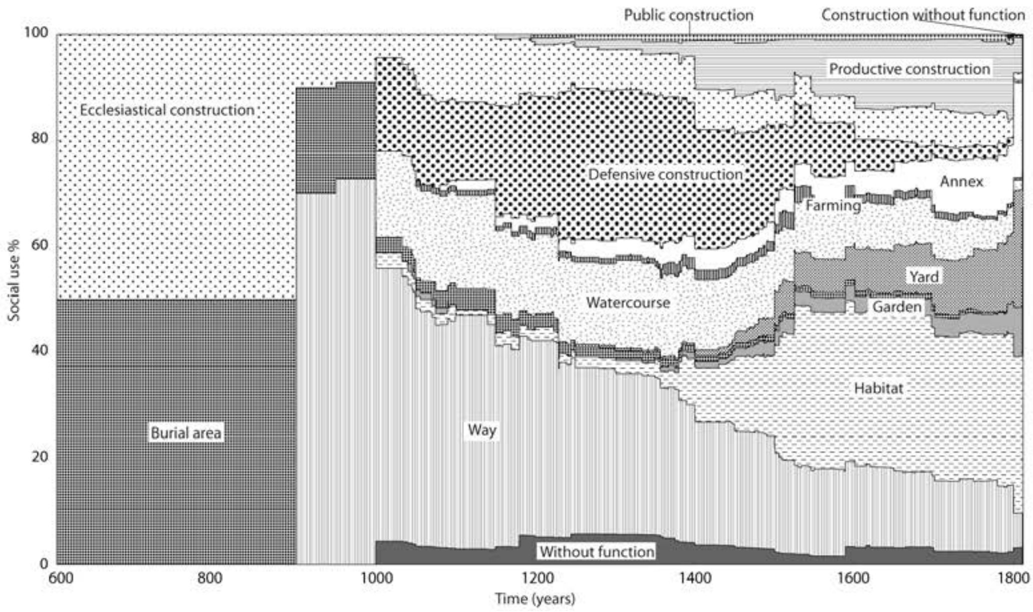


Figure 18: Distribution of social use over time

entity the proportion of historical objects that appear, disappear or are maintained. Between two event-type temporal entities, the duration-type temporal entities are characterised by two items of information: the functional variation (the light values show a reduction in the ratio, the ratio is low when there are more historical objects than functional entities) and the number of historical objects during this phase (Fig. 20).

This plot for the Trinity Abbey district of Vendôme alone reveals several phenomena that can explain the urban fabric.

1. As few historical objects disappear over the whole of the study period, the focus is on appearances. The second half of the 11th century, the period when the abbey was founded, saw many creations at a sustained pace. The 13th century, although the apogee for the Trinity, appears to have been a period when constructions stagnated. Creations of historical objects accelerated from the 14th to the 16th centuries, although they were times of crisis. The 17th and 18th centuries, which are relatively well documented times, appear to have been a stable phase with few objects being created or transformed.
2. Although the state of knowledge greatly affects this representation, the functional diversification of space is very real over the long term.
3. This functional diversification seems to have ended in the 15th century, with later urban space retaining the same structure; even if buildings were regularly updated, their functions changed little. The Maurists who greatly transformed the abbey architecturally from the 17th century onwards did not change its function as indicated by the functional variability index which remains low in this period.

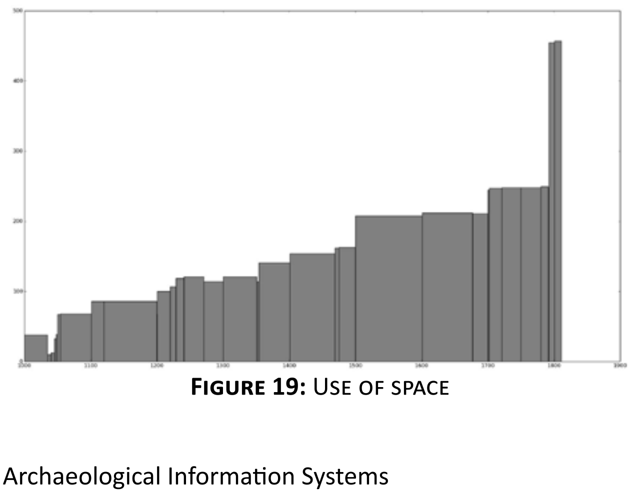


Figure 19: Use of space

Conclusion

The development of the OH_FET application supplies the extension to the model that was missing. It renders accessible analytical tools for studying cities based on a set of input data complying with a loosely restrictive structure. It can be used for any urban space on the basis of a set of historical objects described by an item in a thesaurus, a geometry (area or point) and a chronology in the form of an appearance date and a disappearance date. On this basis, the application performs the spatial and temporal deconstructions recommended by the OH_FET model and produces the outputs, focused mainly on the time dimension in the form of documents to support interpretation of the data set.

Prospects for exploiting the OH_FET application are of several orders. First, the application can be used to multiply the number of study cases essential to test the model's robustness. In addition, comparison of examples will make it possible to envision the initial objective of the modelling of comparing urban trajectories. The aim is not to compare cities term for term but to identify types of transformation of the urban space over the long term and so trajectories.

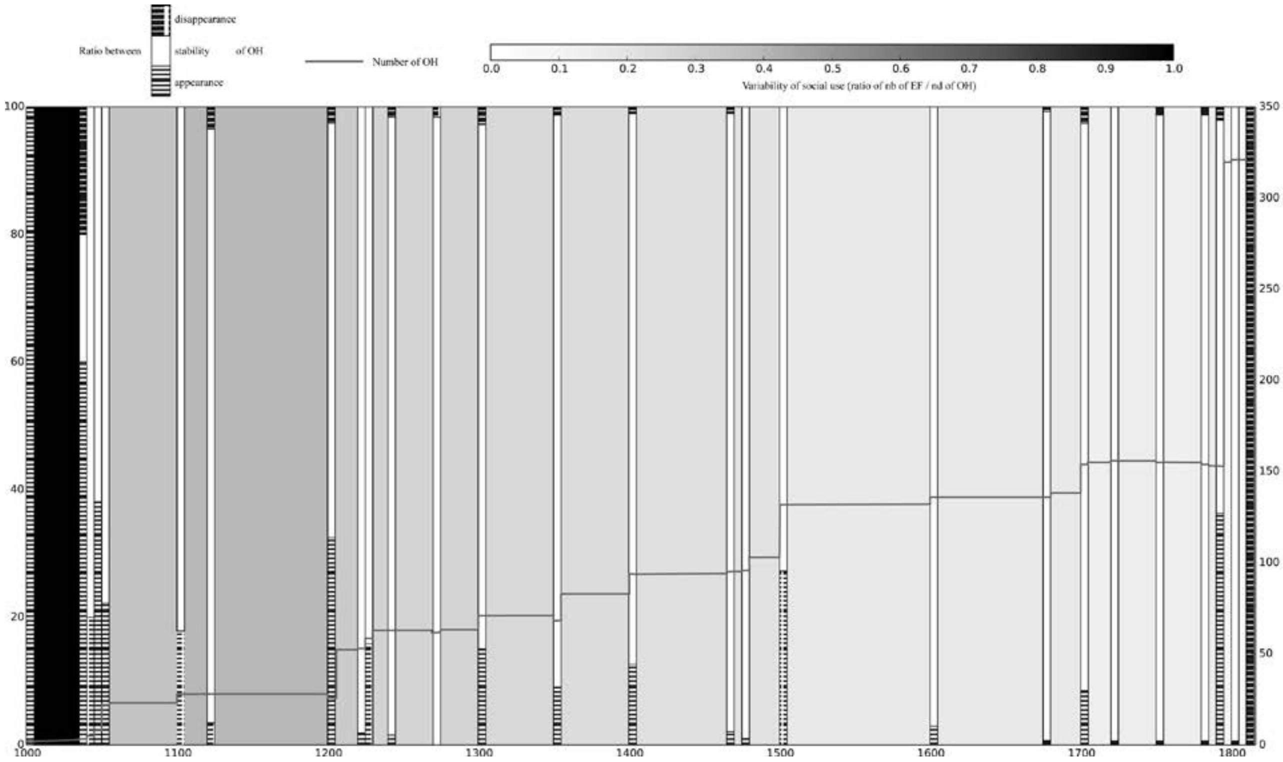


Figure 20: Temporal map.

Next, it is envisaged to use the application on different scales, that of the city, a district or an excavation. Exploratory tests (Lefebvre, Rodier and Saligny, 2012) have shown the model is relevant to all three scales. However, only the scale of the district or a portion of the urban space offering areal and continuous spatial information has been the subject of more advanced studies (Lefebvre, 2008, 2009 ; Simon, 2012). Other sites are currently being studied on this scale.

Lastly, the structure of the data base can be used to propose outputs from the model such as the lines of force of the plan as Bastien Lefebvre contemplated in his thesis, analysis of networks or graphs, and even dynamic representations, the heuristic representations of which are to be found in the various explorations of a data set.

Beyond these possibilities for the application, other avenues for further development lie open especially for managing hypotheses and taking account of uncertainty.

Acknowledgements

The software has been developed by Maison des Sciences de l'Homme, Dijon, with support from the European Regional Development Fund and Burgundy Regional Council.

We are grateful to Bruno Dessachy, Thomas Devogèle, Hélène Mathian and Julien Perret for their collaboration and their proposals making transcription of the model possible.

We thank Pascal Leroux for the *PlanarGraph* script.

References

ALLEN, J. F. (1984). 'Towards a General Theory of Action and Time' *Artificial Intelligence* 23(2), pp. 123-154,

EGENHOFER, J. M., GOLLEDGE, G R. (1998). *Spatial and Temporal Reasoning in Geographic Information*, Oxford : University Press.

GALINIÉ, H. (2000). Ville, espace urbain et archéologie, Collection *Sciences de la Ville* n°16, Maison des Sciences de la Ville, de l'Urbanisme et des Paysages, Tours : Université François Rabelais.

GALINIÉ, H. (ed.) (2007). *Tours antique et médiéval : lieux de vie, temps de la ville : 40 ans d'archéologie urbaine,* Tours : FERACF, .

LANGLOIS, P. (2005). 'Complexité et systèmes spatiaux' In: Guermond Y. (ed.). *Modélisations Géographiques*, Paris: Traité IGAT, Hermès, pp. 299-320.

LARDON, S., LIBOUREL, T., CHEYLAN, J.-P. (1999). 'Concevoir la dynamique des entités spatio-temporelles', In: Représentation de l'espace et du temps dans les SIG, *Revue internationale de géomatique* 9, pp. 45-65.

LEFEBVRE, B. (2008). La formation d'un tissu urbain dans la Cité de Tours : du site de l'amphithéâtre antique au quartier canonial (5e-18e s.). History PhD, Archaeology, under the supervision of Henri Galinié and Élisabeth Lorans - Université François-Rabelais, Tours, 4 vol., 904 p. Available from: http://tel.archives-ouvertes.fr/tel-00349580/en/

LEFEBVRE, B. (2009). 'How to show Dynamics of Urban Fabric : Cartography and Chronography', In: *Online proceedings of the 37th annual international conference on Computer Applications and Quantitative Methods in Archaeology (CAA) 2009*, Willamsburg, Virginia, USA, March 22-26, 2009. Available from: http://www.caa2009.org/articles/Lefebvre_Contribution224_a.pdf

LEFEBVRE, B, RODIER, X., SALIGNY, L. (2012). 'La modélisation de l'information spatio-temporelle', In: Archambault de Beaune S. et Francfort H.-P. (eds.). *L'archéologie à découvert : hommes, objets et temporalités*, Paris: CNRS Editions Alpha. pp. 249-258.

LEFEBVRE, B, RODIER, X., SALIGNY, L., (2008). Understanding Urban Fabric with the OH_FET model based on Social Use, Space and Time , *Archeologia e Calcolatori,* 19, Rome: CNR, Available from: http://soi.cnr.it/archcalc/indice/PDF19/16_Lefebvre.pdf pp. 195-214

OTT, T., SWIACZNY, F. (2001). *Time-Integrative Geographic Information Systems, Management and Analysis of Spatio-Temporal Data*, Berlin: Springer.

PEUQUET, D. J. (1994). 'It's about time: A conceptual framework for the representation of temporal dynamics in geographic information systems', In: *Annals of the Association of the American Geographers* 3, Cambridge and Oxford: Blackwell Publishers, pp. 441-461.

RODIER, X., SALIGNY L. (2008). 'Social features, Spatial features, Time features: An urban archaeological data model', In: Posluschny A., Lambers K., Hertog I. (eds.). *Layers of Perception, Proceedings of the 35th International Conference on Computer Applications and Quantitative Methods in Archaeology (CAA), Berlin, Germany, April 2-6, 2007*, Kolloquien zur Vor-und Frühgeschichte, vol. 10, Bonn: Dr. Rudolf Habelt GmbH, pp.

RODIER, X., SALIGNY, L. 'Modélisation des objets historiques selon la fonction, l'espace et le temps pour l'étude des dynamiques urbaines dans la longue durée'. In, *Cybergeo: European Journal of Geography, Systems, Modelling, Geostatistics*, document 502, online since 17June 2010. Available from: http://cybergeo.revues.org/index23175.html.

RODIER, X., SALIGNY, L., LEFEBVRE, B., POULIOT, J. (2010). 'ToToPI (Topographie de Tours Pré-Industriel), a GIS for understanding urban dynamics based on the OH_FET model (Social Use, Space and Time)'. In *Making History Interactive*, edited by, Fricher B., Crawford J. and Koler D. (Computer Application and Quantitative Methods in Archaeology (CAA), Proceedings of the 37th International Conference, Williamsburg, Virginia, USA, March 22-26, 2009, BAR International Series 2079, Oxford: Archaeopress, pp. 337-349.

SAINT-GÉRAND, T. (2005). Comprendre pour mesurer... ou mesurer pour comprendre ?, In: Guermond Y. (ed.) *Modélisations Géographiques,* Paris: Traité IGAT, Hermès, pp. 49-74.

SIMON, G. (2012). 'Modélisations multi-scalaires des dynamiques urbaines dans la longue durée : l'exemple du quartier abbatial de Vendôme (41)', *Cybergeo: European Journal of Geography, Systems, Modelling, Geostatistics*, document 589, Online since 06 February, Available from: http://cybergeo.revues.org/25110 ; DOI : 10.4000/cybergeo.25110

THÉRIAULT, M. & CLARAMUNT, C., (1999). La représentation du temps et des processus dans les SIG : une nécessité pour la recherche interdisciplinaire, In *Représentation de l'espace et du temps dans les SIG*, Revue internationale de géomatique 9, Paris: Hermès, pp. 67-99.

An 'Alphabet' to Describe the Spatio-Temporal Dynamics of Settlement Systems: A Relevant Representation of Time?

Marie-Jeanne OURIACHI,[1] Frédérique BERTONCELLO,[1]
with the collaboration of Pierre NOUVEL,[2] Laure NUNINGER,[3] Elise FOVET,[3] Stephane ALIX[4]

[1]Université de Nice-Sophia Antipolis, CEPAM-UMR 7264, GDR 3359-MoDys, France
[2]Université de Franche-Comté, Chrono-Environnement-UMR 6249, France
[3]Université de Franche-Comté, Chrono-Environnement-UMR 6249, GDR 3359-MoDys, France
[4]INRAP, France
marie-jeanne.ouriachi@unice.fr, Frederique.Bertoncello@cepam.cnrs.fr, pierre.nouvel@univ-fcomte.fr, laure.nuninger@univ-fcomte.fr, elisefovet@yahoo.fr, stephane.alix@inrap.fr

Abstract
This paper presents a work in progress within the ANR project Archaedyn, which concerns the development of an 'alphabet' designed to characterize and compare the evolution of the settlement pattern in several French microregions during the Roman period.
The 'alphabet' is a composite indicator, built from 3 criteria relating to the quantitative, hierarchical and spatial evolution of the settlement pattern. The different combinations of these criteria form the characters of the 'alphabet'. The evolution of the settlement system of each microregion between two periods can thus be written as a chain of characters.
It is then possible to compare these 'signatures' of the different microregions throughout time.
This indicator has a strong temporal dimension as it is designed in terms of stability/instability with the previous situation. As it Summarizes the evolutions of the settlement pattern on the long term, it is an efficient tool to identify major types of dynamics.

Keywords: Settlement dynamics; hierarchy; spatial structure; composite indicator; time formalisation.

Introduction

Within the Archaedyn project,[1] the workgroup 2 'Settlement patterns and territories' aims to analyse the organisation, intensity and stability of rural settlement from 800 BC to 800 AD in ten microregions of Southern and Central France. The data come from field surveys and only include remains of settlements as defined by the geographer R. Brunet: 'The settlement ('l'habitat') is the grouping and the layout of dwellings in a given space; it can include annexes for livestock and storage, as well as workshops and other buildings for professional use' (Brunet et al., 1992: 229). Diachronic and interregional comparisons require to use common indicators to describe the settlement patterns in each area. Several indicators were developed to characterise the intensity, hierarchical organisation and spatial structure of the settlement, each of them describing the situation of the settlement pattern in an area in one century (Gandini et al., 2008, Bertoncello et al, 2012a, Bertoncello et al., 2012b). In order to describe the dynamics between these states, three of these individual indicators were used to create a composite indicator summarizing the evolution of the settlement patterns in terms of intensity, hierarchical and spatial structures. This paper presents the building of the composite indicator and discusses the time formalisation it induces. It has been applied to four French microregions: the Yonne valley in Burgundy (Yonne), the eastern part of the Languedoc (Hérault, Gard), the Argens valley and Maures massif (Var), and the pre-alpine hills of Grasse ('Préalpes', Alpes-Maritimes).

1. The individual indicators

1.1 Quantitative dynamic of the settlement

The dynamics underlying the evolution of the number of settlements throughout time can be assessed by calculating the rate of pre-existing, created and deserted settlements on the total number of occupied settlements per century in each microregion. For example, in the case of the Languedoc area (fig.1), the curve of the number of occupied settlements peaks on the 1st c. AD and on the 4th c. AD. Although the number of settlements increases regularly from the 2nd c. BC to the 1st c. AD, the relative proportions of pre-existing, created and abandoned settlements during this period show a strong instability: in each century, a lot of settlements disappear while new ones are created. On the 4th century, the settlements inherited from the 3rd century play a more important role in the increase of the number of occupations.

1.2 Hierarchical profile of the settlement

In a systemic approach, the structure and the evolution of the system depend on the interactions between the system's components. In the case of the settlement systems, the relative hierarchical position of the settlements determines the nature of their interactions (Durand-Dastès et al., 1998). We assume that the hierarchical range of the settlements can be inferred from the field survey data (surface area, building materials,

[1] 'Spatial dynamics of territories from Neolithic to Middle Ages': ACI 'Spaces and Territories', French Ministry for Research and New Technologies, contract ET28 2005-2008; Program of the French National Agency of Research, contract ANR-08-BLAN-0157-01, 2009-2012, coordinated by F. Favory and L. Nuninger (UMR 6249- USR 3124, Besancon, France).

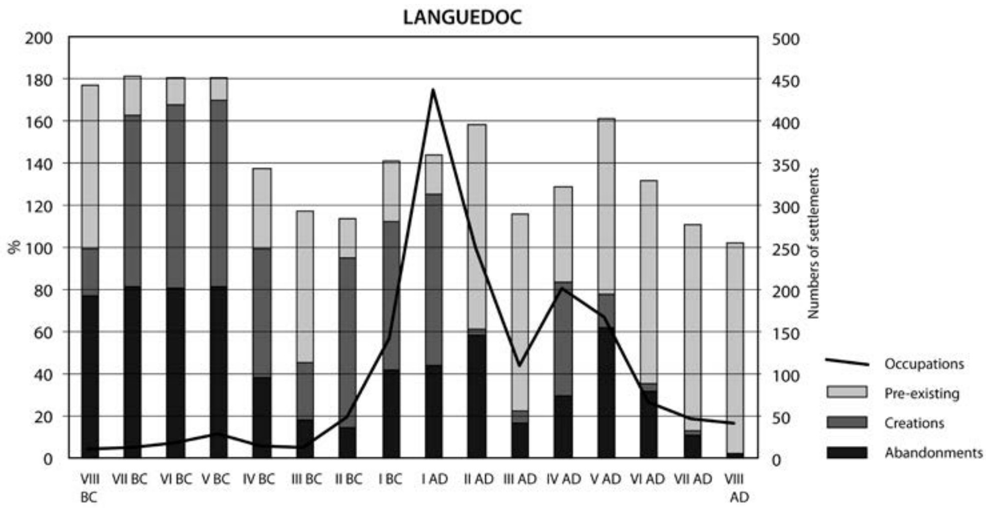


Figure 1: Quantitative indicator of the settlement dynamic in the Languedoc area (the rates of pre-existing, created and deserted settlement are calculated on the total number of occupied settlements per century. As a settlement can be occupied and abandoned during the same century, the total of the three categories can exceed 100 %).

duration, function: Gandini et al, 2008). The classification (Factor Analysis and Agglomerative Hierarchical Clustering) of 1491 settlements from the ten microregions according to these four criteria defined five classes of settlements, which we interpreted as five hierarchical levels (tab. 1).

The distribution of the settlements per class defines the hierarchical profile of each area in each century. The regional histograms show four types of distributions, which correspond to different hierarchical structure of the settlement (fig. 2).

1.3 Spatial structure of the settlement

Several indicators were used to characterise the spatial structure of the settlement. Among them, we selected the local spacing to build the composite indicator: in each microregion, the average distance from each settlement to its five nearest neighbours is calculated per century. The resulting histograms show the level of concentration / dispersion of the settlement pattern (fig. 3).

2. Elaboration of a composite indicator

The individual indicators mentioned above describe the settlement pattern in an area at one century: they thus refer to specific states of the settlement system. These indicators were calculated for four centuries (2nd c. BC, 1st c. AD, 3rd c. AD and 5th c. AD), which correspond to major changes in the evolution of the number of settlements in several areas.

The composite indicator conversely aims to describe the change between these four states: is the settlement pattern stable or unstable? What is the significance of the changes?

Two approaches were possible to create the composite indicator: starting from the analysis of the observed regional situations ('data driven') or identifying all the possible theoretical situations ('concept driven'). In this particular case, we combined these two approaches and went back and forth between them.

More specifically, the composite indicator results from the combination of the various modalities of the three selected individual indicators: the modalities form the signs of the 'alphabet'. The building of the 'alphabet' had to be parsimonious in order to avoid too many signs. This is why we limited the number of modalities by defining only the main types of evolution for each indicator.

Classes	% of settlements	Surface area	Length of occupation	Building materials	Function	Interpretation
A	4%	> 2 ha or 1 to 2 ha	> 5 centuries	Elaborated (mosaics, marble, sculpted elements) or only stones	political and/or religious and/or symbolic function	High hierachical level
B	21%	1 to 2 ha or > 2 ha	> 5 centuries	Elaborated (mosaics, marble, sculpted elements) or more ordinary (stones and tiles)	agricultural or craft activity or no known function	
C	19%	0,5 to 1 ha	3 to 5 centuries	tiles and/or stones, or more elaborated (tubuli, hypocaust bricks, painted coatings)	agricultural or craft activity or no known function	
D	28%	0,1 to 0,3 ha, or 0,3 to 0,5 ha	1 to 3 centuries	tiles, or tiles and stones	agricultural or craft activity or no known function	
E	27%	< 0,1 ha	< 1 century	tiles and stones or perishable materials or absence of materials	agricultural or craft activity or no known function	Low hierachical level

Table 1: Hierarchical classification of the settlements.

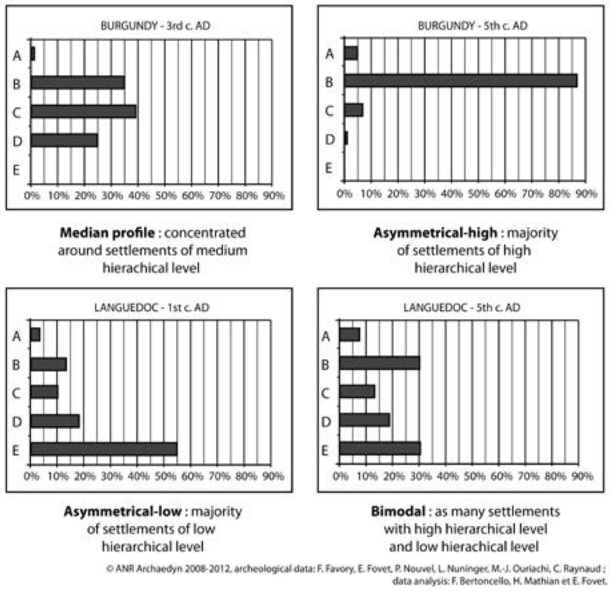


Figure 2: The four types of hierarchical profile.

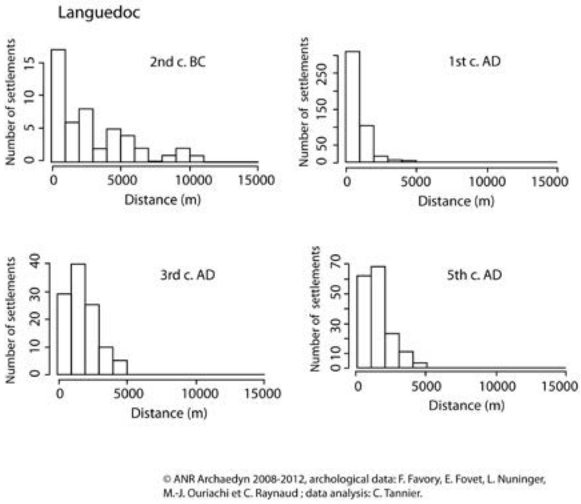


Figure 3: Local spacing indicator: the example of the Languedoc area.

Seven modalities were identified for the quantitative indicator (tab. 2): each of them combines one level of instability (high, medium, low) and one type of evolution of the number of settlements (increase, decrease or stagnation). The level of instability depends on the rate of deserted settlements during the previous century, the rate of pre-existing settlements that are still occupied during the current century, and the rate of new settlements. For example (tab. 2), a strong rate of desertions during the previous century would lead to a low rate of pre-existing settlements persisting during the century under review. If this is combined to a strong rate of creations during this century, then it shows a strong instability of the settlement pattern. However, as creations can balance the desertions, the number of settlements would not necessarily change. In case of low instability (i.e. stability), we have decided to group the three possible evolutions of the number of settlements in one modality in order to limit the number of signs.

The evolution of the hierarchical structure is evaluated through the comparison of the hierarchical profiles of the settlement in two different centuries. The sixteen possible combinations of the four types of profile (fig. 2) are interpreted in terms of stagnation, strengthening or reduction of the hierarchy (tab. 3). For instance, the profiles of the Burgundy study area (fig. 2) show a reduction of the hierarchy: the settlements distributed between three classes of medium hierarchical level during the 3rd c. AD (median type profile), concentrate during the 5th c. AD on one class (B), which corresponds to a high hierarchical level (asymmetrical-high profile).

Regarding the spatial indicator, the stagnation, increase or decrease of the average distance from each settlement to its five nearest neighbours respectively shows the stability, a tendency to dispersion or to concentration of the settlement pattern (tab. 4). For instance, the settlement pattern in the Languedoc area undergoes a strong concentration between the 2nd c. BC and the 1st c. AD, whereas it disperses slightly during the next period (fig. 3).

The addition of the modalities of these three indicators composes a thirteen characters 'alphabet' (seven for the quantitative indicator, three for the hierarchy and three for the spatial dynamic). The association of each of the

Rate of created settlements during the current century	Rate of deserted settlements during the previous century	Rate of pre-existing settlements	Number of settlements	Instability			Quantitative dynamic (instability + number of settlements)	Characters
				Low	Medium	High		
high	high	low	stagnation			X	Strong instability =	Fi =
high	medium	medium	increase			X	Strong instability +	Fi +
medium	high	low	decrease			X	Strong instability -	Fi -
medium	medium	medium	stagnation		X		Medium instability =	Im =
high	low	high	increase		X		Medium instability +	Im +
low	high	low	decrease		X		Medium instability -	Im -
low	medium	medium	decrease	X			Stability -	
medium	low	high	increase	X			Stability +	S
low	low	high	stagnation	X			Stability =	

→ 7 characters

Fi= Strong instability and stagnation of the number of settlements
Fi+ Strong instability and increase of the number of settlements
Fi- Strong instability and decrease of the number of settlements
Im= Medium instability and stagnation of the number of settlements
Im+ Medium instability and increase of the number of settlements
Im- Medium instability and decrease of the number of settlements
S- Stability and decrease of the number of settlements
S+ Stability and increase of the number of settlements
S= Stability and stagnation of the number of settlements
} **S :** stability

Table 2: Modalities of the quantitative indicator.

HIERARCHICAL PROFILE - PERIOD 1	HIERARCHICAL PROFILE - PERIOD 2	Period 1	Period 2	TENDENCY TO HIERARCHY	Characters
Asymmetrical profile with a majority of settlements of low hierarchical level	Asymmetrical profile with a majority of settlements of low hierarchical level	Asym-	Asym-	Stagnation	H=
Asymmetrical profile with a majority of settlements of high hierarchical level	Asymmetrical profile with a majority of settlements of high hierarchical level	Asym+	Asym+	Stagnation	
Median profile: concentrated around settlements of medium hierarchical level	Median profile: concentrated around settlements of medium hierarchical level	Med	Med	Stagnation	
Bimodal profile: as many settlements with high hierarchical level and low hierarchical level	Bimodal profile: as many settlements with high hierarchical level and low hierarchical level	Bim	Bim	Stagnation	
Median profile: concentrated around settlements of medium hierarchical level	Asymmetrical profile with a majority of settlements of low hierarchical level	Med	Asym-	Strengthening of the hierarchy	H+
Bimodal profile: as many settlements with high hierarchical level and low hierarchical level	Asymmetrical profile with a majority of settlements of low hierarchical level	Bim	Asym-	Strengthening of the hierarchy	
Median profile: concentrated around settlements of medium hierarchical level	Asymmetrical profile with a majority of settlements of high hierarchical level	Med	Asym+	Reduction of the hierarchy	H-
Bimodal profile: as many settlements with high hierarchical level and low hierarchical level	Asymmetrical profile with a majority of settlements of high hierarchical level	Bim	Asym+	Reduction of the hierarchy	
Median profile: concentrated around settlements of medium hierarchical level	Bimodal profile: as many settlements with high hierarchical level and low hierarchical level	Med	Bim	Strengthening of the hierarchy	H+
Asymmetrical profile with a majority of settlements of low hierarchical level	Median profile: concentrated around settlements of medium hierarchical level	Asym-	Med	Reduction of the hierarchy	H-
Asymmetrical profile with a majority of settlements of high hierarchical level	Median profile: concentrated around settlements of medium hierarchical level	Asym+	Med	Strengthening of the hierarchy	H+
Bimodal profile: as many settlements with high hierarchical level and low hierarchical level	Median profile: concentrated around settlements of medium hierarchical level	Bim	Med	Reduction of the hierarchy	H-
Asymmetrical profile with a majority of settlements of low hierarchical level	Bimodal profile: as many settlements with high hierarchical level and low hierarchical level	Asym-	Bim	Reduction of the hierarchy	
Asymmetrical profile with a majority of settlements of high hierarchical level	Bimodal profile: as many settlements with high hierarchical level and low hierarchical level	Asym+	Bim	Strengthening of the hierarchy	H+
Asymmetrical profile with a majority of settlements of high hierarchical level	Asymmetrical profile with a majority of settlements of low hierarchical level	Asym+	Asym-	Strengthening of the hierarchy	
Asymmetrical profile with a majority of settlements of low hierarchical level	Asymmetrical profile with a majority of settlements of high hierarchical level	Asym-	Asym+	Reduction of the hierarchy	H-

→ {
H= : Stability of the hierarchical profile
H+ : Strenghtening of the hierarchy
H- : Reduction of the hierarchy

Table 3: Modalities of the hierarchical indicator.

modalities from each of the three indicators produces sixty-three possible combinations or 'trigrams'. As these 'trigrams' integrate the different indicators we retained to describe the settlement's dynamic, we have called them the 'signatures' of the evolution of the settlement pattern in each area within a period of three centuries (tab. 5).

For example, the signatures of the Argens, Languedoc and Burgundy areas are similar between the 2nd c. BC and the 1st c. AD: 'Fi+C+H+', which means that the settlement dynamic is characterized by a strong instability with an increase of the number of settlements, leading to a more concentrated pattern and a stronger hierarchy. The Préalpes area shows the same evolution except for the hierarchical profile (H=), which does not change between the 2nd c. BC and the 1st c. AD (tab. 5).

To make the 'trigrams' more explicit, we have translated them into 'ideograms': a sign or a grey scale represents each modality of the indicators. In this new representation, a distinction is made between the number of settlements and the settlements dynamics, defined in terms of stability or instability (fig. 4).

The application of these 'ideograms' to four microregions shows the interest of this representation to visualise the evolution of the settlement pattern in each area throughout time and to compare the microregional dynamics (fig. 5). At a microregional scale, the variety of the 'ideograms' shows the diversity of the evolutions from one period to another. On an interregional level, the 'ideograms' highlight the singularity of the Burgundy area between the 1st and the 3rd c. AD: whereas all the areas experience a medium instability characterised by the decrease of the number of settlements, a more dispersed settlement pattern and a reduction of the hierarchy, the situation remains very stable in Burgundy during this period. Indeed this type of evolution is delayed to the next period in Burgundy (3rd c. AD – 5th c. AD). It is also interesting to note the diversity of the microregional

Spacing (average distance from each settlement of its five nearest neighbours)	Concentration / dispersion	Instability	Characters
Strong increase	Strong dispersion	X	C-
Moderate increase	Moderate dispersion	X	C-
Small increase	Small dispersion		C=
Strong decrease	Strong concentration	X	C+
Moderate decrease	Moderate concentration	X	C+
Small decrease	Small concentration		C=
Stagnation	Stagnation		C=

→ {
C+ : Concentration of the settlement
C- : Dispersion of the settlement
C= : stagnation of the spacing

Table 4: Modalities of the spatial indicator.

	Quantitative dynamic			Spacing			Hierarchical profile			Signature		
	2BC-1AD	1AD-3AD	3AD-5AD	2BC-1AD	1AD-3AD	3AD-5AD	2BC-1AD	1AD-3AD	3AD-5AD	2BC-1AD	1AD-3AD	3AD-5A
Argens	Fi+	Im-	Im=	C+	C-	C=	H+	H-	H=	Fi+C+H+	Im-C-H-	Im=C=H
Languedoc	Fi+	Im-	Im+	C+	C-	C=	H+	H-	H+	Fi+C+H+	Im-C-H-	Im+C+H
Burgundy	Fi+	S	Im-	C+	C=	C-	H+	H=	H-	Fi+C+H+	SC=H=	Im-C-H
Préalpes	Fi+	Im-	Im=	C+	C-	C+	H=	H-	H=	Fi+C+H=	Im-C-H-	Im=C+H

Table 5: 'Signatures' of the four study areas between the 2nd c. BC and the 5th c. AD.

dynamics during this last period, as shown by the wide range of 'ideograms'.

3. Time formalisation in the 'alphabet' / 'ideograms'

The 'alphabet' and the 'ideograms' are two ways of formalising spatio-temporal processes, that allow to go from the description of states to that of the settlement dynamics. Both of them include different levels of perception of the dynamics (fig. 6).

At the data level (the settlements), the century is the minimum temporal object due to the smallest available chronological resolution of the archaeological remains in the various microregions. This time granularity implies to consider each century as one state, fixing the settlement pattern in a stable situation, although it necessarily underwent transformations during these hundred years. We are aware of the possible impacts of this time discretisation on our perception of the settlement dynamics and it would be interesting to test the robustness of our model by reducing the time granularity when possible.

In our model, the 'ideograms' synthesise the settlement dynamic within a period of three centuries: each microregional signature thus corresponds to a first level of generalisation, i.e. intra-periodic. The comparison of the microregional signatures throughout time is a second level of generalisation (inter-periodic). At the first level, the dynamic results from the comparison of states (the centuries), whereas it results from the comparison of dynamics (signatures) at the second level (fig. 6). As an example, the settlement dynamic in the Argens area between the 2nd c. BC and the 5th c. AD can be described as follows (fig. 5): after a phase (2nd c. BC – 1st c. AD) of intense restructuring of the settlement pattern (strong instability due to high rates of desertions and creations, which leads to an increase of the number of occupations, the concentration of the settlement pattern and the strengthening of the hierarchy), the second period (1st c. AD – 3rd c. AD) corresponds to a phase of skimming (medium instability characterised by a lot of desertions with very few creations, which leads to a decrease of the number of settlements, a more dispersed pattern and a reduction of the hierarchy towards the higher hierarchical levels). The last period (3rd c. AD – 5th c. AD) shows the stabilisation of this settlement structure (medium instability due to persisting desertions compensated by some creations, stagnation of the settlements' number, spacing and hierarchical range).

In such a formalisation, the long duration (i.e. the settlement dynamic during seven centuries) is perceived through the addition of short durations (i.e. three phases of three centuries). This is thus a discontinuous approach, which does not necessarily reflect the possible different temporalities and overlaps of the various phenomena implied in the evolution

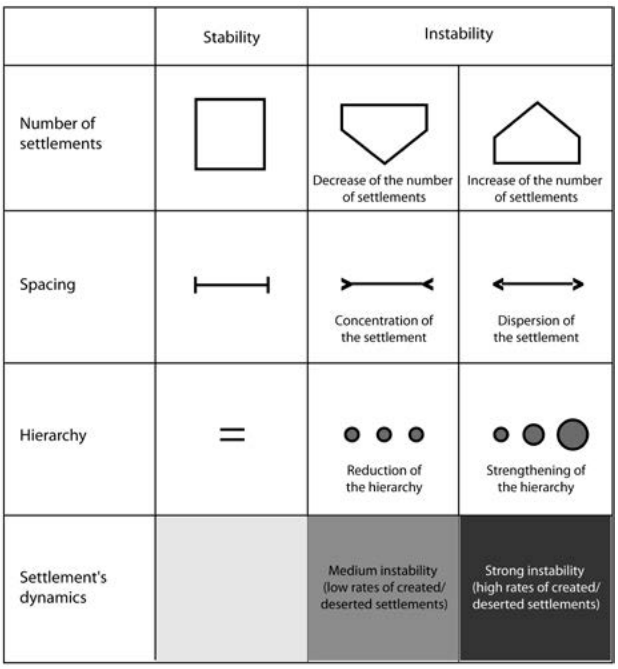


Figure 4: Table of the 'ideograms'.

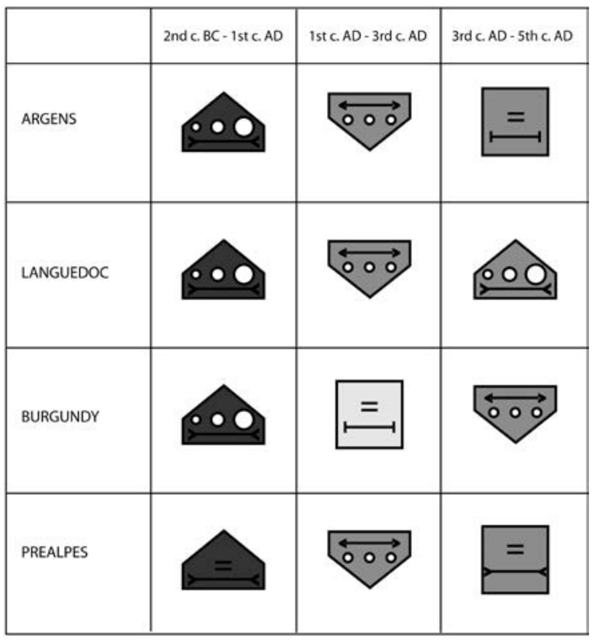


Figure 5: Representation of the microregional dynamics of the settlement using the 'ideograms'.

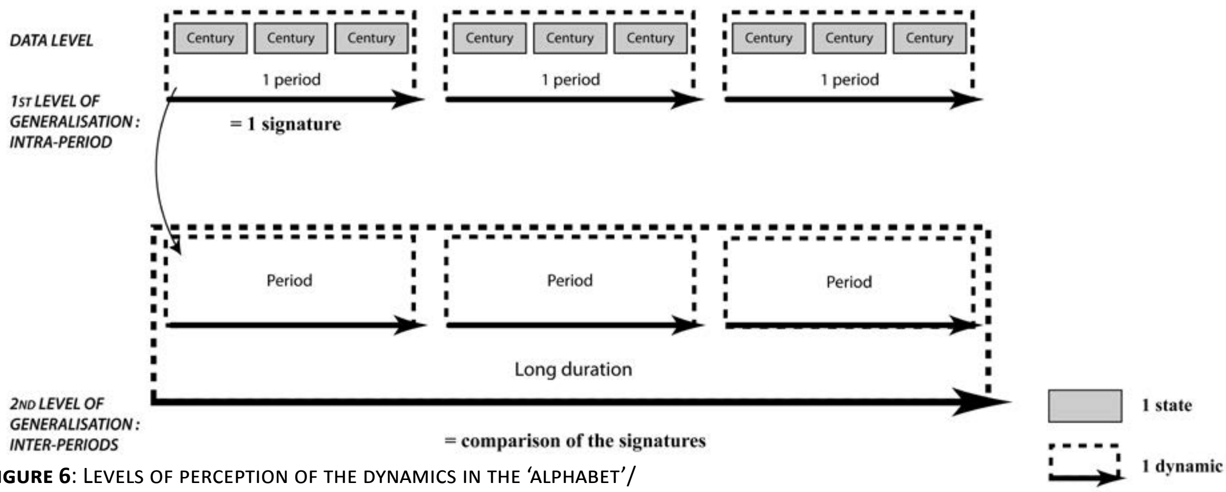


FIGURE 6: LEVELS OF PERCEPTION OF THE DYNAMICS IN THE 'ALPHABET'/ 'IDEOGRAMS'.

of the settlement pattern. For instance, the temporality of the process of settlement creation is much shorter than that of hierarchical structuration. Nevertheless, this discontinuity is balanced by the focus we put on the dynamics: as each signature is compared to the previous one in order to qualify the change, this formalisation forces to link the different stages of the settlement system.

Conclusion

The 'alphabet' and the derived 'ideograms' respond to the purpose to describe the spatio-temporal dynamics of the settlement system rather than its successive states. Summarizing various indicators, this tool allows to characterise the dynamics by integrating their multiple components instead of deducing them from the juxtaposed descriptions of the individual indicators. This mode of visualisation of the dynamics greatly facilitates the comparisons through time and space, both at interregional and intraregional scales and in a diachronic or synchronic perspective.

Concerning time, the 'alphabet'/ 'ideograms' articulate several durations, from the century to the 'canonical' period i.e. the Antiquity (2nd c. BC – 5th c. AD). The dynamics are first perceived within sequences of three centuries (the 'signatures') but the necessity to articulate and link them to compare them counterbalances this discontinuous approach. Allowing to synthetically represent the evolutions of the settlement pattern on the long term, the 'alphabet' and 'ideograms' appear to be efficient tools to identify generic types of dynamics. In the present application, only seven of the eighty-one possible ideograms are used. It would be interesting to test this tool with a higher temporal granularity (century by century) or on a larger time span (from the Iron Age to the early Middle Age) and on other areas in order to identify other types of settlement evolution. This'alphabet'/'ideograms' is thus a first step towards the identification of 'stylised facts' of the settlement dynamic, that can be used to model the evolution of the settlement systems on the long term (Sanders, 2006: 154).

References

BERTONCELLO, F., FOVET, E., GANDINI, C., TRÉMENT, F. & NUNINGER, L. (2012a). The spatio-temporal dynamics of settlement patterns from 800 BC to 800 AD in Central and Southern Gaul: models for long-term interregional comparison. In: Gandini, C., Favory, F. & Nuninger, L. (eds.). *Archaedyn: 7 millennia of territorial dynamics: settlement pattern, production and trades from Neolithic to Middle Ages.* Colloque de Dijon 23-25 juin 2008. Oxford (UK), Archaeopress.

BERTONCELLO, F., FOVET, É., TANNIER, C., GANDINI, C., LAUTIER, L., NOUVEL, P. & NUNINGER, L. (2012b). Configurations spatiales et hiérarchiques du peuplement antique : des indicateurs quantitatifs pour une confrontation interrégionale. In: Bertoncello, F. & Braemer, F. (eds.). *Variabilités environnementales, mutations sociales : Nature, intensités, échelles et temporalités des changements*. Antibes: APDCA.

BRUNET, R., FERRAS, R. & THÉRY, H. (eds.) (1992) *Les mots de la géographie : dictionnaire critique*, Paris, GIP RECLUS / La Documentation Française.

DURAND-DASTÈS, F., FAVORY, F., FICHES, J.-L., MATHIAN, H., PUMAIN, D., RAYNAUD, C., SANDERS, L. & VAN DER LEEUW, S. (1998). *Des oppida aux métropoles : Archéologues et géographes en vallée du Rhône*, Paris, Economica.

GANDINI, C., BERTONCELLO, F., with the contribution of GAUTHIER, E., NUNINGER, L. & TRÉMENT, F. (2008). Hierarchical typology and settlement patterns modeling at inter-regional scale. In: Posluschny, A., Lambers, K. & Herzog, I. (Eds.). *Layers of Perception. Proceedings of the 35th International Conference on Computer Applications and Quantitative Methods in Archaeology (CAA)*. Berlin, Dr. Rudolf Habelt GmbH, Bonn.

SANDERS, L. (2006). Les modèles agent en géographie urbaine. In: Amblard, F., Phan, D. (eds.) *Modélisation et simulation multiagents ; applications pour les Sciences de l'Homme et de la Société.* Londres, Hermes-Sciences & Lavoisier.

Chapter 6. GIS and Spatial Analysis

ArkeoGIS,
Merging Geographical and Archaeological Datas Online

Loup BERNARD,[1] Damien ERTLEN,[2] Dominique SCHWARTZ[2]

[1]Université de Strasbourg, UMR 7044 ARCHIMEDE, France
[2]Université de Strasbourg, UMR 7362 LIVE, France
loup.bernard@unistra.fr, damien.ertlen@live-cnrs.unistra.fr, dominique.schwartz@live-cnrs.unistra.fr

Abstract
After a few years of existence, ArkeoGIS is now fully operational in its version 3. This webGIS adds over 20 databases in French and German, with over 10 000 archaeological sites and environmental features. Inventorial databases, searchers databases, students and laboratories cooperate in the Rhine valley, implementing each other datas, and having access to cross-boundary information and literature in two languages. Users and contributors, once logged, have the possibility to build quick queries with a handy interface, questioning several informations on four depth levels. For this paper, the focus will be on the benefit happening while showing archaeological sites and finds on an interactive map where the user can add geographical inventories, existing analysis and how this helps us finding new questionings.

Keywords*:* Gis, Geoarchaeology, Cross-Border Studies, Arkeogis

1. Tool inception issues

ArkeoGIS software was incepted in 2009 in order to gather archeological and environmental datas regarding the Upper Rhine valley, an area providing a whole range of excellent research centres in archeology as well as geography, although they are not well connected. Besides linguistic problems, traditional research methods and archeology-linked legislations result in a state of datas particularly hard to merge with already existing digital tools.

This area presents many assets: its fair amount of already georeferenced sites inside inventory tools depending on the ministry of Culture (French Patriarche, German AdaB, either providing several thousands referenced sites); and research dynamics inside universities and research centres (research groups, researchers and research-professors, doctoral students' works, Master students) as well as rescue archeology teams, mainly in France. INRAP (Institut National de Recherches Archéologiques Préventives) and PAIR (Pôle Archéologique Interdépartemental Rhénan) archeologists and geo-archeologists are also important providers of data while these are not systematically published, remaining mostly a part of « grey literature ». Geographers' works are no less scattered around (Universities, BRGM – Bureau de Recherches Géologiques et Minières…)

Accessing to this datapool is therefore terribly time-consuming, while it is a basic tool for all (mainly young researchers) who work earnestly on relations between man and environment from Neolithic to Medieval Ages in the Rhine Valley. Besides, many colleagues are not bilingual, which prevents them from easy access – even a synthetical one – to the across-river state of research.

Such a general conclusion led us to quickly set online a simple, free and bilingual software, allowing users to find accurate access to researchers, sites, analyses, as well as pooling all or any of archeologists' and geographers' datas. Interdisciplinarity – or rather, working together – is an especially basic fact in the Rhine Valley, where geographical strains and soil erosion and capping are strong issues. In the following article, I wish to put forward the operation of the software and on-going issues.

2. Software operating

ArkeoGIS (www.arkeogis.org) is an opensource online free software, for which users need only an Internet access and a navigator. The coding of up-to-date 3.0 version is available on github (http://arkeogis.org/page/logiciel). This latter version allowed us to adapt and develop a few more tools than the former versions (cf. Bernard 2012 and 2014) and it should not evolve before some time. Many difficulties regarding language, chronology, typology and map-designing traditions are now under control. Protected through a login/password, the website provides a sole demonstration version (http://demo.arkeogis.org/login) to the public at large, so as to prevent site-plunderers to access sensitive archeological data. The Alsace historical atlas (http://www.atlas.historique.alsace.uha.fr/) provides online the first maps designed with our free software, along with a comment note, and general broadcasting to the public at large.

ArkeoGIS has no database per se. ; the software works as a databases feed agregator. Today, more than 25 French and German bases provide data about several tens of thousands archeological sites and objects, as well as many analyses such as sedimentology and pedology. One of the software's main callings is to provide wide access to all researchers through an interface allowing intuitive ways of putting a request, which answer will be displayed instantly on a chosen outline map. Too many archeologists are expecting this from a SIG (a system that has sometimes been called SIA – Archeological Data System). For the hundred people or so who are using the site now, the general conclusion is that these maps and requests are sufficient to chart raw data, to emit hypotheses, to visualize them and to cross-examin results on various outline maps. Experts in databases or SIG have direct access to files, enabling them to include said files into their usual softwares.

Basically, once connected, users access to an interface centered on a map of the Rhine Valley ; they then choose a working language (French or German) and, using a specific toolbox, proceed to build their request, which can be saved later. Three "yes/no" boxes allow the data quality to be refined, depending on their being accurately georeferenced or situated on the municipality's centroid ; another box sorts out searched, surveyed, assessed or mentioned-in-literature sites. A last box allows focusing on unique-occupated sites.

In order to build a request, users must choose at least one database, one age and one criterium among: real estate (features, sites), movable items, production items (mining sites, natural resources...) and landscapes. To this day, 28 bases have been incorporated to ArkeoGIS, as various as: inventory bases from regional services (Patriarche and AdaB), researchers' bases, advanced students, engineers (mainly for geographical data). Some published works have also been made into databases so as to receive direct requests; other analyses listings and negative assessment results stand waiting to be integrated to the software. Each base is labelled with: its creator's or administrator's name (for the inventory bases), the last update, the precision scale, the number of sites, chronological and geographical range of the study as well as contents description; quotations will soon be made easier thanks to an incoming ISSN. We have received very positive comments; the possibility to incorporate a database has renewed interest for achieved works, which are cited more often in bibliography. Researchers choose to put online all or part of their results, leaving to ArkeoGIS users the opportunity to get in touch with them, should they need access to further information.

A selection of research area was settled, allowing to choose between the whole displayed map or reducing the area pointed by the software. This has quickly prooved a necessity, given the huge amount of available sites on ArkeoGIS. Users are then able to search a whole map or a mere rectangle or a disc, or even inside a given radius around a given point (usually the studied site).

After they have selected one or several bases, users choose one or several ages, thanks to a four-level directory system; for instance, « Iron Age, Hallstatt D, HaD1, Ha D1a ». The chronological system is therefore a relative one, very convenient for archeologists; geographers data are using an undeterminate dating system, time scales being too much different. The directory system also allows the exclusion of any given age (Hallstatt C out / Hallstatt D in, for instance) which helps in not missing wrongly dated sites, since a request about an accurate periodization will provide all the defaulting results linked to the requested one. To continue with the same example, a HaD1a site will also appear in a request to identify all Iron Age ore deposits. An accurate chronology, which absolute dating value is a matter of debate, allows searchers who work on movable items (clothes, weapons, ceramics, etc.) to incorporate their bases with no loss in data quality. Obviously, habitat-scaled or necropoles databases are documented with wider dating values, or even undeterminate ones whenever dealing with necropoles or fortified walls that have not yet been searched. ArkeoGIS was developed by and for searchers whose working areas range from the Bronze Age to Upper Middle Ages; thus, historical and pre-historical data are not accounted for.

Users must then select one or several criteria regarding: the site, movable items, production resources and landscapes. These categories are subdivised in several levels of accuracy. This versatility allows requesting all habitat sites as well as all thermal baths or all amphoras, for instance, but also to cross-reference criteria in order to reveal the richest sites (in this case, romano-gaulish habitat and amphora and baths linked to a villa). An "and/or" operator allows more accurate requests – resulting in fewer outcomes – or more open requests – resulting in near-to-exhaustive lists –, depending on documented bases. All users thus utilize the same criteria to launch requests on their own database or their working group ones, as well as other colleagues' databases. This is particularly interesting for specific data regarding some artefacts studies, of which detailed re-description proove to be onerous when it amounts to repeat step-by-step an academic catalogue, as well as regarding production data. For instance, a site, known to be linked with metal extraction during Medieval Ages, will allow antiquity researchers to look for former exploitation of this resource which could have been overlooked.

The landscape request-box allows to learn about negative assessments, analyses, surface formations or modern structures. These data are simply listed, so as to make it easier for archeologists to know whether their working site has or has not been already analysed, and if said analyses are intra- or extra-site. This allows to hone new requests for further analyses and to plan them more efficiently. Boundary areas are also entitled to reveal whether across-river colleagues already have studied their area. We have listed the analyses most frequently made on sites (14C, anthracology...) as well as works regarding surface formations (paleocanals, paleosoils, petrology, sedimentology) and related analyses (granulometry, phosphates, etc.). More recent land structures, like « murgers / drystone walls », fossilized track-lots (e.g.: Celtic fields), convex fields, etc., are also listed; their main interest is to allow the selection of new digging areas, since these have often sealed former levels.

Once the databases, Ages and criteria are determined, the request can be designed; users may then display whether a table or a map of results.

The designed request will be displayed on the selected outline map; beside free outlines (openstreetmap, mapquest, etc.), several other free outlines are available through partner .wms servers, allowing access to several sets of aerial views, more or less recent maps, digital models of grounds, etc. This proves to be very useful for archeologists, whose access to cartographic data is less open than to geographers. Icons are created at will, and linked to the state of research (the bigger the dot, the more the site has been researched), a color-coded chronology, and the accuracy of dots location (with a « # » standing for the municipality centroid). Users may present up to eight various requests in a row and save the more relevant ones.

ArkeoGIS is not an online maps publishing website; sites and landscape items are indeed displayed on outline maps, but the editing and set-up tools are not included in the project. For those who wish to create a publishable map, dots matching the request must be forwarded to a matching CAD or GIS software. To achieve this as well as to improve research, users can export their request under .csv format.

The raw .csv format allows importing and exporting databases; users are free to import or re-import one or any databases depending on: their work in progress, the implementation of their bases from their colleagues' bases, or even brand new results. Once exported, the file corresponding to the ArkeoGIS request can easily be imported to a spreadsheet, a database or a GIS software, so as to be edited or used to create maps.

Now that the tool has been described, we will deal with some of the issues for which it prooves particularly useful, for geographers as well as archeologists.

3. A few instances of geo-archeological issues:

The availability of a wide archeological database of deposits and landscape items as well as related analyses allow the scientific community working on the area to be more efficient on some issues. I wish now to stress upon a few instances regarding empty plots on maps and their interpretation through geomorpholgy and pedology as well as relations between men, environment, landscape, and their processing by ArkeoGIS.

3.1 Empty spots

Empty spots on any archeological site bring out certain questions. Whenever this is not due to the state of research, one might question the existence of forests, or erosion or overburden ground. Conversely to previous issues, this work is rather undertaken by archeologists, when they need to know whether an area is really empty or if the information have not been recorded into the ground because of taphonomic biases. Regarding this point, listing paleosoils (which formation may take several centuries, showing a steady environment), analyses made by rescuing geo-archeologists are particularly strong entry points. Buried depth of ruins, depending on period and location between mountain, piedmont, floodplain and river, obviously complements these issues. The first working sessions on this issue reveal that, were the general models of mountainside erosion (thus capping the piedmont) be valid, the line corridor studies (TGV) show that modelization is more complex than it seems at first.

Before the Rhine's course was artificially traced in the XIXth Century, the river was able to freely change its course in the Valley. So far, geographers had hard times accessing to relatively simple data in horizontal stratigraphy related to archeological sites datings. Thanks to ArkeoGIS, a few clicks offer the possibility to list sites by wide Age ranges, thus revealing areas devoid of all settlements, surely corresponding to some ancient riverbed track, whether the latter has eroded former sites or no settlement ever occurred. Conversely, a well-preserved two thousand years-old Roman settlement is

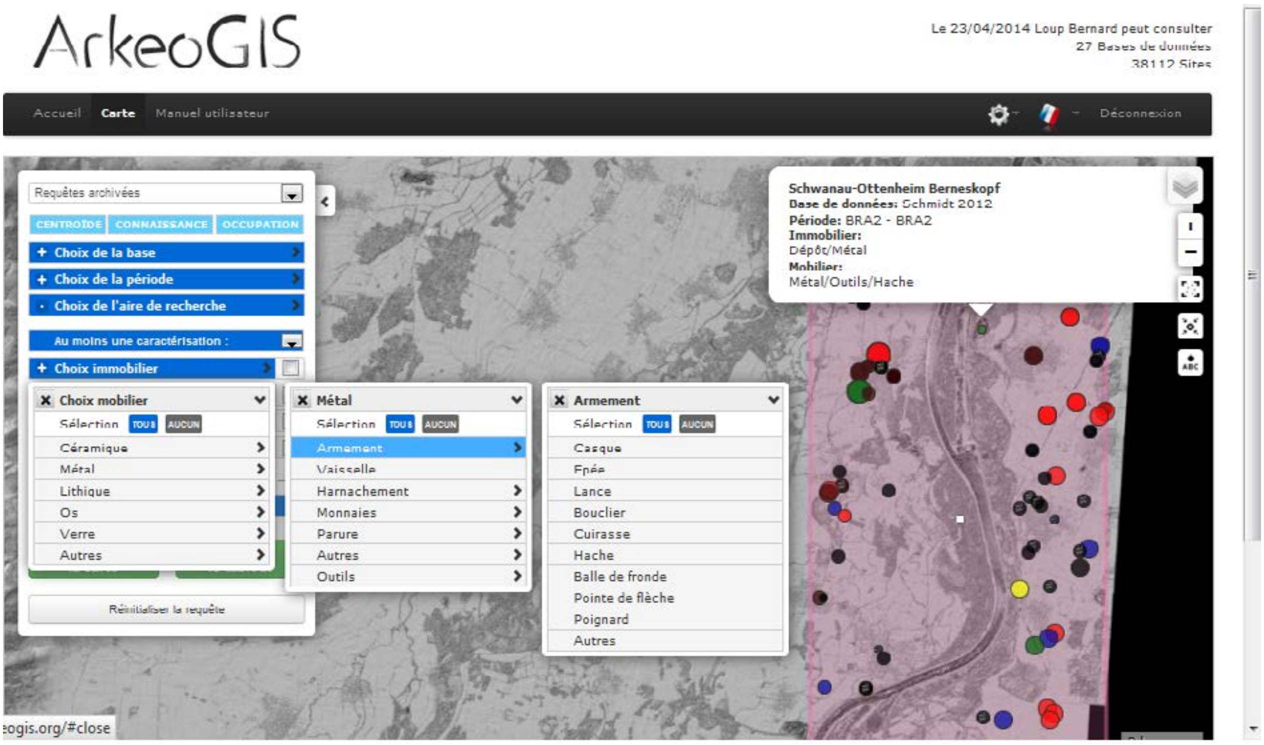


Figure 1 : Shows this fact, projected on a digital model of terrain; sites on either side of the river reveal a wide band that must match ancient riverbeds. Roman sites account for a stabilisation of the riverbends, compared to the previous period.

a sure sign that neither the river or its tributaries have had any activity on this site ever since. ArkeoGIS thus prooves to allow a first relative dating value for some ancient meanderings of the Rhine river. Insofar, it allows identifying areas where further surveys or analyses will lead to adress Rhine course issues with more interdisciplinary methodologies (Figure 1).

We will now mention another issue, regarding the alluvial fan located at the outlet of the Fecht valley in the Rhine Graben, near the modern town of Colmar and the adjoining Ill alluvial plain. The density of known archeological sites and their stratigraphical locations inside the river fan provides various answers to issues addressing mountainside sites erosion and capping of piedmont and plain sites. An interdisciplinary study is now in undertaking (Bernard, Campaner, Ertlen, Schneider, to be published) so as to put several hypotheses to the test, regarding both rivers courses meanderings in this particular area (Figure 2a, 2b, 2c).

3.2 Relations between men and environment

Archeologists have too often overlooked the rendering of today landscapes as ancient. The difficulty of pooling data concerning recent and ancient evolutions related to farming habits and general anthropisation accounts for related works still being in the bud.

We will now deal with two works in progress to show how advantageous it is to agregate recordings of this men/environment interactions.

The first of these works in progress (Schwartz *et al.,* 201) deals with murgers, i.e. linear shaped drystone walls which location have widely fossilized the track-lots and ancient soils. Their settlement period is being analysed (it is usually thought to be between the Xth and the XIIIth centuries), but in this case too, ArkeoGIS allows an access to contextual environment that may or may not deliver dating values wherever these drystone walls are capping vestiges. Regarding rescue archeology, the knowledge that a many century-old drystone work is present on location should be an incentive to dig near it or ideally below it, inasmuch the most ancient levels would have been preserved from damages due to erosion, farming and anthropisation (Figure 3).

Yet another case are the slot-pits, i.e. deep and narrow excavations, mostly dated from the Neolithic or the Bronze Age (Achard-Corompt, 2013) which were variously studied of late, in that they usually deliver no artefact or few ecofacts. They are nonetheless very interesting in that their filling is a kind of ideal trap for superficial soil horizons which are then less anthropised but well preserved thanks to the structure depth. They provide rare information about the soil cover and its farming potential at the time of occupation. (Ertlen *et al.,* 2013). Thanks to the systematical cartography of these sites in ArkeoGIS, it is now possible to get the overall surrounding sites really faster. We may wager that soon, the understanding of this type of enigmatic features will be easier thanks to their integration to a more comprehensive archeological environment (Figure 5).

Conclusion

As a conclusion, this software aims at pooling information, while everyone remains in control of their own finds; besides, knowing that there is such a website is useful to geographers (horizontal stratigraphy, dating values, paleocanals anthropisation = Terminus Post Quem...) as well as to archeologists (existing analyses, paleosoils, floodings...), all of which enhance transboundary interests. Beyond this, our free software allows an easier overview of everyone's issues, bringing out new issues and communal works. Direct access to everyone's sites and databases enhances access and citation of works that were so far remaining confidential; it allows us to get closer to a comprehensiveness in studies for which preliminary bibliographical approach is very much facilitated by ArkeoGIS. ArkeoGIS may only be a tool, but it's functional, it's online and it's free: join the community and bring in your databases !

References

ACHARD-COROMPT N. et RIQUIER V. (dir.) (2013). *Chasse, culte ou artisanat ? les fosses « à profil en Y-V-W » : structures énigmatiques et récurrentes du Néolithique aux âges des Métaux en France et alentour, Actes de la table-ronde de Châlons-en-Champagne, 15-16 novembre 2010. 33ème supplément à la Revue Archéologique de l'Est.* Dijon: S.A.E.

BERNARD L. (2010). ArkeoGIS, développement d'un WEBSIG transfrontalier : contraintes et premiers résultats. In: Giligny F., Costa L., Djindjian F., Ciezar P., Desachy B. (eds.). *Actes des 2èmes Journées d'Informatique et Archéologie de Paris - JIAP 2010* (Paris, 11-12 juin 2010), Archeologia e Calculatori supplemento 3, pp. 153-160.

BERNARD L. (2014). ArkeoGIS v2.0, Elements d'analyse de la mise en ligne de bases multilingues sur fond cartographique: fonctionnalités, apports et limites, In : L. Costa, F. Djindjian, F. Giligny (eds.). *Actes des 3èmes Journées d'Informatique et Archéologie de Paris, JIAP 2012*, Archeologia e Calcolatori supplemento 5. pp.228-237.

ERTLEN D., FELIU C., MICHLER M., SCHNEIDER-SCHWIEN N., SCHNEIKERT F., THOMAS Y. (2013). avec la coll. de BRAGUIER S., MINNI D. et VEBER C., Les structures de type « fente » dans le Kochersberg (Alsace) : un programme de recherche en cours de développement. In: Achard-Corompt N. et Riquier V. (eds.). *Chasse, culte ou artisanat ? les fosses « à profil en Y-V-W » : structures énigmatiques et récurrentes du Néolithique aux âges des Métaux en France et alentour, Actes de la table-ronde de Châlons-en-Champagne, 15-16 novembre 2010.* 33ème supplément à la Revue Archéologique de l'Est. Dijon: S.A.E.

SCHWARTZ D., ERTLEN D., BATTMANN J., CASPARD M., GEBHART A., GOEPP S., BASOGE F., KOUPALIANTZ L., METZ B. (2010). Etudes actuelles sur un type de paysage agraire encore très peu connu en Alsace : les paysages d'enclos médiévaux Extension, typologie, éléments de datation. In: Charpentier V., Marcigny C. (dir.). *Des hommes aux champs : pour une archéologie des espaces ruraux du Néolithique au Moyen Age*, Rennes : Presses Universitaire de Rennes, Collection Archéologie & Culture

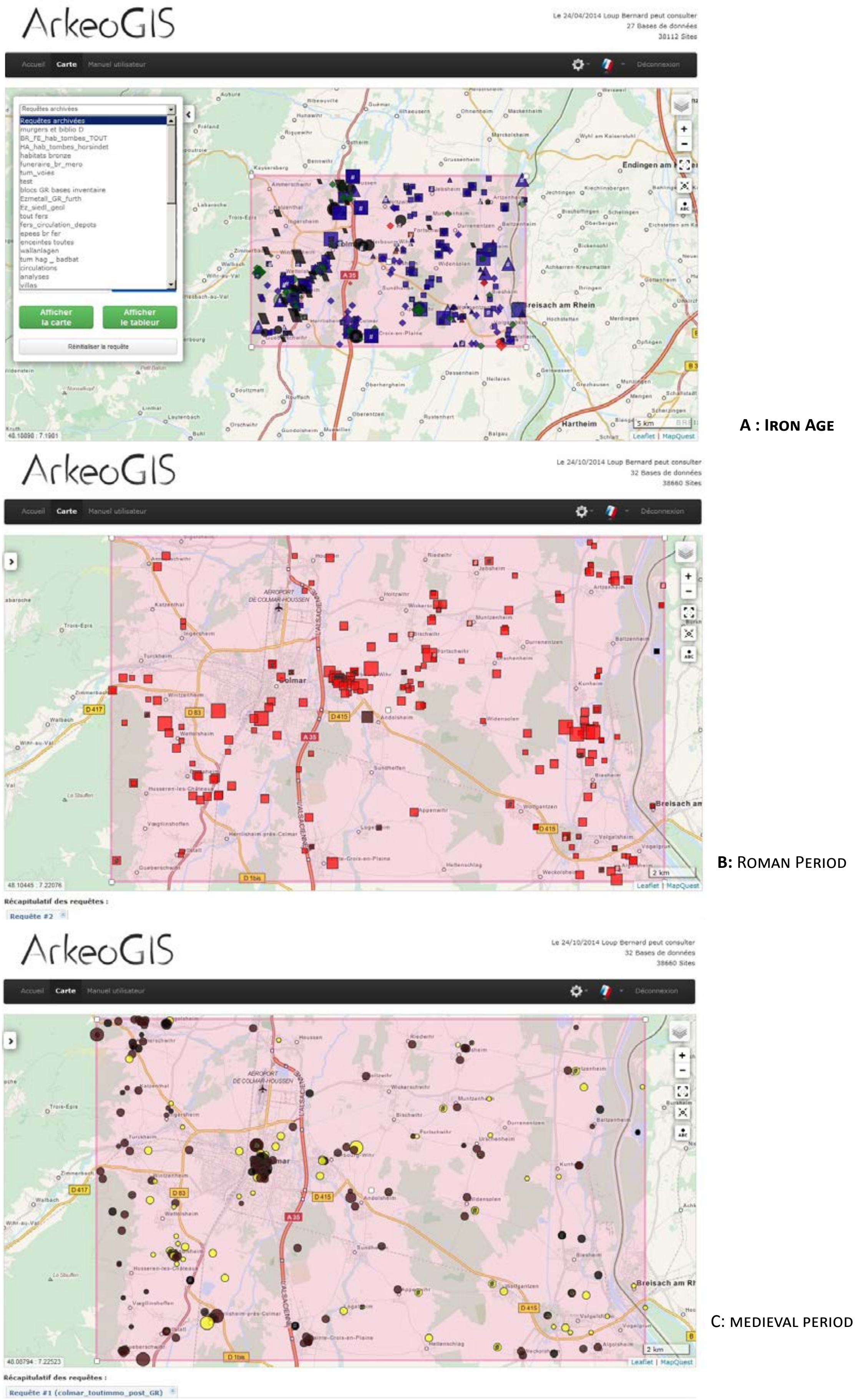


Figure 2: Notice how the north-eastern part of the area seems to be devoid of sites during the Iron Age (A) and the Roman Period (B) until the medieval period (C). A more detailed study of the aera is being undertaken.

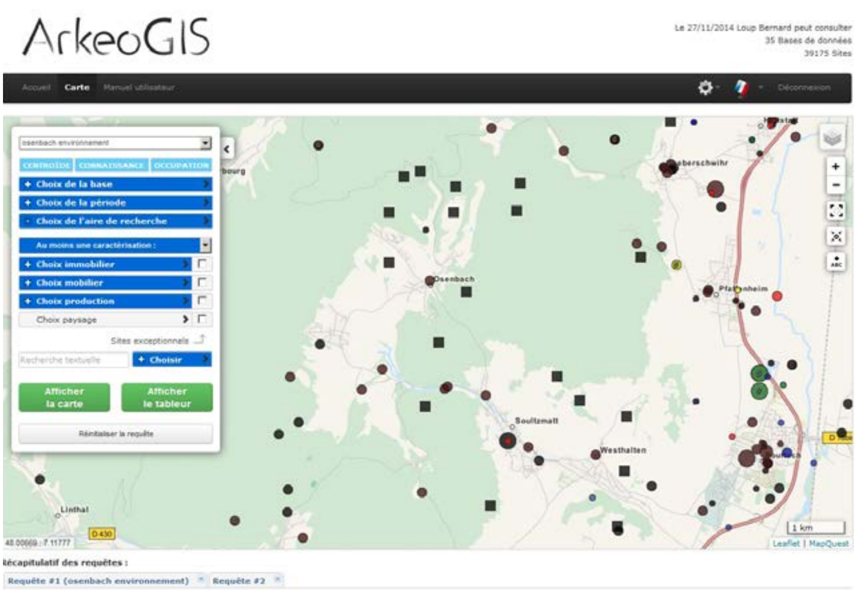


Figure 3 : displays known districts presenting murgers (squares) and their –mostly medieval- archeological *environment (dots).*

Figure 3*bis*

Figure 3*ter*

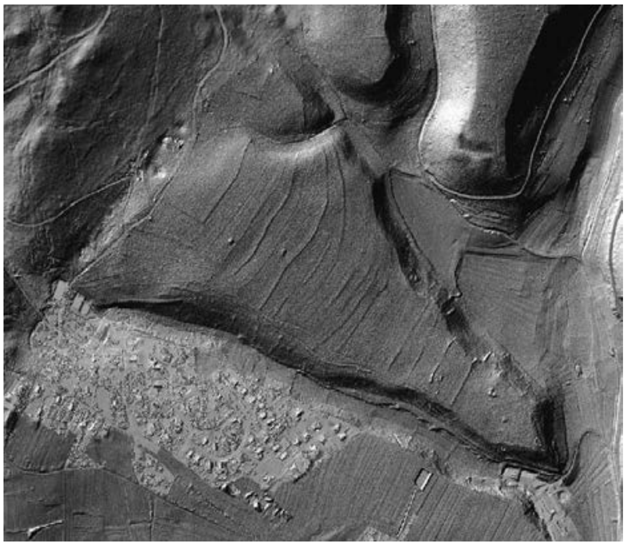


Figure 4 : LIDAR picture of the Kalkhoffen (CG68, infogeo68.fr)

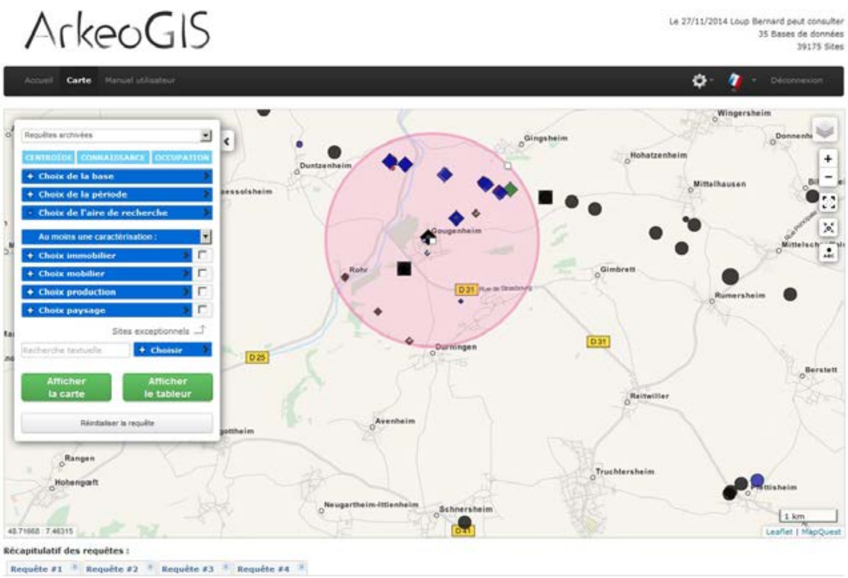


Figure 5 displays how many prehistoric sites (black dots), existing analysis (squares) and other sites around the slot-pit appear inside the applied searching radius.

Counting Sheep Without Falling Asleep: Using GIS to Calculate the Minimum Number of Skeletal Elements (MNE) and Other Archaeozoological Measures at Schöningen 13Ii-4 'Spear Horizon'

Alejandro GARCÍA-MORENO, Jarod M. HUTSON,
Aritza VILLALUENGA, Elaine TURNER, Sabine GAUDZINSKI-WINDHEUSER

Archaeological Research Centre and Museum for Human Behavioural Evolution, RGZM, MONREPOS, Germany
garcia@rgzm.de, hutson@rgzm.de, villaluenga@rgzm.de, turner@rgzm.de, gaudzinski@rgzm.de

Abstract

In this paper, we describe a GIS-based methodology for estimating the minimum number of skeletal elements (MNE) and other archaeozoological measures, such as cut mark distribution and density. As a case study, we present a preliminary application to the Middle Pleistocene site of Schöningen 13II-4, the so-called "Spear Horizon", where a large and exceptionally preserved faunal assemblage imposes difficulties for quantifying skeletal element abundances. We base our methodology on a series of digital templates introduced in a GIS, where each identifiable bone specimen is drawn and the number of overlapping bone fragments is calculated. This methodology yields a direct and accurate calculation of MNE and provides a foundation for assessing other critical archaeozoological measures. Analysis and interpretation of these measures is essential for understanding Palaeolithic subsistence strategies and hominin behaviour.

Keywords*:* GIS, Archaeozoology, Schöningen, Middle Pleistocene.

Introduction

In the last decades, the application of Geographical Information Systems (GIS) and computing has led to great developments in Spatial and Landscape Archaeology, Heritage Management, predictive modelling and Virtual Archaeology. Today, GIS capacities allow for an even wider range of applications in Archaeology (Scianna & Villa, 2011). The combination of database management and image or vector entities representation makes GIS a useful tool for documentation and management of archaeological collections over a variety of scales. The multi-scalar capacities of GIS make possible its application to different types of analysis, from continent-wide to regional, local and intra-site comparisons, and even the study of single artefacts.

In this paper, we present a methodology for analysing individual bones and bone fragments in a quantitative and automatic way. The aim of such analysis is to calculate the minimum number of skeletal elements (MNE), defined as 'the minimum number of skeletal portions necessary to account for the specimens representing that portion' (Lyman, 1994:102). The calculation of MNE is important as it forms the basis of more complicated measures of abundance, such as minimum number of individuals. Together, these measures estimate the number of animals or portions of animals present at a site and provide one piece of evidence to interpret the taphonomic histories of faunal assemblages and associated human behaviours. One downfall is that the calculation of MNE is often difficult and time consuming when dealing with large assemblages. Using GIS, these calculations can be made quicker and more precise through a batch process. In addition to measuring skeletal part abundance, this methodology forms a basis for other analyses, including the location and density of bones surface modifications (cut marks, percussion marks, carnivore damage, etc.). As a case study, we applied this methodology to the Middle Pleistocene faunal accumulation from Schöningen 13II-4. However, the methodology explained here can be applied to any faunal assemblage, regardless of chronology and context.

2. Materials and methods

2.1. Case study: the Middle Pleistocene Schöningen 13II-4 'Spear Horizon'

Schöningen, located in Lower Saxony, Germany, is considered one of the most significant Palaeolithic sites from Central Europe owing to the discovery well-preserved wooden spears in association with a large assemblage of Middle Pleistocene fauna (Fig. 1). Open-cast lignite mining works led to the discovery of numerous archaeological sites within several erosional 'channels' indicating a paleolake environment. The discovery in 1995 of a series of wooden spears, considered the oldest known spears in the world (Thieme, 1997), drove worldwide attention to the so called 'Spear Horizon'. This horizon was originally dated to around 400 kaBP (Richter & Thieme, 2012), although recent U/Th dating resulted in an age of 290±5 kaBP (Sierralta, Frechen & Urban, 2012). This date places level 13II-4 with Marine Isotope Stage 9, which is consistent with biostratigraphic and environmental data (Urban & Sierralta, 2012). Excavations of the 'Spear Horizon', which continued until 2007, extended over an area of ca. 3.900 m^2 and yielded roughly 15.000 archaeological remains (Serangeli *et al.,* 2012), including a large, exceptionally preserved faunal assemblage (van Kolfschoten, 2014; Voormolen, 2008).

As part of a research theme regarding hominin adaptations to interglacial environments, ongoing archaeozoological studies conducted by the MONREPOS Archaeological Research Centre and Museum for Human Behavioural Evolution, show that the Schoningen 13II-4 faunal assemblage is overwhelmingly dominated by horse (*Equus mosbachensis*). Other ungulates, such as aurochs (*Bos primigenius*), bison (*Bison priscus*), red deer (*Cervus elaphus*), giant deer (*Megaloceros giganteus*), roe deer (*Capreolus capreolus*) and rhinoceros (*Stephanorhinus* sp.), are also present in addition to carnivores, small mammals, birds and fishes, suggesting a mosaic interglacial environment around the lakeshore site. The composition and taphonomical analysis of the faunal assemblage suggest an *in-situ*, intensive exploitation of carcasses by hominin groups, probably during several hunting episodes, followed by secondary scavenging by medium-sized carnivores.

While the exceptional preservation of faunal remains allows for extensive, high-resolution archaeozoological analyses, the thousands of identifiable bones impose difficulties for quantification. This is especially true when estimating MNE, where all bone fragments from the same skeletal element must be considered and compared to each other. For that reason, Schöningen 13II-4 is an excellent site to test and develop a GIS-based methodology for estimating archaeozoological measures, such as MNE.

2.2. Methodology: MNE calculation using GIS

The method presented in this paper is based on the previous work by Marean and colleagues (Marean *et al.,* 2001; Abe *et al.,* 2002) who first presented the possibility of using a GIS to estimate MNE using an image analysis approach using digital templates. This method proved to be very useful, but difficult to implement due to limitations in GIS software since it required extensive preparation of bone templates and GIS processing. For that reason, we decided to develop a new method by taking advantage of the newest software and tools in order to simplify the calculations. In a basic sense, this method is based on counting how many times cells with the same value overlap within a series of rasters.

As part of the documentation and archaeozoological analysis of the Schöningen 13II-4 faunal assemblage, every identifiable bone or bone fragment was drawn in a digital template. These templates include multiple views (lateral, medial, cranial, caudal, etc.) of every bone (e.g., right and left femur or mandible) for all species represented in the faunal assemblage. We used simple line drawings of each bone that include all major anatomical features for orientation. Our templates were previously created just for reference purposes, but were perfectly suited to our GIS analysis. One advantage of this new methodology is that any templates can be used; scanned images from an anatomical atlas, hand drawings, photographs, etc., can be adapted to accommodate the needs of any archaeozoological analysis. The level of template preparation is up to the user, but the only requirement is that the templates remain unchanged throughout the entire process. Templates were used as a base outline where bone fragments were drawn in their corresponding location within that skeletal element. Using a digitalizing pad and image processing software (Adobe Photoshop CS4 Extended), bone fragments were directly drawn on the blank templates (Fig. 2), using identifiable anatomical features to locate, orient and scale bone fragments and fit them to templates. The resulting drawings were recorded as a .jpg image file, using the bone's ID number as the file name. That way, an image dataset of bones and bone fragments was created, which not only documents the faunal remains, but also can be integrated directly into a GIS as a raster layer. During the drawing process, special care was taken to represent the bone or bone fragment as accurately as possible. Scaling the background template image with the zoom feature to approximate the size of the bone was helpful to create accurate drawings. The ruler or grid feature included in many graphics programs further aids in drawing bone fragments accurately. Additionally, bone modifications, such as cut marks, percussion notches, carnivore damage, etc., were also drawn. A predefined colour scheme was used in order to ensure that the same colour (defined by Red-Green-Blue values) was used for the same kinds of modifications. Image size and resolution were fixed for

Figure 1: Location of Schöningen, and view of the 13II-4 site.

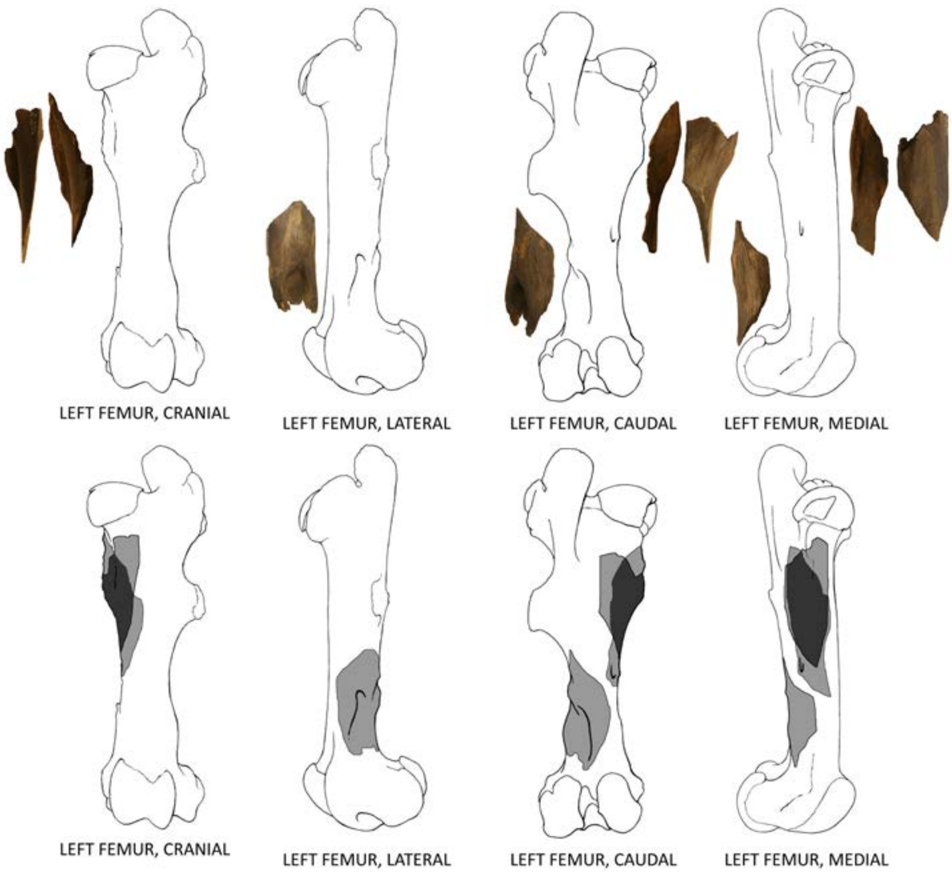


Figure 2: Example of three left femur fragments drawn onto a template. For illustrative purposes, each bone was photographed in several views relative to the template (top). Each bone was drawn onto individual templates and later combined into a single template (bottom). The darker shading represents areas of overlap between the fragments.

each template so that all raster layers included the same extension once introduced in a GIS.

Once bone fragments were drawn, image files were incorporated in a GIS as raster layers. With the fixed colour scheme, every cell corresponding to 'bone' had the same value in all rasters in the dataset, different from any other cells representing blank spaces or bone outlines. Similarly, cells corresponding to bone surface modifications, such as cut marks, had the same value in all rasters. In other words, if a colour value of R: 180, G: 180, B: 180 was used for drawing bone fragments, all cells in the raster dataset corresponding to bone fragment had a value of 180.

As explained previously, MNE is estimated by counting the number of times bone fragments overlap within a faunal assemblage, considering every skeletal element and every species separately. That is, if any portion of two or more bones or bone fragments overlap, they cannot have originated from the same bone. Since all the raster layers corresponding to the same kind of template have the same extension, every cell in the raster shares the same position with another cell in a different raster. Considering that cells representing bone fragments have the same value, overlaps can be calculated easily by counting how many times cells with that value appear in every cell location.

In order to make that calculation, the Equal to Frequency tool from ArcGIS's Spatial Analyst extension was used. This tool 'evaluates on a cell-by-cell basis the number of times the values in a set of rasters are equal to another raster' [1] (Fig. 3). As an input layer, a raster with the same extension as the templates was created, where all cells had the same value as the cells corresponding to 'bone fragment' in the raster dataset.

The resulting raster layer shows how many times the value from the input layer appears in every cell position within the raster dataset. However, since templates have three colour bands (red, green and blue), the Equal to Frequency tool considers every raster three times (the searched value appears three times in every raster, one for every colour band). For that reason, the resulting raster layer must be divided by three (using the Raster Calculation tool). The final raster shows the number of times bone fragments from the template data set overlap (Fig. 4). Each cell records a maximum value that corresponds to the maximum number of times cells corresponding to bone fragments overlap, and therefore to MNE. To avoid negligible overlaps that could overestimate MNE, the raster histogram was checked, and highest values represented in very low frequency (just in a few cells) were ignored (Fig. 5). Ignoring the highest values could turn out in a less precise estimation, although the resulting MNE estimation will be more reliable, since the probability that the estimation is correct is higher when MNE value is lower and its frequency is higher.

[1] ArcGIS Help Library, 2010, access 01.07.2014.

2.3. Further developments

The method presented here can form the basis for calculating other archaeozoological measures. The use of a predefined colour scheme for drawing every bone modification is used to evaluate the location and frequency of each modification type within the raster dataset. However, with this method, templates' cells have a different value in every colour band, which impedes the use of the Equal to Frequency tool used to estimate the MNE. Due to this limitation, additional processing of drawings is required.

One solution is to reclassify, using a batch process, templates' cells values. In this case, values of cells corresponding to modifications, such as cut marks, are changed to '1', while the rest of the cells are given a value of '0'. Once all drawings have been reclassified, raster layers can be summed, using ArcGIS's Raster Calculator Addition function, which sums on a cell-by-cell basis the values of two or more rasters. The resulting raster layer shows the location of bone modifications for each bone, including any locations where modifications overlap (Fig. 6a). Once locations of bone modifications are known, further analyses can be undertaken, such as evaluating the spatial distribution of cut marks across each skeletal element. In this sense, a density analysis will show the areas where higher concentrations of modifications appear (Fig. 6b), allowing for more accurate and meaningful taphonomical analyses.

Despite the need of previous processing, this method makes possible deeper archaeozoological analysis using a single drawing for every bone fragment, instead of using different templates for every modification. Analysing the preferential locations of bone modifications allows for a better understanding of human behaviour, such as butchering techniques and exploitation of carcasses in the case of cut marks or percussion notches.

Discussion and conclusions

Estimating MNE and other zooarchaeological measures, such as cut mark frequency and location, is a rather complex and time consuming process. This is especially difficult when dealing with very large faunal assemblages where hundreds, or even thousands, of bone fragments must be compared. This is the case for the large and well-preserved faunal assemblage from the Middle Pleistocene site of Schöningen 13II-4.

In order to improve MNE estimation, a GIS-based method was developed and tested with the Schöningen 13II-4 faunal assemblage. This method was based on the use of

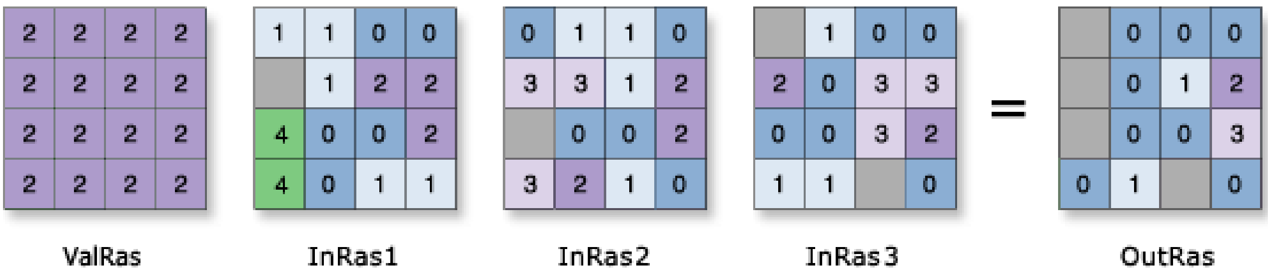


Figure 3: Schema of the Equal to Frequency tool calculation process (Image: ArcGIS Help 10.1).

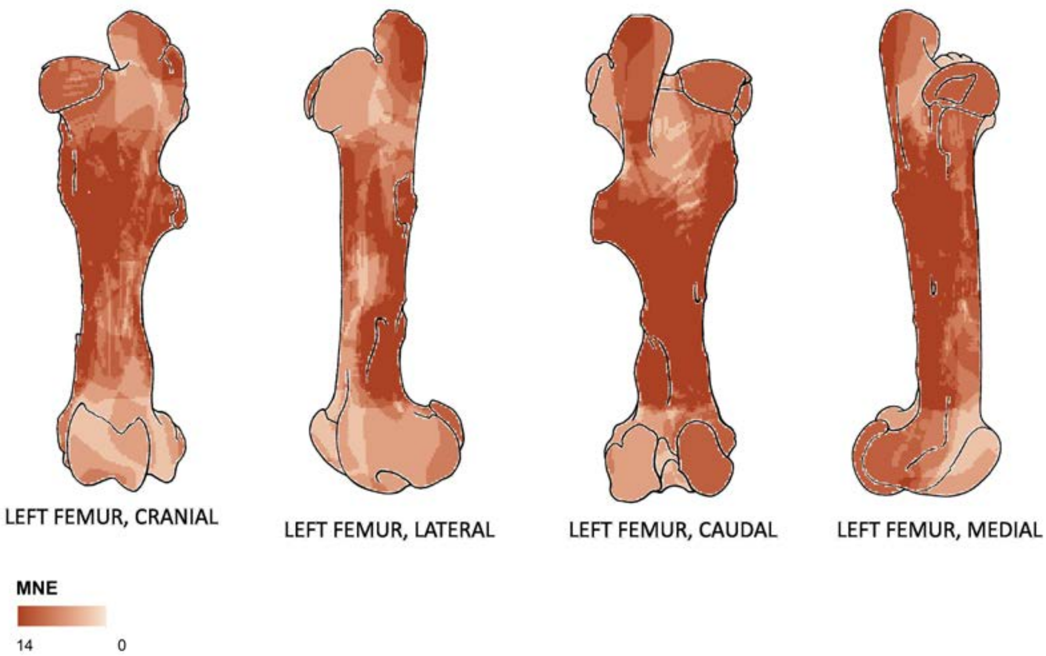


Figure 4: Raster layer showing the final number of bone fragments overlapping for Schöningen 13II-4 horse left femora.

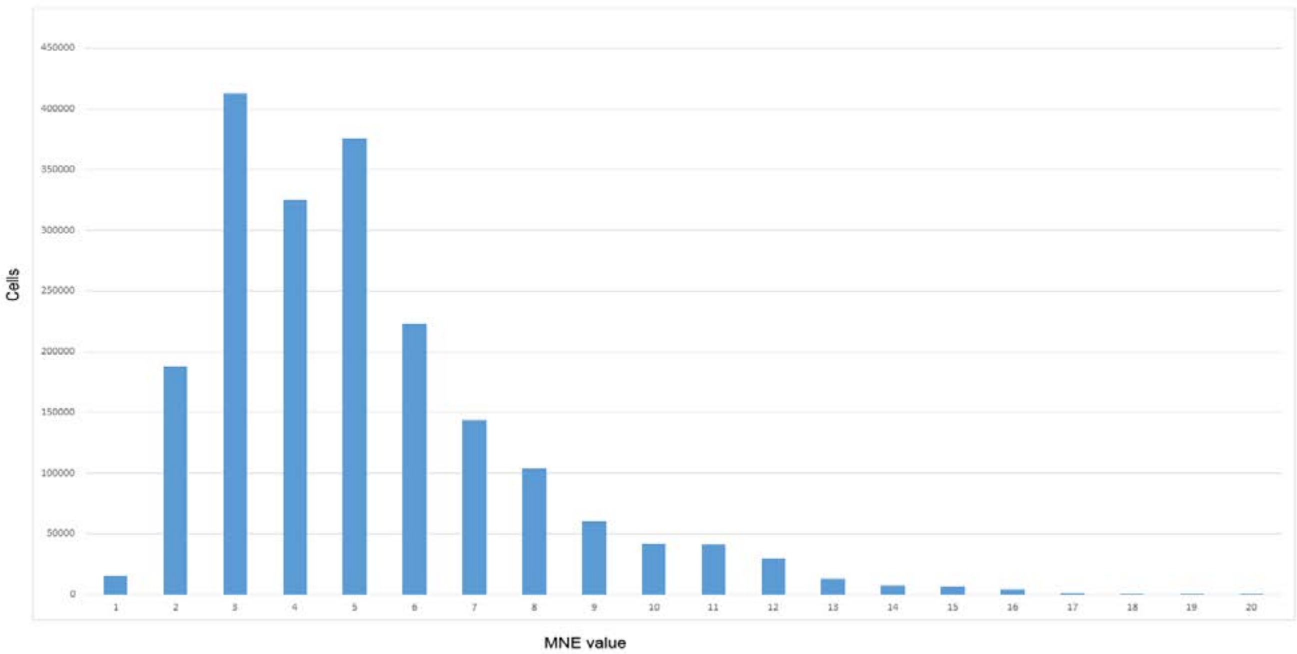


Figure 5: Raster's histogram, showing cells' value frequency. Negligible frequencies in higher values can be ignored, in order to obtain a more accurate MNE estimation.

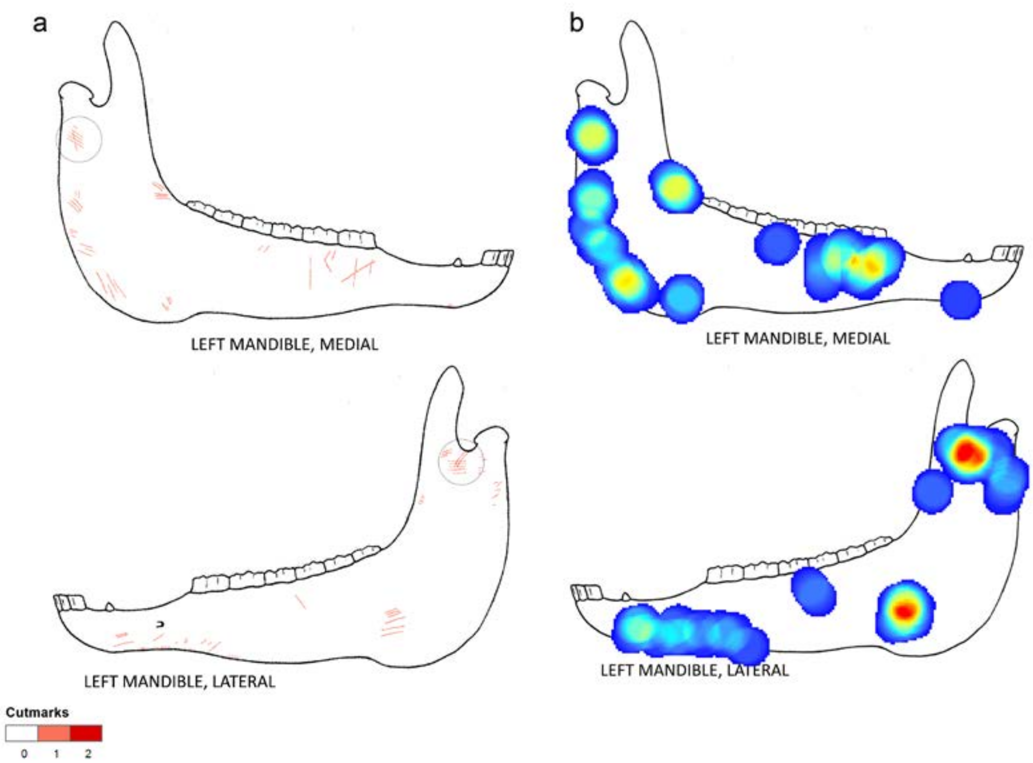


Figure 6: Raster's histogram, showing cells' value frequency. Negligible frequencies in higher values can be ignored, in order to obtain a more accurate MNE estimation.

templates, where bone fragments were drawn as part of the documentation process and the archaeological analysis. Those drawings were then introduced as raster layers in to a GIS, and the number of times cells corresponding to bone fragments overlap was calculated. Templates can also be used to evaluate bone modifications, analysing the distribution, concentration and density of alterations such as cut marks.

The application of GIS allows for the calculation of different archaeozoological measures in an easy, batch-like way using any digital template. The method presented here is not entirely automatic since it still requires a few calculations to obtain the final layer showing bone fragment overlaps. In addition, the evaluation of other archaeozoological measures, such as cut marks, needs previous processing of the templates. Despite some additional processing, the method presented here represents an improvement of traditional (manual), extremely time consuming methods and allows for more complex archaeozoological analyses of human subsistence strategies and behaviour. Further development of the methodology presented here could improve the level of automation for this procedure, for example creating a workflow model stringing together the sequence of steps and calculations required. The use of free software with similar capacities and tools as the ones used here would allow the application of this method to any archaeozoological analysis, regardless of the availability of funds. Advancing the method tested on the Schöningen 13II-4 assemblage will improve archaeozoological analyses and will allow for a better understanding of past human behaviour.

Acknowledgements

We would like to thank Curtis W. Marean (Arizona State University) for providing us the details of his previous work, which is the basis for the methodology provided in this paper. The project *Subsistenzstrategien mittelpleistozäner Homininen in Schöningen* is funded by the *Deutsche Forschungsgemeinschaft* (DFG; GA6839/-1). We would like to thank the *Niedersächsisches Landesamt für Denkmalpflege* for their support of the project. AGM attended the CAA2014 Conference thanks to a *Deutscher Akademischer Austauschdients*.

References

ABE, Y., MAREAN, C.W., NILSSEN, P.J., ASSEFA, Z., STONE, E.C. (2002). 'The analysis of cutmarks on archaeofauna: a review and critique of quantification procedures, and a new image-analysis GIS approach', *American Antiquity* 67, pp. 643-663.

LYMAN, R.L. (1994). *Vertebrate Taphonomy*. Cambridge : Cambridge University Press.

MAREAN, C.W., ABE, Y, NILSSEN, P.J, STONE, EC (2001). 'Estimating the minimum number of skeletal elements (MNE) in zooarchaeology: a review and a new image-analysis GIS approach', *American Antiquity* 66, pp. 333-348.

RICHTER, D., THIEME, H. (2012), 'One first chronometric date for the Lower Palaeolithic occupation at Schöningen 13I' In: K.E. Behre (ed.). *The chronological setting of the Palaeolithic sites of Schöningen*, Mainz: Römisch-Germanisches Zentralmuseum Mainz, pp. 171-182.

SCIANNA, A., VILLA, B. (2011). 'GIS applications in archaeology', *Archeologia e Calcolatori* 22, pp. 337-363.

SERANGELI, J., BÖHNER, U., HASSMANN, H., CONARD, N. (2012). 'Die pleistozänen Fundstellen in Schöningen – eine Einführung', In: K.E. Behre (ed.). *The chronological setting of the Palaeolithic sites of Schöningen*, Mainz: Römisch-Germanisches Zentralmuseum Mainz, pp. 1-22.

SIERRALTA, M., FRECHEN, M., URBAN, B. (2012). '230Th/U dating results from opencast mine Schöningen', in *The chronological setting of the Palaeolithic sites of Schöningen*, ed s K.E. BEHRE, Römisch-Germanisches Zentralmuseum Mainz, Mainz, pp. 143-154.

THIEME, H. (1997). 'Lower Palaeolithic hunting spears from Germany', *Nature* 385, pp. 807-810.

URBAN, B., SIERRALTA, M. (2012). 'New palynological evidence and correlation of Early Palaeolithic sites Schöningen 12 B and 13 II, Schöningen open lignite mine', in *The chronological setting of the Palaeolithic sites of Schöningen*, ed s K.E. BEHRE, Römisch-Germanisches Zentralmuseum Mainz, Mainz, pp. 77-96.

VAN KOLFSCHOTEN, T. (2014). 'The Palaeolithic locality Schöningen (Germany): A review of the mammalian record', *Quaternary International* 326-327, pp. 469-480.

VOORMOLEN, B. (2008). *Ancient hunters, modern butchers. Schöningen 13II – 4, a kill – butchery site dating from the northwest European Lower Palaeolithic*. Universiteit van Leiden, Leiden.

Looking for the Best. A Comparison Between GIS and PageRank Based Algorithms for Preventive Archaeology in Urban Areas

Dubbini NEVIO,[1] Gabriele GATTIGLIA,[1] Augusto PALOMBINI[2]

[1]University of Pisa – MAPPA Lab, Italy
[2]CNR (National Research Council of Italy), ITABC - Institute for Technologies applied to Cultural Heritage, Italy
g.gattiglia@arch.unipi.it

Abstract
In this paper we present a comparison between the MAPPA algorithm and some more standard predictive models implementable within the GIS: kernel density estimation (KDE), kriging/co-kriging and r.finder. The MAPPA algorithm is a mathematical model, based on PageRank, that imitates the modus operandi of archaeological practice, reproducing the relations between the different natural and anthropogenic elements. Kde is a non-parametric way to estimate the density of a random variable, in which a function creates a continuous approximated distribution. Kriging is an interpolation method, based on spatial autocorrelation. Co-kriging allows to add further spatial information in order to improve the interpolation. r.finder is a GRASS-GIS script created to check the analogies among the places where some known items are located, in terms of similar cell-values in the frame of a series of thematic raster maps. The obtained results confirm the better suitability of predictive models specifically developed by and for archaeologists.

Keywords: Predictive Archaeology, Map Algebra, Mathematical Models, Geostatistics

Introduction: Predictive modelling and archaeological potential

In the general framework of predictive archaeology it is useful to focus on the features of different tools and their suitability for specific contexts. Predictive modelling takes its starting point in the 40's, while the use of predictive modeling in archaeology can be connected with the rise of the New Archeology in the late 60's (Verhagen & Whitley, 2011). By the 70's, more rigorous approaches have been tried (Gumerman, 1971). By the 80's, two primary lines of models were developed: models finalised to identify spatial suitability and models targeted to correlative statistical summaries that could be applied in unsurveyed areas (Verhagen & Whitley, 2011). The approaches used so far were mostly based on statistical modelling techniques, with a number of different methods based on regression, correlation, Bayesian statistics, Kriging/coKriging models. Mathematical models instead were tried as an alternative, since they allow to add modelling working principles besides statistical information, and include techniques like map algebra, trend surface analysis, cost distance models, Dempster-Schafer theory, agent-based models (see Drennan, 2010; Hodder & Orton, 1976; Kamermans, van Leusen & Verhagen, 2009; Wheatley & Gillings, 2002, and references therein). The idea to include parameters in predictive archaeological models concerns the basic linear form of most predictive models, this has been done in (Dubbini & Gattiglia, 2013), where the PageRank algorithm (Langville & Meyer, 2006) was used to predict the archaeological potential. Within predictive modelling, accuracy measures how well sites are captured, and precision how narrowly areas of high probability are limited (Kamermans, van Leusen & Verhagen, 2009). Accuracy and precision are often use to validate or test the predictions with some new available data. The gain statistic (Judge & Sebastian, 1988) is the mostly used to measure and combine these attributes. Other measures are described in (Lieskovsky, Ďuračiováa, & Karell, 2013).

In the last 20 years, predictive modeling has been used mainly as a decision-making tool in cultural resources management and in land management, and less for the definition of site location or the interpretation of the spatial patterning of archaeological sites. The use of predictive models in land management has produced both enthusiasm and criticism. Conversely, the recent practice of preventive archaeology shows that the use of predictive models in the early stages of land management planning is very successful for the protection of the archaeological heritage (Verhagen & Whitley, 2011). Hitherto, predictive models were primarily applied for the creation of maps of archaeological potential on a landscape scale, rather than on urban scale, even if the risk of pauperising the urban archaeological deposits is actually very high. With archaeological potential of an urban area, we means the probability that a more or less significant archaeological stratification is preserved. The archaeological potential of an urban area is the determined by the following overall parameters: type of settlement, i.e. the presence of settlement structures and their different typology; density of settlement; multi-layering of deposits; removable or non-removable nature of the archaeological deposit; degree of preservation of the deposit, calculated according to the presence of anthropic and natural removals. Finally, the archaeological potential is independent of any other

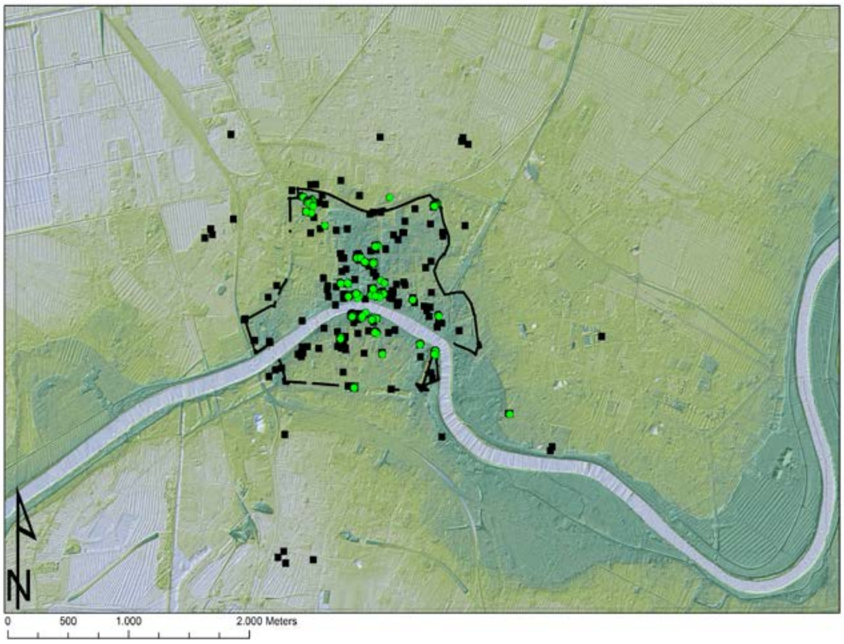


Figure 1: The spatial distribution of the vector data of the late medieval archaeological findings with Certain geolocation (circle) and with Uncertain geolocation (square). The black line highlights the path of the medieval urban walls that enclose the city centre.

Figure 2: The spatial distribution of the vector data of Shapes of the late medieval archaeological findings.

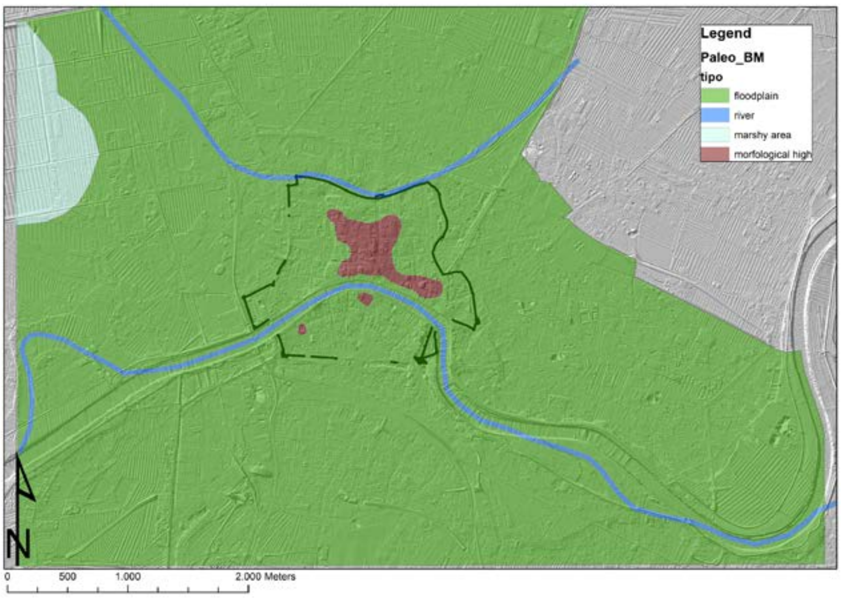


Figure 3: The vector data (polygons) of late medieval geomorphology with the different geomorphological facies deduced from geological surveys (river, floodplain, marshy area, geomorphological high).

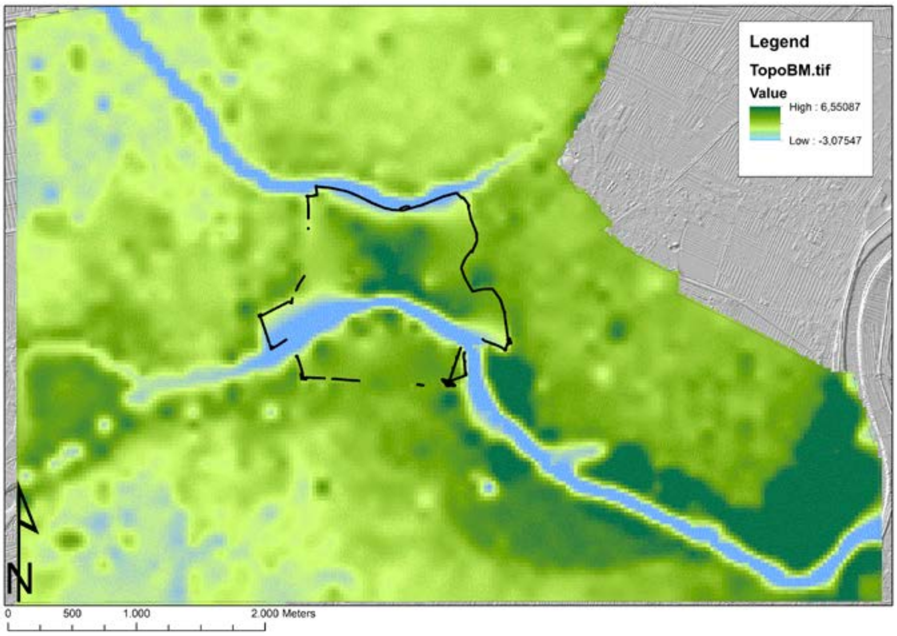


Figure 4: The late medieval DEM created with a double realibility scale (high) for urban (area enclosed by the city walls) and (low) for extraurban areas. The extra-urban DEM was developed in a regressive manner as descripted in (Gattiglia, 2014, p. 49).

following intervention that is carried out, which must be regarded as a contingent risk factor (Dubbini & Gattiglia, 2013).

In this paper the authors present a comparison test between kernel density estimation, kriging/co-kriging, r.finder and the MAPPA algorithm finalised to the definition of the most useful model for preventive archaeology in an urban area. As a matter of fact, it is meaningless and misleading to address good or bad models, instead, looking for the best result means the ability to check the most suitable tool to give meaning to available data in specific contexts, i.e. to estimate the archaeological potential depending of the datasets. Moreover, '*It is important to understand here that the method's conceptual base as well as its units of measure are provided by theory. Stripped of its theory a method provides measurements, but no meaning. Effort that stress the development of method as a route toward theory building confuse measurements with understanding*' (Moore and Keene 1983, p. XIV).

The datasets

To compare the results of the different methods applied for the estimation of archaeological potential, we used the Late Medieval data of the urban area of Pisa (Dubbini & Gattiglia, 2013), composed of these datasets:

- vector data (points derived from polygons) of archaeological findings with Certain geolocation (1205 features);
- vector data (points derived from polygons) of archaeological findings with Uncertain geolocation, i.e. data for which we only knew that they are located in a certain region of the space (344 features) (Fig. 1);
- vector data (points derived from polygons) of Shapes of archaeological findings: i.e. from certain geolocation data, the archaeological team has tried to give a shape to the find, e.g. to outline a house from a wall or a floor, or the continuation of a street from a piece of it, and so on (66872 features) (Fig. 02);
- vector data (polygons) of geomorphology: these data were deduced from geological surveys, in order to identify, for each archaeological period, the diverse geomorphological facies, which were distinguished in river, floodplain, wetland, marshy area, geomorphological high (10 features) (Fig. 03);
- raster data of the Late Medieval DEM (20 m cells) derived by elevation points available from archaeological excavations (Gattiglia, 2014, p. 49) (Fig. 4)

All data are associated with a value of archaeological potential. To compute the value of this parameter a list of 19 areas of interest has been drawn up, corresponding to the main informative fields on which the finds can provide information. The identified areas of interest are: production, building techniques, trade, food, agriculture/breeding, worship, waste management, political/institutional aspects, social and gender aspects, physical anthropology, fauna/flora, geomorphology, viability/transport, health and hygiene, warfare, land management, leisure, tradition, water system. Finds were so assigned a value of absolute potential by summing, for each area of interest, the value 1 if the category provided information on that informative field, and 0 otherwise. Also the geomorphological datum was given an absolute potential value, by summing the absolute potential of all the categories of finds that can be present in a facies (Anichini *et al.*, 2013; Dubbini & Gattiglia, 2013).

The archaeological data relating to the late medieval period are 519 overall, amounting to 24% of all archaeological data collected, so they represent a good sample for our test.

Topographically, they are located in the northern-centre section of the area under study and are highly concentrated in the current city centre, where they are spread out quite evenly, with the exception of the western and south-eastern sector. With regard to the type of finds, the categories more greatly attested are *Areas for private use* (23%), attributed almost entirely to houses, for the majority tower houses, and to contexts which cannot be better interpreted, such as those relating to generic *Frequentations* and to *Indefinite structures* which together amount to 21% of the total finds. *Infrastructures*, especially roads, squares and alleys, refer to 16% of finds; *Areas for public use*, almost entirely places of worship, to 10%; *Production areas*, represented for the most part by metal manufacturing, to 9%; *Natural contexts*, mainly marsh and overflow areas, to 7%; areas defined as *Non-places*, corresponding to contexts related to cancelling, abandonment and spoliation, to 5%. Smaller percentages of finds (4%) regard *Areas with military function*, mainly referred to the urban defensive system; *Agricultural/ vegetable gardening areas*, corresponding mainly to vegetable gardens, and *Funerary areas* amounting to 2%. The late medieval geomorphology was reconstructed in great detail. The more elevated area developed in correspondence with the stretch where the two watercourses (Arno and Auser river) were the closest to one another. To the east and west, respectively, of this morphological high, two lows, represent the morphological trace of two former marshy areas. The DEM has been interpolated processing the elevation points available from archaeological investigations with the ANUDEM (Australian National University Digital Elevation Model) algorithm created for the development of hydrogeologically correct DEM (Hutchinson, 1989), selected following quantitative comparisons with other algorithms of the algorithms general purpose type (for example, Spline, IDW), in order to obtain the greatest elevation accuracy. The algorithms were compared using reiterative cross-validation techniques, which consist of creating a new DEM and omitting certain points of the original dataset and subsequently testing the elevation of these points with their original one. The result provides a representation of a city that is clearly divided by the Arno River, with elevations included between current -1 m a.s.l. and + 5.6 m a.s.l. (Gattiglia, 2014, p. 107).

The Methods

Due to the spatial distribution of our datasets, we decide to restrict the analysis area at the area enclosed by the medieval city walls and at the closest nearby. The datasets were analysed with four different kinds of spatial analysis methods. The first two, i.e. kernel density estimation and kriging/co-kriging, are available in the majority of GIS software; the third, i.e. r.finder, consists in a GRASS-GIS script created for predictive archaeology; and the fourth method is based on PageRank algorithm, and was specifically developed during the MAPPA project.

Kernel density estimation (KDE) is a non-parametric statistical tool used to estimate the density distribution of a random variable and belongs to the point pattern analysis family (Hodder & Orton, 1976, p. 30). It is a model in which a two-dimensional probability density function acts among observed values to create a smooth approximation of the distribution of the available data from the centre outwards, weighing the events depending on their distance from the point where their intensity is estimated. The probability density function is called the kernel as it acts as a kernel in a mathematical scalar product, giving also the name to the technique. When processing the data, various parameters

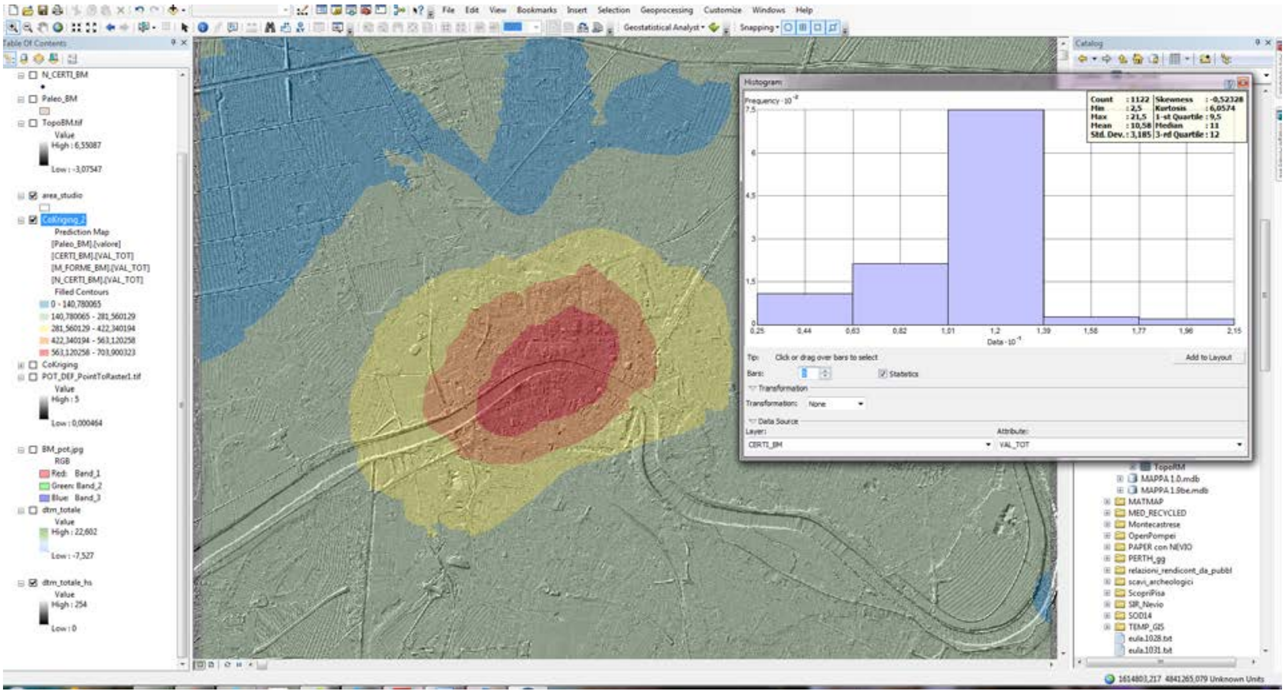


Figure 5: The data analysis carried out before the application of co-kriging interpolation, that depict how the data are distributed according to a Gaussian curve. This is, together with numerical and spatial relevance of the data, one of the main prerequisites for applying correctly the kriging interpolation.

or features of the kernel function can be chosen, like for instance its shape and bandwidth (or radius). The function is directly influenced by the radius within which the density of the point is calculated: the greater it is, the closer the result is, until coinciding with a continuous surface; the smaller it is, the more the result will capture single starting events only. The density value of each cell is calculated by summing the distribution density values that overlap in that cell, producing a continuous surface that can be more easily interpreted than that obtained with simple density and in which clusters are more evident (Beardah & Baxter, 1996).

Since it is not possible to weigh the single data, for the aim of the analysis conducted, we merged the three vector data datasets of archaeological findings (Certain geolocation, Uncertain geolocation, Shapes) applying a circular-shaped kernel, with a radius of 150 m, a distance within which it is logical to suppose continuity of the urban settlement, and an output with 20 m resolution cells.

Kriging is a geostatistical weighted interpolation based on spatial autocorrelation of the data. Kriging interpolation is often applied when both the distribution and density of the points are irregular, since it allows to interpolate data with a reasonable amount of flexibility. It describes spatial variation as the product of a deterministic and a stochastic component, according to the regionalized variable theory (Matheron, 1971). The deterministic component is a trend, whilst the stochastic component is composed of two parts: a random component correlated to a global pattern and a highly-localized casual random noise caused by a measurement error or by small-scale processes (Lloyd & Atkinson, 2004). The principle of spatial autocorrelation (Hodder & Orton, 1976, p. 174) is based on the fact that the values observed are related to each other. This means that values tend to be similar when they are spatially close to each other and vice versa. The correlation between the variable's values tends to decrease as distance increases. The weight given to the values depends on the spatial structure and on the degree of spatial autocorrelation of the distribution. For this reason, a semivariogram is used (Lloyd & Atkinson 2004) to evidence qualitatively and quantitatively the spatial dependence degree (the autocorrelation), and to interpolate the variance of the values observed in groups of pairs points (the lag) at set distances. The semivariogram has 3 main parameters: range, i.e. the maximum distance within which autocorrelation is displayed; sill, i.e. the maximum value reached by the semivariance; nugget, i.e. non-explicable part of the semivariance, which is attributable to measurement errors, instrumental errors and spatial variability at distances lower than the minimum sampling distance. In order to obtain a continuous function, the semivariogram must be associated with a mathematical model necessary to describe the general trend of spatial variation. The most frequent models used are circular, spherical and exponential. The first two show a decrease in autocorrelation, until becoming 0 at a specific distance (range). Exponential models are applied when spatial autocorrelation decreases exponentially with increasing distance and then disappears at an infinite distance. The kriging is also able to define the main orientations of diffusion of the values – in this case we speak of anisotropic kriging – or not to define them – in this case we speak of isotropic kriging. The kriging requires that the values to be studied must be in some way continuous, that they must be numerically and topographically relevant, that they are distributed according to a Gaussian curve and that there is no longer correlation from a certain spatial distance onwards, or that the semivariogram curve flattens out beyond a certain threshold, i.e. when it reaches the sill.

For the aim of the analysis conducted, we used co-kriging (stratified kriging) combining all the datasets but the DEM with an Isotropic, Spherical, Ordinary kriging method (Fig. 5).

Here the main parameters for each dataset:

Geomorphological data
TrendType: None; NeighbourSearch: Standard; NeighboursMax: 5; NeighboursMin:2; sectorType: 8;
Certain location data
TrendType: None; NeighbourSearch: Standard; NeighboursMax: 5; NeighboursMin:2; sectorType: 4
Shape
TrendType:None; NeighbourSearch:Standard; NeighboursMax:5; NeighboursMin:2; sectorType: 4
UnCertain location data
TrendType: None; NeighbourSearch Standard; NeighboursMax: 5; Neighbours Min:2; sectorType: 4
Variogram
NumberOfLags 12
LagSize1906.3064255630882

R.finder is a GRASS-GIS (Neteler & ,Mitasova 2004) script created for predictive archeology, but mainly conceived for landscape analysis (Palombini, 2013). The main function of the script is to identify the areas more similar to the ones in which some archaeological findings are located, on the basis of the values of a series of raster thematic maps. r.finder considers the values of the cells where input elements are located and outs a map representing the degree of similarity of any part of the region to those.

R.finder distinguishes two kinds of maps: qualitative and quantitative. For qualitative maps the analysis is performed checking for any single category value of the thematic maps in which the occurences fall. For quantitative maps, the analysis is performed checking for the whole range between the lowest and the highest value (or the standard deviation interval).

The final result is a map resulting for the sum of the single analysis, showing the higher or lesser concentration of analogies with the cells on which elements are.

For the experiment, the typological class of tower-houses was taken into account, analyzing – for the definition of the location – one qualitative map (geological units) and three quantitative maps: DEM, slope, aspect, and distance buffers from water streams and city wall belt.

The mathematical model used in this paper for the comparison was developed during the MAPPA project, and is a mathematical model conceived for the predictive calculation of the archaeological potential of an urban area and it is based on a modification to the PageRank algorithm.

PageRank is a well-known algorithm developed and used by search engines for attributing a value of importance to web pages regardless of the value of their content and solely on the basis of the interconnections between the pages. The criterion used to compute the ranking of the web pages may be summarised as follows: a page i that points to pages, distributes its importance in equal parts to pages , and therefore gives 1/k of its importance to the pages it points to. The rationale in applying a PageRank based model to estimate the archaeological potential is that of imitating the modus operandi of archaeological practice, in order to reproduce the relations (links) between the different natural and anthropogenic elements.

(Langville & Meyer, 2006) contains an explanation of the PageRank model, while the application of PageRank based techniques to the estimation of the archaeological potential can be found in (Dubbini & Gattiglia, 2013). Considering the present dataset, the model works based on a vector D representing available data, and a matrix S representing the weights (value) of the links between cells. The matrix *S* is computed in the following way:

- each cell with a data distributes its importance to a square mask of cells centered in the cell itself. The edge of this square is related to the functional area the cell is in, by means of the so called functional areas (Dubbini & Gattiglia, 2013)
- when a mask as in the previous step is constructed, the total weight (i.e. the sum of weights distributed by a cell) inside the mask is given by a value influenced by the probability of finding high or low valued finds in the nearby (Dubbini & Gattiglia, 2013);
- the distribution of weights in the mask around each cell with data, is given by the uniform distribution weighted by the geomorphological values of the mask around the cell, and weighted by the functional areas values (Dubbini & Gattiglia, 2013).

The implementation of the PageRank algorithm consists of a basic procedure, applied repeatedly:

- The vector D representing available data and the matrix of weights S are generated as described above;
- The following iterations are performed

$$A = S \cdot x + \begin{bmatrix} 1/n & \cdots & 1/n \\ \vdots & \ddots & \vdots \\ 1/n & \cdots & 1/n \end{bmatrix} \cdot x;$$

$$A = (1\text{-yield}) \cdot A + \text{yield} \cdot x;$$

$$u = [1\ 1\ \ldots\ 1];$$

$$y = \text{rel} \cdot (A) + (1\text{-rel}) \cdot [D \cdot (uT \cdot x)];$$

$$y = \frac{y}{\sum_{i=1}^{n} y_i}$$

$$x = y;$$

end

$$D = \text{speed\_up} \cdot x + (1\text{-speed\_up}) \cdot D;$$

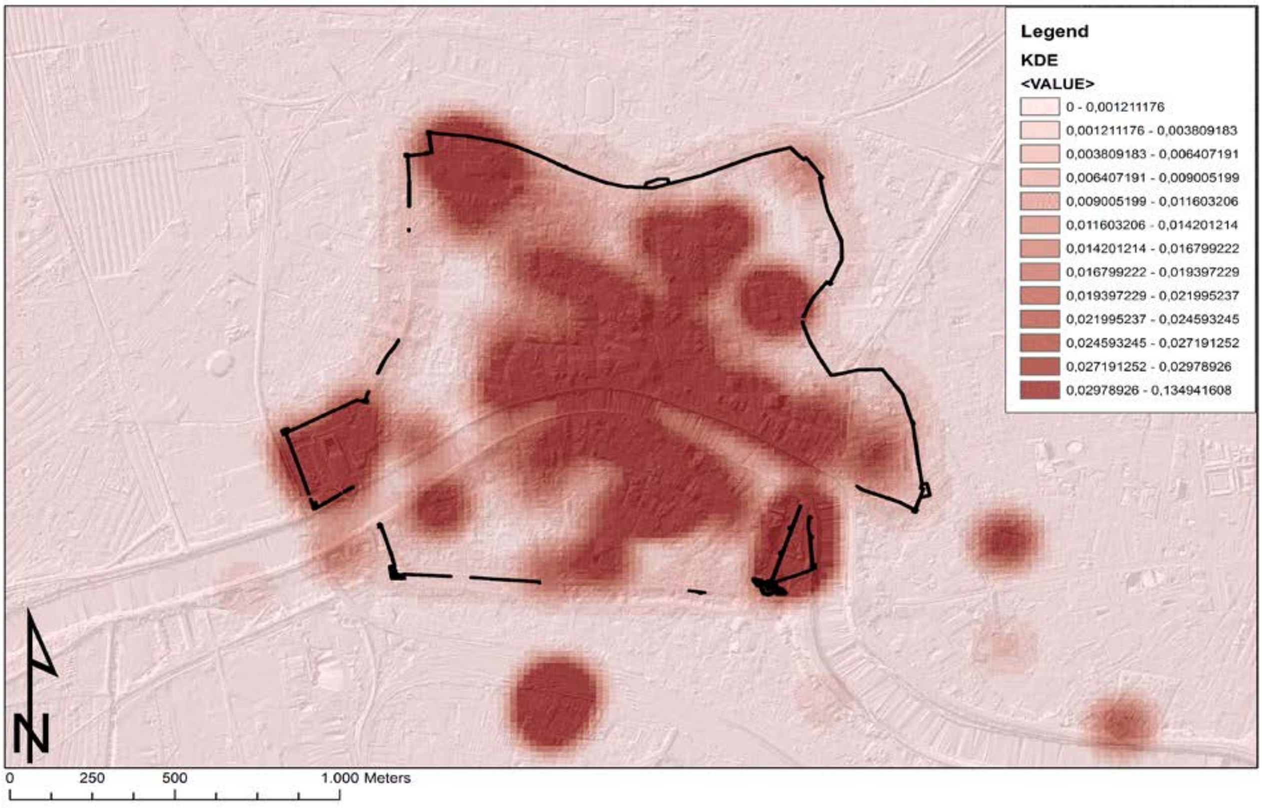


Figure 6: The KDE predictive test visualised as a ¼ standard deviation. The result obtained is an approximate idea of the data spatial structure, i.e. of the concentration of late medieval archaeological findings, but, practically, its non-parametric function depict a sort of self-fulfilling prophecy.

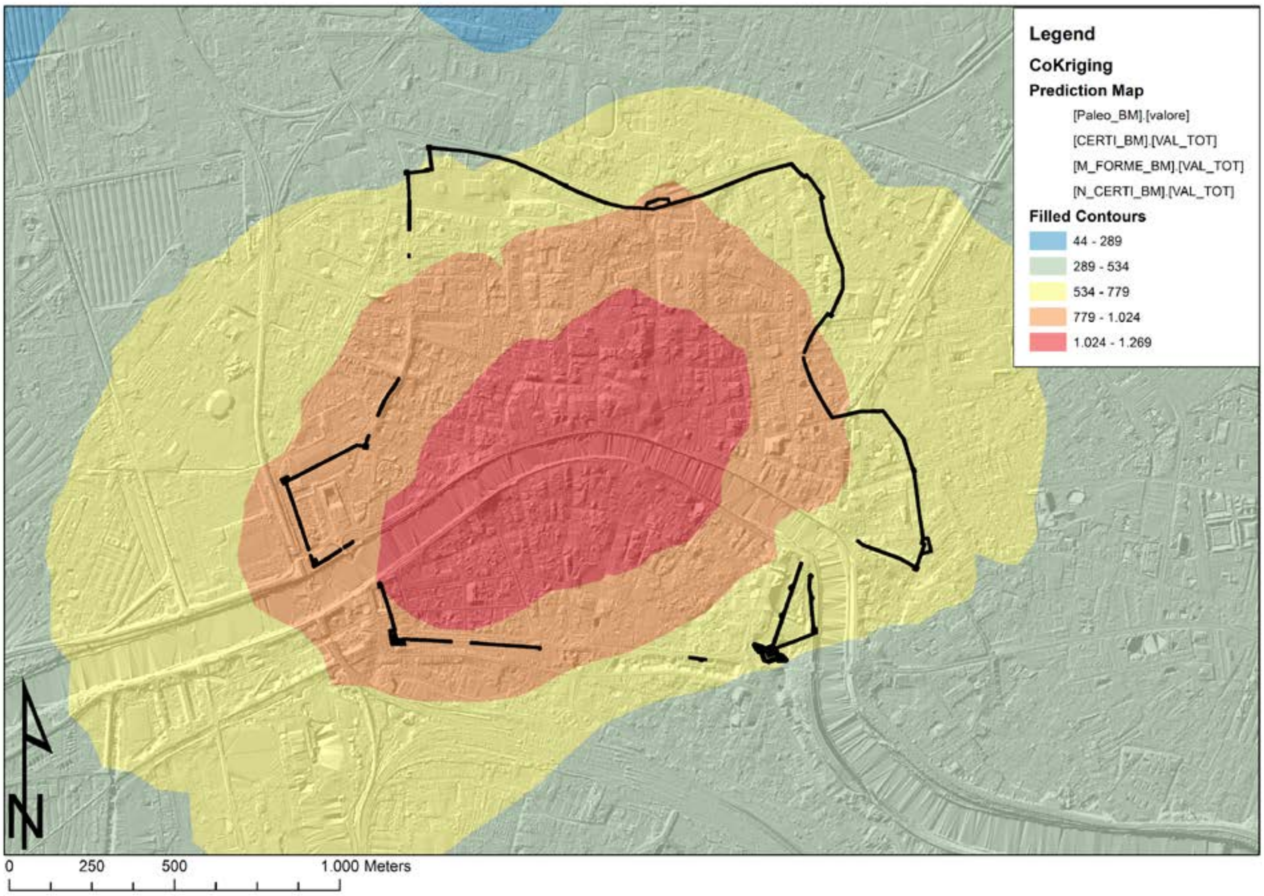


Figure 7: The predictive map of archaeological potential realised with the co-kriging interpolation. The results highlights how the method spreads the values as a continuum from the centre of the city were the archaeological findings with highest potential values are conserved. The prediction is unable to appreciate sharp interruptions such as the presence of city walls, or of the river.

- In these formulas *x* is a stochastic (i.e. the sum of its component equals 1) random column vector of dimension n, used as an initial condition for the application of the iteration described in the ‘for’ cycle (the result of these iterations are independent of the initial condition). The following are the tunable parameters used in the algorithm:
- *maxit* ∊ *N* is the number of times the algorithmic procedure is executed. Each time the algorithm makes a step in the prediction of archaeological potential, and after each step the result is taken as the new starting point for the next step. So, the greater maxit, the more the prediction ‘turn from’ the original data;
- *speed’_’ up* ∊ *[0,1]* is the weight expressing the part of the new absolute potential due to the results of the application of steps 1. and 2., and the part due to the absolute potential of the previous step. So, the more speed_up approaches 1, the less the new computation is due to the data from the previous step;
- *rel* ∊ *[0,1]* is the parameter ruling how much we take in consideration the potential given by the weight matrix S (relations), with respect to how much we take in consideration the potential given by the values of absolute potential. So, the more rel is near 1, the less the absolute value of the potential are taken in consideration, and the more the matrix of weights S (i.e. the relations) is preeminent in determining the archaeological potential;
- *yield* ∊ *[0,1]* is the amount of potential each cell keeps for itself, with respect to the rest, which is distributed on the basis of the weight matrix S. So the more yield is near 1, the more each cell keeps potential for itself.

For the aim of the analysis conducted the actual values of the parameters used were *maxit = 3, speed_up = 0.7, rel = 0.5, yield = 0.5.*

Results

In the analysis with the kernel density estimation, the greater the number of archaeological finds falling within the radius, the greater the density will be in that area, so the result obtained is an approximate idea of the data’s spatial structure, which in our case means the varying of the presence of finds and the greater or lesser concentration. Their representativeness is greater in areas where many archaeological interventions have been carried out, also in the absence of data regarding that chronological period (Fig. 6).

The final outcome of the co-kriging interpolation is an archaeological potential map for late medieval period

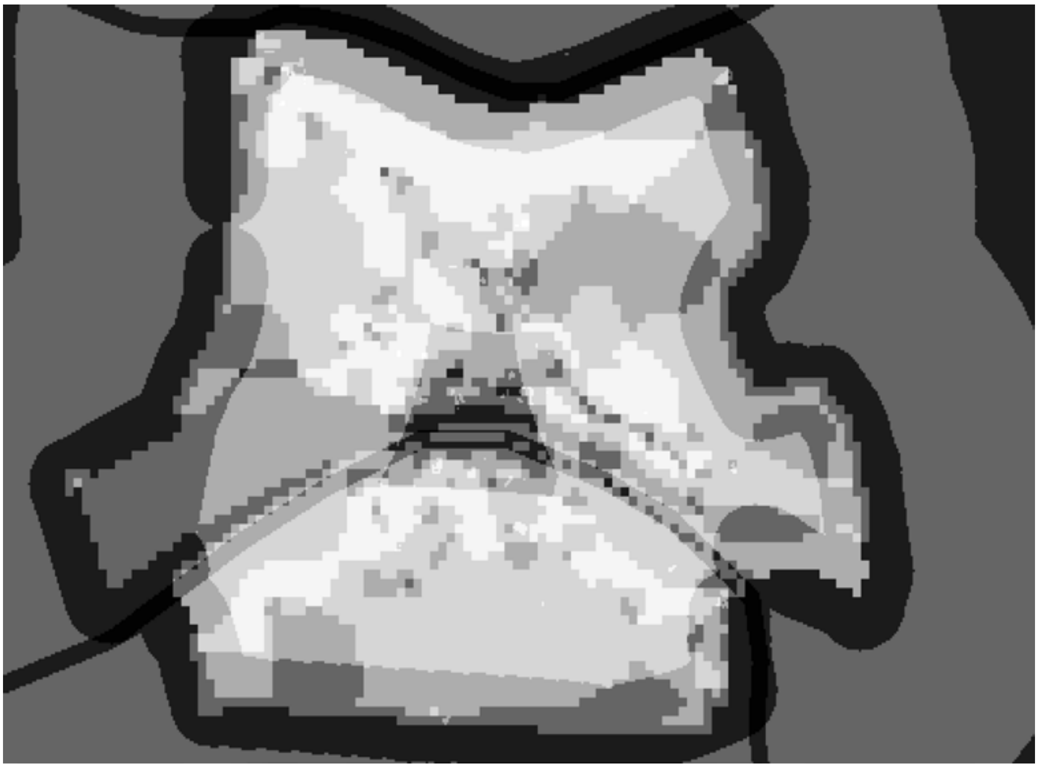


Figure 8: r.finder predictive test for tower-houses distribution (considering thematic maps of geological units (qualitative), DEM, slope, aspect, and distance buffers from water streams and city wall belt (quantitative). The result indicates areas half-way between the city wall belt and the river as the most suitable for tower-houses placement.archaeological findings with highest potential values are conserved. The prediction is unable to appreciate sharp interruptions such as the presence of city walls, or of the river.

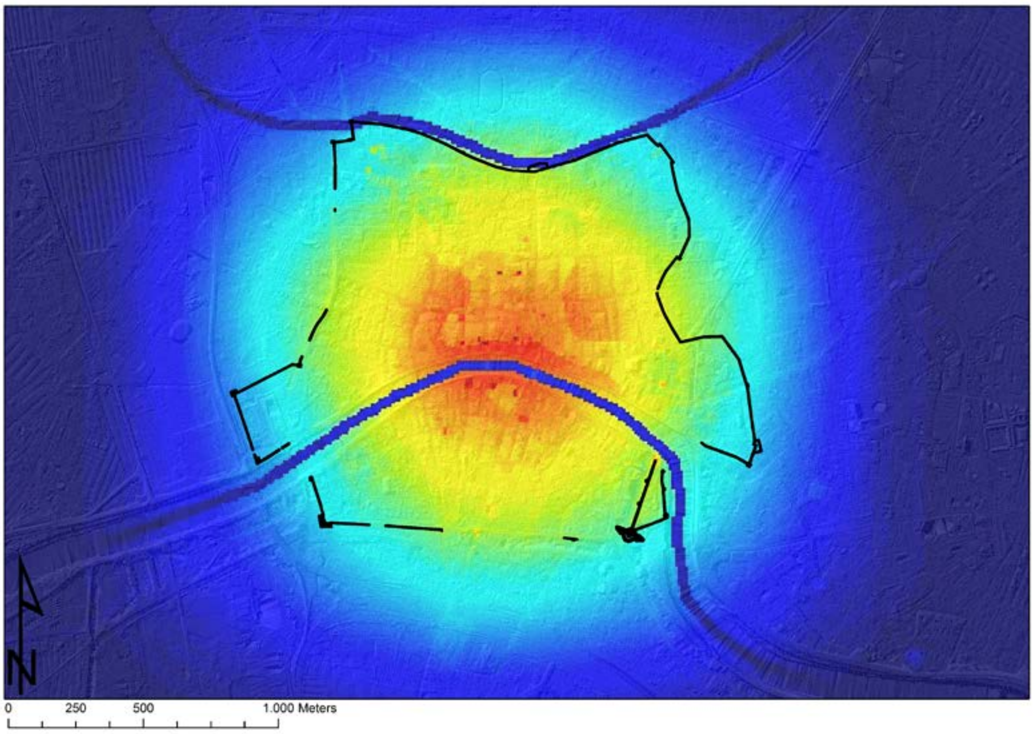


Figure 9: The absolute archaeological potential map estimated with the MAPPA PageRank based algorithm that is characterized by a clear role in enclosing the urban areas of the walls, which act as a border, and the ability to consider geographical element, such as the rivers, as a barrier.

that spreads its values from the middle of the city were the archaeological findings with highest potential values are conserved. As we have seen, this interpolation produce a continuous model, whilst we are aware that the archaeological potential does not always develop as a continuum. In fact, the outcome prediction is unable to appreciate sharp interruptions such as the presence of city walls, or of a geographical barrier even if we consider the geomorphological datum as the base datum in the interpolation (Fig. 7).

As noted before, for the experimental r.finder analysis, the typological class of tower-houses has been considered, and – as thematic layers - one qualitative map (geological units) and three quantitative maps: DEM, slope, aspect, and distance buffers from water streams and city wall belt.

Such thematic layers clearly rely on the standard landscape approach, whereas in this case, the 'closed' nature of urban space implies stronger constrains which forces anthropic elements. As a matter of fact, the analysis didn't show a strong relevance of such thematic layer, and the only meaningful, however interesting output, seems to be the indication of areas half-way between the city wall belt and the river as the most suitable for tower-houses placement (fig.08).

The final outcome of the MAPPA algorithm is an archaeological potential map for late medieval period, that is characterized by a clear role in enclosing the urban areas of the walls, which act as a border, and the ability to consider geographical element, such as the rivers, as a barrier. Anyway, both in the functional area picture, and in the estimated archaeological potential a suburban area can be noted outside the walls, or on both the rivers banks (fig.09).

Conclusions

From an archaeological point of view, the aim of these comparison is to illustrate which method is more useful depending on the datasets available. The Kernel density estimation output map represent simply the spatial concentration of archaeological finds depending on the radius we choose. This method seems useful when we have a single-non parametric dataset, but its non-parametric function depict a sort of self-fulfilling prophecy. The necessity to manage a lot of different parameter in the co-kriging interpolation and the difficulty to consider sudden interruption can easily produce untrustworthy maps. From this point of view the co-kriging seems the less useful method for the estimation of the archaeological potential. For a better functioning of this method you need to use different parameters as tested in the evolution of the urban fabric in (Gattiglia, 2014). R.finder and the PageRank MAPPA algorithm are able to catch the archaeological complexity and the sharp geographical or anthropic barrier, this demonstrates the necessity to develop predictive models specifically made by and for archaeologists.

From the technical poiny of view, the aspect the authors wants to highlight is the difference in the performance between statistical and mathematical methods. Kernel density estimation and co-kriging outs maps reflecting only the statistical properties of the input variables like the density distribution of the finds in the case of KDE and the spatial autocorrelation in the case of kriging. In a certain sense, a statistical property can be appropriate or not for a dataset, but cannot distinguish between the nature of different objects in the dataset, among which basic working rules can be imposed through equations in a mathematical models. The kernel density estimation provides a map with an approximate idea of the data spatial structure, mostly given by the distribution of available data. In areas with scarce quantity of data this method seems to be poor in estimating a reasonable predicted density of finds giving an approximation of the archaeological potential. The cokriging highlights the fact that the purely statistics methods cannot put in place the modeling capabilities of a mathematical method, where instead you can put some assumptions that satisfy some basic mechanisms of modeling :in the case of the archaeological potential the assumption that 'most links lead more potential '. The co-kriging tends to concentrate the value of archaeological potential in the area with more data and tends to decrease further one moves away from the concentration. R.finder was developed to work in a landscape context, but nonetheless gives interesting results here, where it is tested on a urban area. In such a context, the case-study opens also a useful perspective of discussion on different approaches to both territorial settings. This method shows the typical strengths and weaknesses of a map algebra based algorithm, which can combine factors in great quantity and of different nature, but these combinations are obtained with binary logical operations, so any complexity of the dataset read to be beyond such rigidity cannot be taken in consideration. The PageRank based algorithm is expressly realised to estimate the archaeological potential of an urban area. The model is suited for working with datasets of different nature (archaeological, geomorphological, historical, and so on), and it treats all the data as if they are in a complex network. In this way the archaeological potential rises through the interactions between the different elements. The PageRank model is capable of assigning archaeological potential to cells, ranking them on the basis of their interactions, and it turned out to be a good choice in estimating the archaeological potential. However, the PageRank model seems to work better for those data which concentrate in an area, than for 'polycentric' data. The PageRank model was tested through a series of core-drillings, but we have no general method, e.g. a minimum amount of data or relations, to decide whether a PageRank based model could be appropriate, for instance with respect to a more standard one.

References

ANICHINI, F., FABIANI, F., GATTIGLIA, G. and GHIZZANI MARCÌA, F. (2013). Questione di pesi: valori parametric, relazioni per il calcolo del potenziale archeologico. In: Anichini, F., Dubbini, N., Fabiani, F., Gattiglia, G. and Gualandi, M.L. Mappa. *Metodologie Applicate alla Predittività del Potenziale Archaeologico*. 2. Roma: Edizioni Nuova Cultura.

BEARDAH, C. and BAXTER, M. (1996). The archaeological use of kernel density estimates. *Internet Archaeology 1*. Available from: http://intarch.ac.uk/journal/issue1/beardah_index.html [Accessed: 24th July 2014]

DRENNAN, R.D. (2010). *Statistics for archaeologists.* NewYork: Springer.

DUBBINI, N. and GATTIGLIA, G. (2013). *A PageRank based predictive model for the estimation of the archaeological potential of an urban area*. In: Addison, A.C., Guidi, G., De Luca, L. and Pescarini, S. (eds). *Proceedings of the 2013 Digital Heritage International Congress. Marseille, 28 October – 1 November 2013.* Danver: IEEE. pp. 571-578

GATTIGLIA, G. (2014). *Pisa in the Middle Ages: archaeology, spatial analysis and predictive modelling.* Roma: Edizioni Nuova Cultura.

GUMERMAN, G.J. (1971). *The Distribution of Prehistoric Population Aggregates.* Prescott: Prescott College Press.

HODDER, I. and ORTON, C. (1976). *Spatial analysis in archaeology.* Cambridge: Cambridge University Press.

HUTCHINSON, M. F. (1989). *A new procedure for gridding elevation and stream line data with automatic removal of spurious pits.* Journal of Hydrology. 106. pp. 211–232

JUDGE, W.J. and SEBASTIAN, L. (1988). *Quantifying the Present and Predicting the Past*. Denver: U.S. Department of the Interior, Bureau of Land Management.

KAMERMANS, H., VAN LEUSEN, M. and VERHAGEN P. (eds) (2009). *Archaeological prediction and risk management*. Leiden: Leiden University Press.

LANGVILLE, A.N. and MEYER, C.D. (2006) *Google's PageRank and Beyond*. Princeton: Princeton University Press.

LIESKOVSK, T., ĎURAČIOVÁA, R., and KARELL, L. (2013). *Selected Mathematical Principles of Archaeological Predictive Models*. Interdisciplinaria Archaeologica Natural Sciences In Archaeology. IV (2). pp. 177-190

LLOYD, C. D. and ATKINSON, P. M. (2004). *Archaeology and geostatistics.* Journal of Archaeological Science. 31. pp. 151-165

MATHERON, G. (1971). *The Theory of Regionalized Variables and its Applications.* Fontainbleau: École Nationale Supérieure des Mines de Paris.

MOORE, J.A. and KEENE, S. (1983). *Archaeological hammers and theory.* New York: Academic Press.

NETELER, M., MITASOVA, H. (2004). *GRASS-Gis: an open source Gis approach*. New York: Springer.

PALOMBINI, A. (2013). r.finder: uno strumento per l'analisi geografica predittiva. In: NEGRETTI, M., MARIN, V., SALMONA, P. (eds). *Geomatics Workbooks*, 1. XIV meeting degli utenti italiani GRASS e GFOSS. Genova, 6-9 febbraio 2013. Milano: Politecnico di Milano.

VERHAGEN, P. and WHITLEY, T. G. (2011). *Integrating Archaeological Theory and Predictive Modeling: a Live Report from the Scene*. Journal of Archaeological Method and Theory. 12 (1). pp. 49-100

WHEATLEY, D. and GILLINGS, M. (2002). *Spatial Technology and Archaeology.* London: Taylor and Francis.

Proceedings of the 42nd Annual Conference on Computer Applications and Quantitative Methods in Archaeology
CAA 2014 - 21st Century Archaeology/F. Giligny, F. Djindjian, L. Costa, P. Moscati, S. Robert (eds.)

Analyses of Bone Modifications on Human Remains: A GIS Approach

Elsa CIESIELSKI, Hervé BOHBOT

ASM - Archéologie des Sociétés Méditerranéennes, UMR5140, Univ Montpellier 3, CNRS, MCC, Montpellier, France
elsa.ciesielski@gmail.com , herve.bohbot@cnrs.fr

Abstract

The analysis of fragmentation and cut marks on human remains usually relies on extensive textual descriptions, sketches on bone diagrams and recording in data tables, and most often on a combination of these methods, which each have their disadvantages. Some zooarchaeologists have found ways of facilitating these charts analyses, and a few authors have occasionally adapted them to human bones: using Geographic Information Systems (GIS) technology, where 'base maps' are sketches of skeletons instead of geographical areas. We applied this approach to a set of Gallic severed heads coming from Le Cailar (Gard, France) on recent GIS software to investigate the cutting and fracturing. Since it is an ongoing study, we mainly discuss our method and questions raised by the creation of MNE (minimum number of elements) and density maps. The preliminary results indicate a great accuracy of MNE estimating, and a similarity between the different estimation methods of cut marks or fractures densities.

Keywords: GIS, human bone, minimum number of elements (MNE), cut mark, fracture, density

1. Archaeological context

Archaeological excavations at 'place de la Saint-Jean' (Le Cailar, Gard, France, fig. 1) started in 2003 under the direction of R. Roure (Associate Professor, Université Paul-Valéry Montpellier 3). Archaeologists have unearthed the remains of a fortified settlement occupied since the 6th century BC. At the foot of the fortifications, in an interior plaza, several layers containing thousands items (8,000 of which have accurate coordinates) dated from the 3rd century BC have been discovered, including human skulls and metallic objects (where the majority is typical Celtic warrior equipment, see Girard & Roure 2009, 2010), but also coins, faunal remains and ceramics (Roure et al., 2006). While fauna and ceramics appear to come from successive backfills, metal objects, human skulls and coins are evidence of a well-known practice in the Gallic world: the displaying of severed heads. Classical authors such as Posidonius, whose works are known from references in the texts of Diodorus of Sicily (Library of History, V, 29) and Strabo (Geography, Book IV, 4, 5), assert that the Galls bring back the heads of their enemies, hanging from their horse's necks, and nailing them at the entrance of their settlements (see Ciesielski et al. 2011 for discussion about the different translations). In southern France, this practice is known both by lapidary representations and by human remains (fig. 2). There are several types of stone carvings: steles or low reliefs showing severed heads, freestanding sculptures portraying illustrious characters cross-legged with severed heads on their knees, and porticoes with cavities in which heads may be slotted into (Arcelin & Plana, 2011, and short notes on archaeological sites supra in Roure & Pernet eds., 2011). Human remains attesting this practice (skull and cervical vertebrae) were found in about fifteen sites, but only six assemblages can actually be considered as severed heads due to the presence of cut marks (skull without question mark on fig. 2). Some remains, like in Entremont, were found with nails used to hang them up. The association of human remains and metal objects is known in Catalonia (e.g. Ullastret, Spain) and in northern France (e.g. Ribemont-sur-Ancre, Somme, where bodies without heads were discovered), but was unprecedented in southern France. In this context, the particularities of the Cailar's bone samples are: from a recent excavation, a large assemblage of remains, and regroup well-preserved bones showing numerous traces.

2. Material

Currently, about 2,700 fragments of human bones (1,000 of which are teeth) were discovered. They represent about 53

FIGURE 1: LOCATION OF LE CAILAR (GARD, FRANCE).

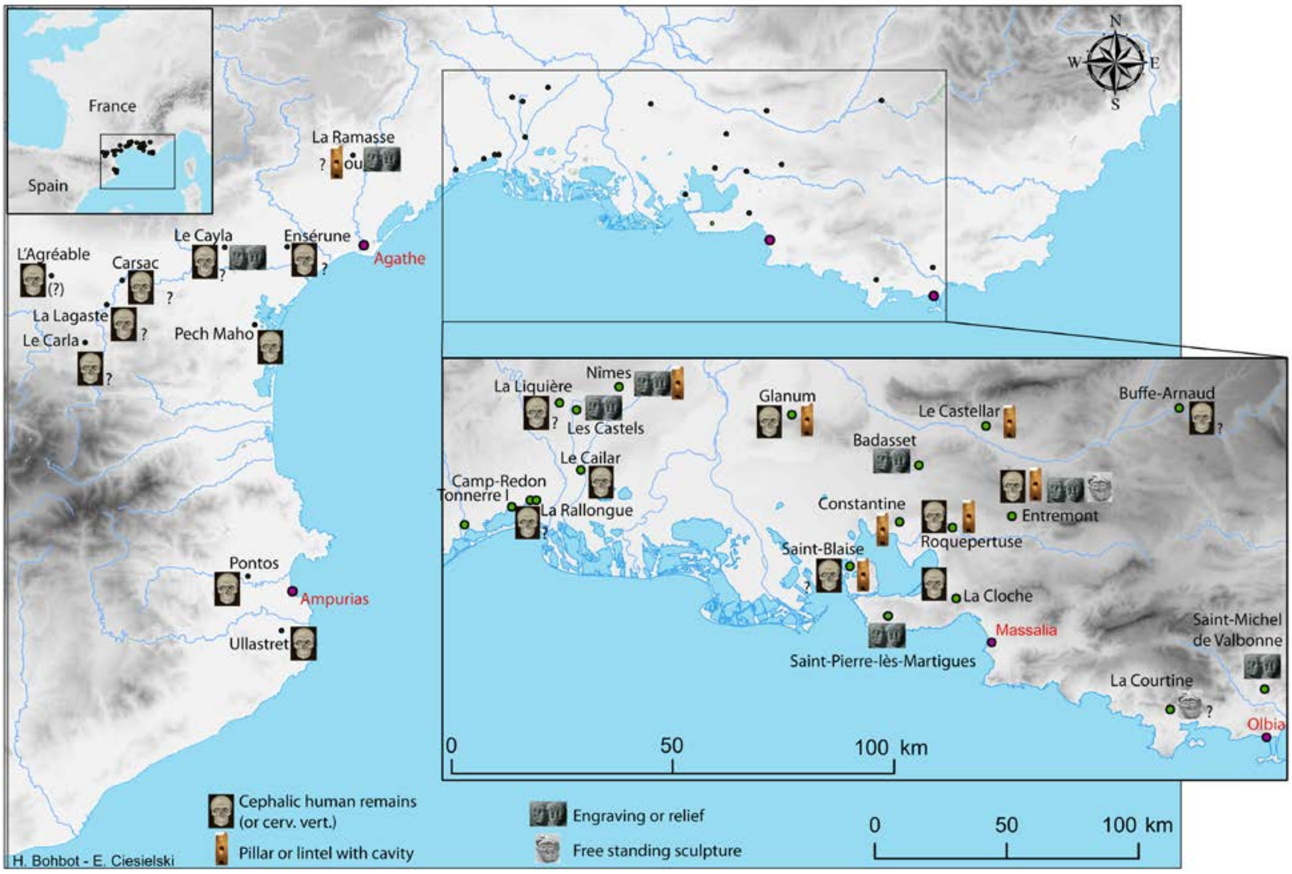


Figure 2: Locations of regional archaeological sites relate to the displaying of human heads during the Iron Age.

individuals. All of them are fragmented and commingled and are exclusively cranial remains except for six fragments of cervical vertebrae. All individuals are adults or possibly adolescents seeing as some have incomplete maturation of the third molars. Some are extremely robust while others have more gracile features (according the definition of cranial features by Acsádi and Nemeskéri, 1970), but as we do not have the entire skull, it is useless to record them. Indeed, given that it is difficult to determine the sex of an individual without coxal bone (Bruzek, Schmitt 2008) this is all the more true for intermingled skull fragments; especially we do not have reference data because incineration is the main funerary practice for these populations.

3. Why geographic information system (GIS)?

The purpose of this work with GIS is to study cut marks and fracturing of a large, extremely fragmented and commingled assemblage. In order to study these bone modifications in physical anthropology or zooarchaeology, we customarily use textual descriptions, recordings in data tables, handmade bone diagrams or often combinations of the three methods (White, 1992; Lyman, 1994; Boulestin, 1999; Lyman, 2008). Each of these methods has their own disadvantages (fig. 3). Textual descriptions may be very detailed, but rapid and relevant syntheses using computer tools are impossible. Storing bone-surface modifications in data tables cause significant data losses due to discretization, especially when it comes to describing fragment morphologies and cut mark locations.

As a consequence, it is usual to represent cut marks on bone diagrams to complement data tables. In any case, summarizing large datasets of both graphic and textual data is difficult if not impossible (fig. 4). Moreover, for fragments of skull remains it is very tricky to input data in tables because, unlike fragments of long bone shafts – which are generally most studied – their shape is more irregular, so it is hard to pick a 'typical' shape to describe a fragment in a database. This step is more efficiently done with computer analysis tools than human subjectivity.

The GIS approach allows to manage, display and analyse spatially referenced data. It may be used for various purposes and at various scales, including non-geographic applications. As long as it comes to analysing spatial relationships between

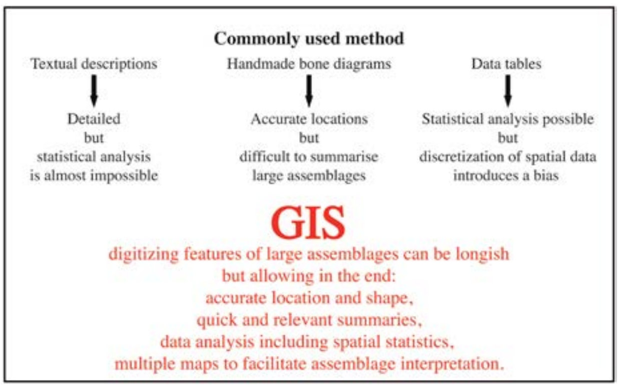


Figure 3: Summary of advantages and disadvantages of commonly used method and GIS approach.

features (points, lines, polygons) it is natural to use GIS in this particular case because features are bones, cut marks and fractures here. Marean et al. (2001) and Abe et al. (2002) were the first to apply GIS techniques to zooarchaeology to quantify fragments and cut marks (BoneEntryGIS, based on discontinued ArcView® 3.x software from ESRI). They draw fragments and cut marks on computerized drawings of skeletal elements, as vector data instead of doing it by hand as usual, allowing further computer analysis.

After that, they proposed to calculate a MNE (minimum number of skeletal elements) by pixel overlapping and to correct cut mark frequencies by the observed surface area. R.L. Lyman (2005, 2008) criticized this approach. According to his own tests he denies the replicability of fragment drawing between observers and he finds an overestimate of 0-50 % than the actual number. Conversely, Parkinson et al. (2014) find consistently replicability in fragment drawing and a slightly underestimated of MNE relative to the direct tallying with bones. She argued that the difference with R.L. Lyman is due to the level of training in bone anatomy of drawers, we concede that, but one may wonder if the difference is not due to the size of fragments and collection too. On the other hand, R. L. Lyman (2005) does not think it appropriate to weight frequency of cut marks by the observed surface area for several reasons, including the fact there is no reason the non-preserved area display the same cut mark ratio as the preserved one. We agree on this point, but we believe it is still needed to take into account the percentage of observed surface.

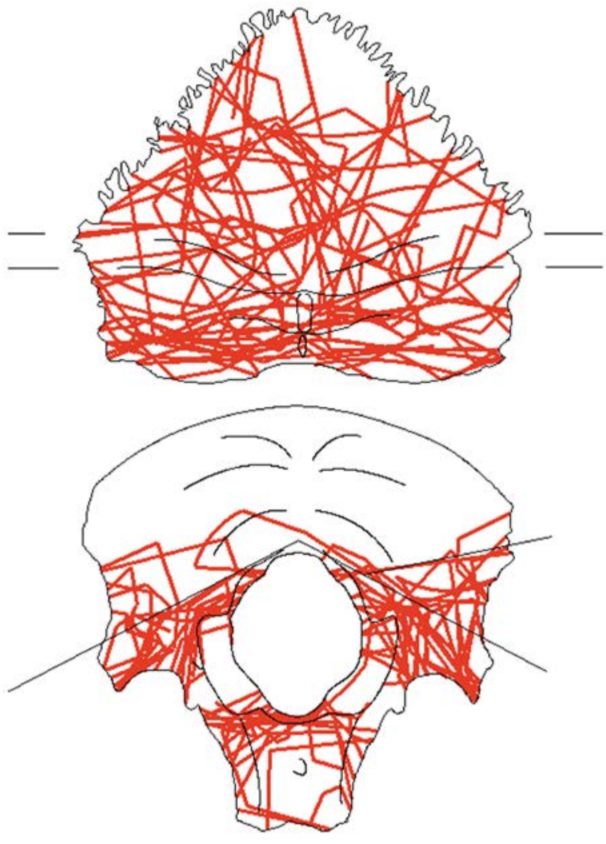


Figure 4: Summary of occipital bone fractures (difficult to analyse with classical hand drawn sketches!).

Finally, although digitizing features of a large assemblage can be longish, GIS has many advantages. It allows to have accurate location and shape, quick and relevant summaries, make data analysis, including spatial statistics, and create multiple maps to facilitate the assemblage interpretation (fig. 3). As we are trying to find if there is a particular location for a given type of cut mark or fracture, it is important to not bias data, before any analysis, by discretizing location or outline to fit data tables items.

4. Methods

Even if we do not use *BoneEntryGIS*, which is based on the discontinued ArcView® 3.x software, we necessarily drew on this software to develop a similar application based on the latest version of ArcGIS (ESRI, Redlands, CA). Our method is slightly different in terms of structure (e.g. all bone outlines are stored in the same layer instead having a layer for each bone in *BoneEntryGIS*) and features (e.g. we added the characterization and the analysis of fractures).

4.1 Model creation

We created the human bone model in ArcMap (name of the mapping application of the ArcGIS framework). We used a polygon feature class to represent the shape of skeletal elements and a polyline feature to represent anatomical landmarks (fig. 5). The latter is important to increase the accuracy to draw the outline – like the photographic template used by Marean *et al.* (2001) – in order to minimize possible under or overestimates.

4.2 Fragment digitizing and data recording

For each specimen (in osteology, it means an isolated fragment) we have drawn their outline, fractures and cut marks with the traditional tools of the *Editor* such as *Create feature*. Because the shapes of bones are complex their outlines were created by copy and paste of a model. We started by drawing fractures, then we used the tools *Reshape Polygon* or *Cut Polygon* on the copy of polygon - being sure that *snapping* is enabled - to have perfect matches between outline and fracture.

Unlike the Marean's method, each feature types is on a dedicated layer, that means that we have one layer for outlines, one for fractures and one for cut marks. To avoid overlay and ease the drawing of features, we simply display or hide them by using the definition query in layer properties. When we create a new feature, a field has a default value matching the definition query, and we change the latter to hide the feature and draw the next one. We think it is much easier to work with only three shape files rather than to have one shape file by specimen, all the more if you have a lot of fragments, whether it is for data creation and management, for querying, or for using geoprocessing and spatial analysis tools. This allows to use recent built-in ArcGIS tools instead of having to develop scripts to handle multiple layers. After digitizing,

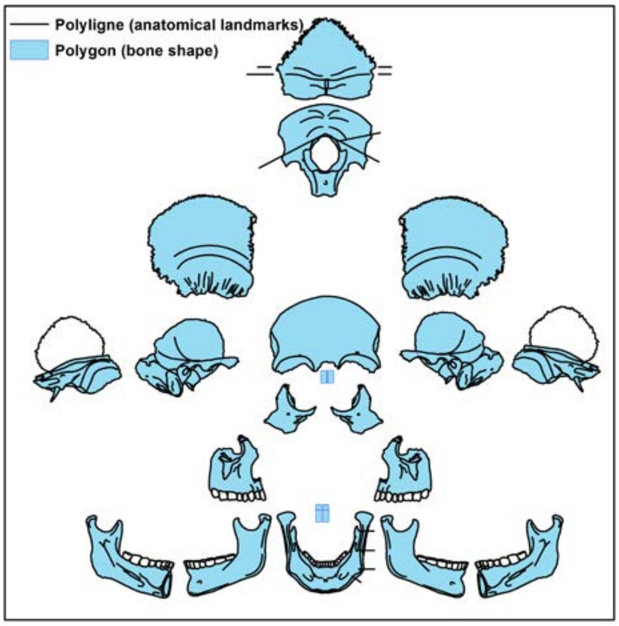


Figure 5: Model of human skull in ArcMap.

the bone references and the characteristics of fractures and cut marks are populated in the attribute table of the layer (via the *Attribute* window in the *Editor*, fig. 6).

4.3 Data analysis

The minimum number of skeletal elements (MNE) is derived from the overlapping parts of polygon outlines. Like Marean *et al.* we can display MNE (fig. 7 and 8) by skeletal element, but we also calculate proportions based on site's MNI (minimum number of individuals) (fig. 9). Our method is different because we do not have one layer by specimen, so we used current ArcGIS tools instead of their scripts.

To create our MNE maps, first we use the *Union* tool to create polygons for each part of bone created by the superimposed outlines. Then we use the *Spatial join* tool in order to have a new layer with a field counting the number of superimposed parts (the counterpart of the MNE). To have this equivalence, the same layer is used for the target and joined features and the match option should be 'are identical to'. Afterwards, a symbology is applied to the 'join count' field to display a map of bone portion survivorship. On large assemblages, this vector layer may be very big and it may be necessary convert it to a raster layer. Then, this raster values may be divided by the site's MNI, thus highlighting potential lacks of skeletal elements.

To analyse the repartition and the density of marks on bones, Parkinson *et al.* (2014) suggest the use of *nearest neighbour* and *kernel density* tools. This is also our approach, but during the process some questions about the tools have arisen. The *Average nearestneighbour* tool is used to know if features are dispersed or clustered. However, this function is designed to analyse the area of

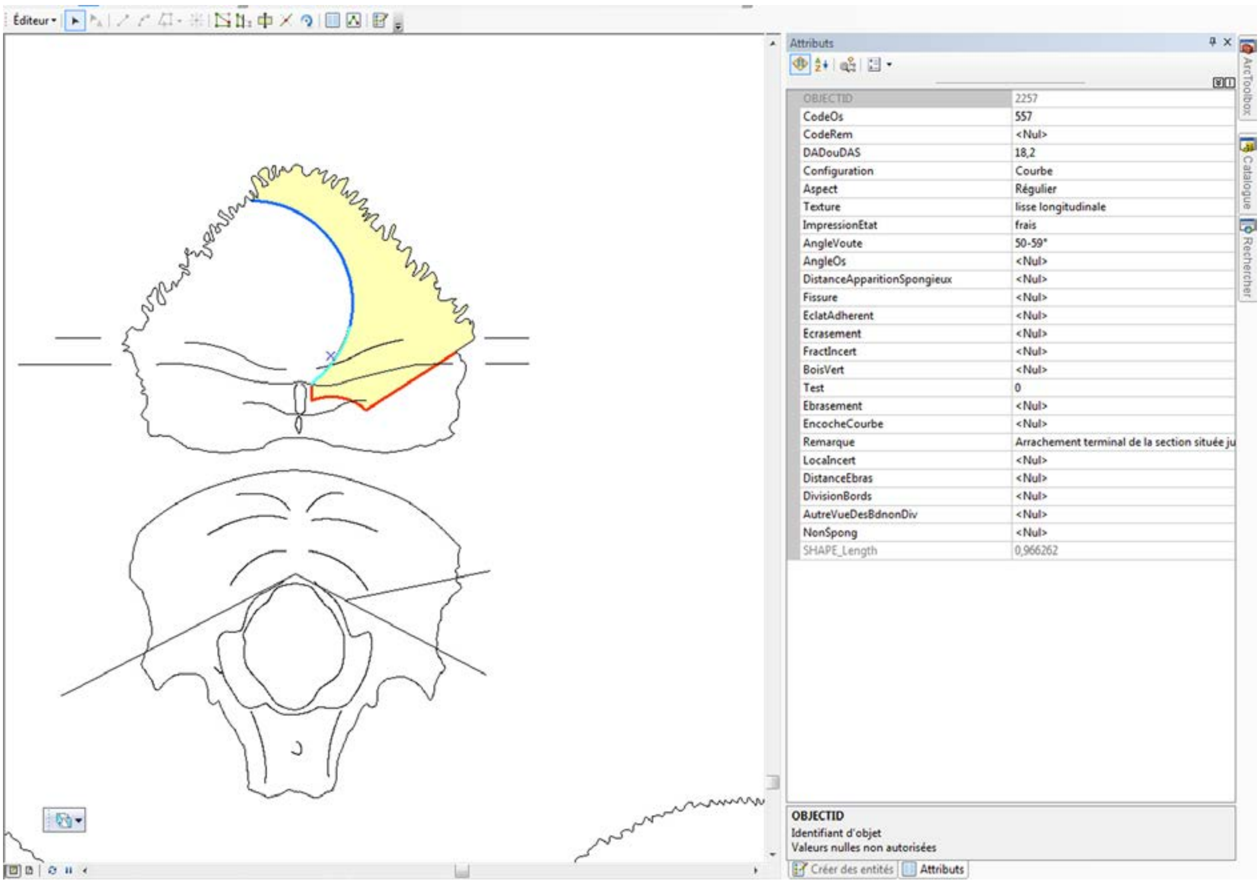


Figure 6: Example of cut mark (blue) and fractures (red) digitizing, and example of attributes entered for the fractures.

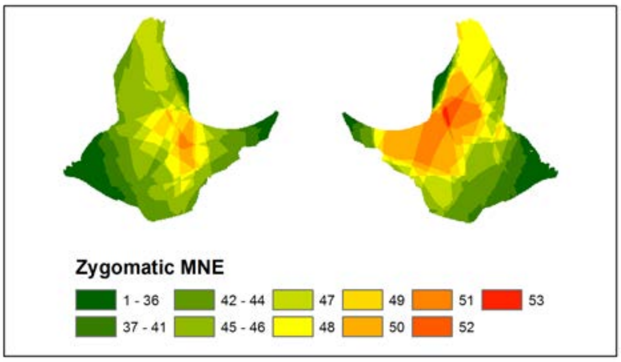


Figure 7: Map of bone portion survivorship for the zygomatic bone.

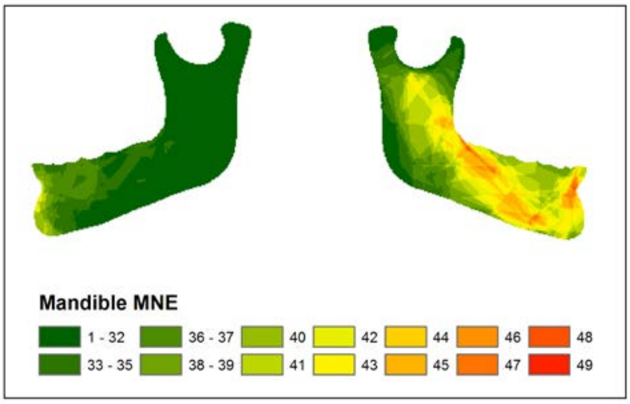


Figure 8: Map of bone portion survivorship for the mandible bone.

the minimum enclosing rectangle that would encompass all features, but bones do not have a regular area. Parkinson *et al.* (2014) suggest calculating the real area with the bone template. We agree, but in our case we think this area must be modified if we want to analyse some more specific areas. For example, in order to analyse the cut marks that are on the posterior edge of the ramus and on the inferior edge of the mandible body, we need to know the real area. Otherwise the result shows a cluster that is actually due to the fact that we look at cut marks that are only located in some parts of the bone. This method is also subject to edge effects, which could be difficult to overcome for non-rectangular areas. As always in statistics, we must be careful with the tools used and always try to better adjust them to our sample.

The same kind of issues arises with the use of density tools. Kernel density could be very useful to summarize data, but further testing is needed to know whether this tool is suitable or not. For example, it calculates a magnitude per unit area, which is difficult to relate to absolute density. We could understand the variation of magnitude, but this one is difficult to transform in number of features especially for line features. Kernel density may be calculated in several ways, but with ArcGIS there is only smoothing equation available (the quadratic function), which is rather for a uniformly distributed phenomenon (Zaninetti, 2005). There is also the possibility to apply only a constant bandwidth (search radius), which does not allow to have

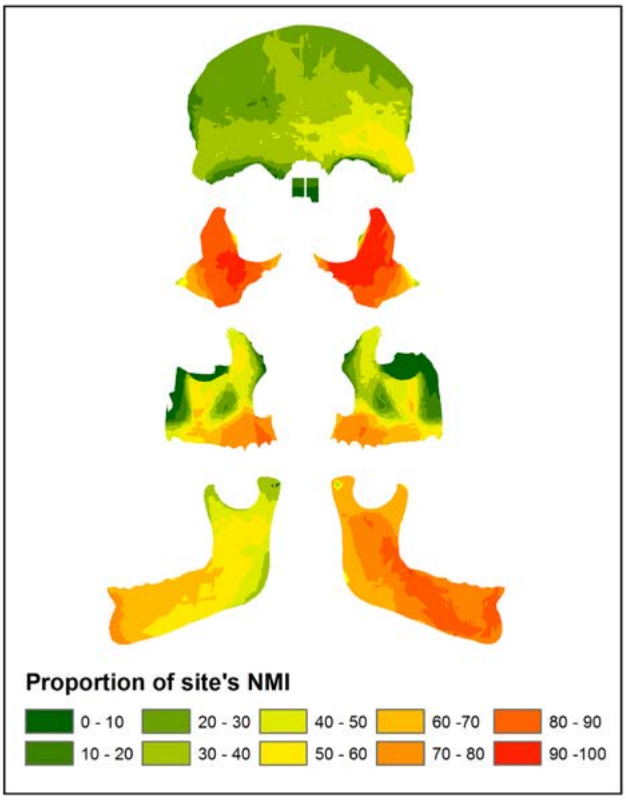


Figure 9: Map of the proportion of bone portion survivorship based on site's NMI.

the same precision in lower density areas (Zaninetti, 2005). The application of this method also assumes that the data is exhaustive and that the pattern studied is areolar and isotropic which it is not in our case. Accordingly, it could be useful to simply calculate an absolute density, which is equivalent to the number of marks per unit area. To this end, we create a fishnet with polygon features, then we do a spatial join between this fishnet and the cut marks. The results are visually similar (fig. 10) but with this method we could have a better idea of the real number of cut marks. It could be a start to apply the quadrat method, even if it is less powerful than nearest neighbour. The next step will be to weigh these density maps with the MNE maps in order to minimize density differences due to an underrepresentation of some bone parts. For example, in figure 11, a lower density is visible on the left side but MNE is also low, indicating a lower conservation rate: as a result, the difference in cut mark density is probably due to conservation rather than human actions.

5. Conclusions and perspective

Currently, we have explored several methods for the quantification of fractures and cut marks, that need more testing to understand if they are really appropriate to bone modification studies. We also use CrimeStat (a software developed by Ned Levine & Associates, Houston, TX, and the National Institute of Justice of Washington, DC) and SADA (developed by the University of Tennessee Research Corporation) that provide additional tools for spatial analysis and allow analysis and creation of shape files.

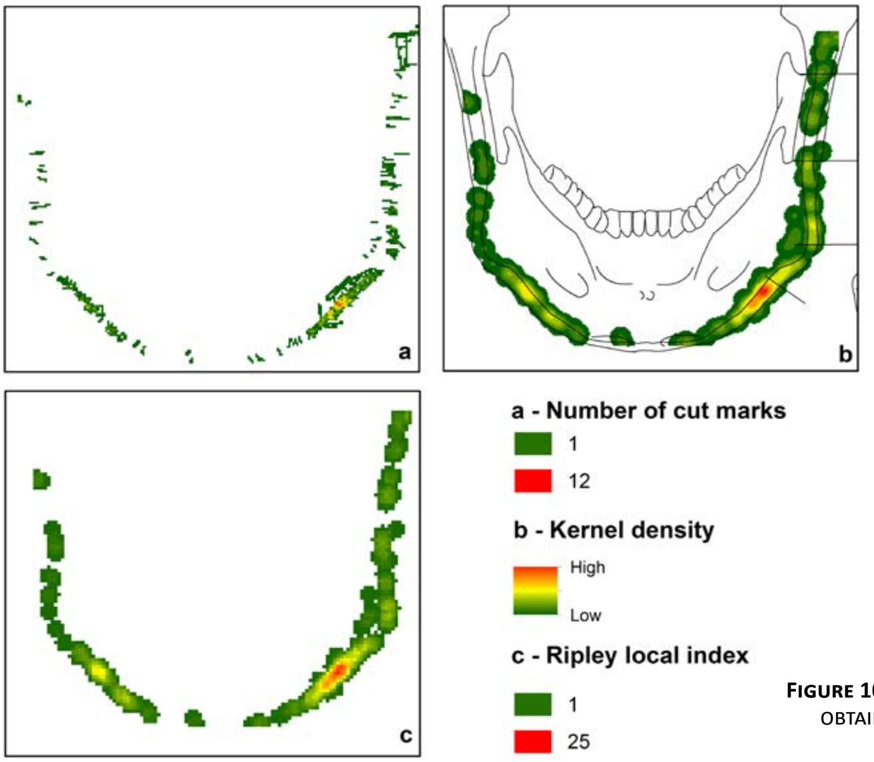


Figure 10: Comparison of the density obtains with several methods.

We conclude, like Parkinson and coll., that MNE calculation with a GIS approach is as accurate as direct tallying with bones because we obtain very similar results with both methods. For example, for the zygomatic bone we obtain a maximum of n=52 individuals with direct tallying and n=53 with the GIS approach (fig. 7 shows that the area with n=53 is very small and probably an artefact due to drawing). Finally, we have similar results but the GIS shows an accurate map of bone portion survivorship. Another advantage of the GIS approach is to facilitate the understanding of cut mark frequency variation by comparing MNE map and density map. The next step will be to study frequency variation of fractures according to their characteristics because one of the goals of our study is to determine if fractures occur on fresh or dry bones. Usually fracture angle is an indicator of freshness: higher values meaning drier bone fractures. With this GIS approach we could easily compare accurate locations of fresh and dry fracture lines in order to propose a statistical assessment of the observed fracture patterns.

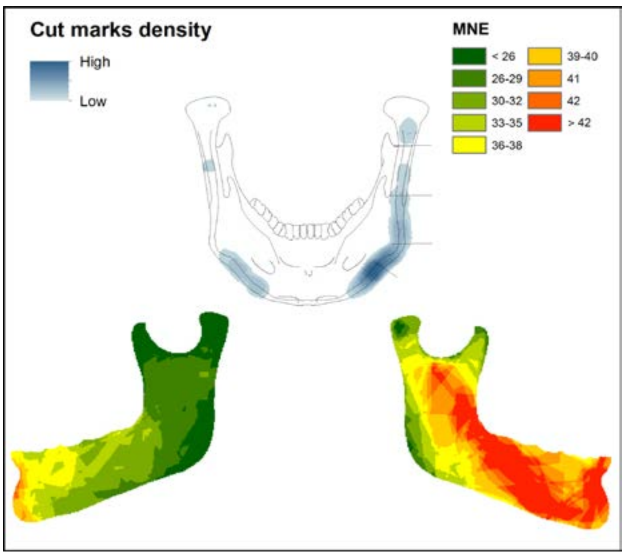


Figure 11: Density of cut marks and bone portion survivorship of the mandible bone.

This methodology is developed as part of a PhD but besides we would like to develop a software add-in for the ArcGIS environment, based on this initial work. This toolbox would be aimed at easily characterising human bone datasets with basic GIS knowledge. Subsequently, the use of a common framework would help anthropologists in order to compare assemblages from various sites.

Acknowledgements

This work is part of a study for a PhD in archaeology at Université Paul Valéry Montpellier 3 (France) under the supervision of T. Janin (Professor, Université Montpellier 3) and H. Duday (Emeritus Research Director, UMR PACEA - Université Bordeaux 1), it will be defended in 2015. This work has received financial support from Labex ARCHIMEDE, program 'Investissements d'Avenir ' ANR-11-LABX-0032-01.

References

ABE, Y., MAREAN, C.W., NILSSEN, P.J., ASSEFA, Z. & STONE, E.C. (2002). The Analysis of Cutmarks on Archaeofauna: A Review and Critique of Quantification Procedures, and a New Image-Analysis GIS Approach. *American Antiquity* 67 (4). pp. 643–663.

ACSÁDI, G. & NEMESKÉRI, J. (1970). *History of Human Life Span and Mortality*. Akadémiai Kiadó. Budapest.

ARCELIN, P. & PLANA, R. (2011). L'expression monumentale des rites protohistoriques en Gaule méditerranéenne et dans la partie nord-est de la Péninsule Ibérique. Stèles, bâtiments cultuels et sculptures. In: R. Roure & L. Pernet (eds.). *Des Rites et des Hommes : les pratiques symboliques des Celtes, des Ibères et des Grecs en Provence, en Languedoc et en Catalogne*. Archéologie de Montpellier Agglomération 2. Errance.

BOULESTIN, B. (1999). *Approche taphonomique des restes humains : le cas des Mésolithiques de la grotte des Perrats et le problème du cannibalisme en préhistoire récente européenne.* BAR International series 776. Oxford, England: Archaeopress.

BRUZEK, J. & SCHMITT, A. (2011). L'identification du sexe d'un individu à partir du squelette. In: Charlier, P. (ed.). *Ostéo-archéologie et techniques médico-légales, Tendances et perspectives, Pour un 'manuel pratique de paléopathologie humaine'.* Coll. Pathographie 2, Paris: De Boccard, pp. 259-267

CIESIELSKI, E., DUDAY, H., GIRARD, B., ROURE, R., MARTIN, A. & AGUSTI, B. (2011) La pratique des têtes coupées et les dépôts d'armes en Gaule méditerranéenne et dans le nord-est de la Péninsule Ibérique. In: R. Roure & L. Pernet (eds.). *Des Rites et des Hommes : les pratiques symboliques des Celtes, des Ibères et des Grecs en Provence, en Languedoc et en Catalogne.* Archéologie de Montpellier Agglomération 2. Paris: Errance.

GIRARD, B. & ROURE, R. (2009). Le mobilier métallique du dépôt du Cailar (Gard) : quantification, composition et traces de manipulations destructives. In: M. Honegger, D. Ramseyer, G. Kaenel, B. Arnold, & M.-A. Kaeser (eds.). *Le site de La Tène : bilan des connaissances - état de la question.* Actes de la table ronde internationale de Neuchâtel, 1-3 novembre 2007. Archéologie Neuchâteloise. 2009, Hauterive, pp. 197–205.

GIRARD, B. & ROURE, R. (2010). Les armes du dépôt du Cailar, IIIe siècle avant notre ère. *Gladius*. XXX. pp. 43–60.

LYMAN, R.L. (1994). *Vertebrate Taphonomy*. Cambridge: Cambridge University Press.

LYMAN, R.L. (2005). Analyzing cut marks: lessons from artiodactyl remains in the northwestern United States. *Journal of Archaeological Science* 32 (12). pp. 1722–1732.

LYMAN, R.L. (2008). *Quantitative paleozoology.* Cambridge: Cambridge University Press.

MAREAN, C.W., ABE, Y., NILSSEN, P.J. & STONE, E.C. (2001). Estimating the Minimum Number of Skeletal Elements (MNE) in Zooarchaeology: A Review and a New Image-Analysis GIS Approach. *American Antiquity.* 66 (2). pp. 333–348.

PARKINSON, J.A., PLUMMER, T.W. & BOSE, R. (2014). A GIS-based approach to documenting large canid damage to bones. *Palaeogeography, Palaeoclimatology, Palaeoecology* 409. pp. 57–71.

ROURE, R. (2006). Armes et têtes coupées au Cailar (Gard) : premiers éléments de réflexion sur un dépôt rituel en Gaule méditerranéenne. In: P. Barral, A. Daubigney, C. Dunning, G. Kaenel, & M.-J. Roulière-Lambert (eds.). *L'âge du Fer dans l'arc jurassien et ses marges. Dépôts, lieux sacrés et territorialité à l'âge du Fer.* Actes du XXIXe colloque international de l'AFEAF. Annales Littéraires ; Série 'Environnement, sociétés et archéologie'. 2006, Bienne, 5-8 mai 2005: Presses Universitaires de Franche-Comté. pp. 653-658

ROURE R. & L. PERNET (eds.) (2011). *Des Rites et des Hommes : les pratiques symboliques des Celtes, des Ibères et des Grecs en Provence, en Languedoc et en Catalogne.* Archéologie de Montpellier Agglomération 2. Paris: Errance.

WHITE, T.D. (1992). *Prehistoric cannibalism at Mancos 5MTUMR-2346.* Princeton: Princeton University Press.

ZANINETTI, J.-M. (2005). *Statistique spatiale : méthodes et applications géomatiques*. Paris, France: Hermès science publications.

Territorial Organisation of the Terramare Culture (Bronze Age, Italy): Use of GIS Methodology to Tackle Societal Issues

Julie BOUDRY

UMR 8215 Trajectoires,Maison de l'Archéologie et de l'Ethnologie Nanterre, France
julie.boudry@gmail.com

Abstract

This paper presents the methodology used to address some territorial issues in the Italian Bronze Age Terramareculture. Through a reconstruction of the Bronze Age drainage network, this study shows close links between terramare sites and watercourses. Based on this re-established environmental context, several spatial analyses were implemented to examine settlement strategies. The spatio-temporal development of the occupation was studied with Ripley's Correlograms, deviational standard ellipses and kernel density.The hierarchical organisation between sites was reviewed through weighted Kernel density, spatial distributions of specific artefacts, settlement duration and architectural features. This paper also proposes a new typology for the mountain sites of this culture, based on data derived from a DTM.Lastly, the results reveal patterns of circulation and exchange, as well as three territories. A number of hypotheses are put forward about the social organisation of the terramare and the sudden appearance and decline of this culture.

Keywords: Terramare culture, statistical and spatial analysis, water network, societal collapse

1. Introduction

1.1. The Terramare culture and its peculiarities

The Terramare culture (Bernabò Brea et al., 1997) emerged and developed during the Middle and Late Bronze Ages (i.e. between 1650 and 1150 BC) in Italy. Its area of origin is located in the Po Plain and more particularly in the provinces of Parma, Reggio Emilia and Modena, in the Emilian region south of the Po River as well as in the low-lying plains of the provinces of Cremona, Mantua and Verona, north of the Po River (Figure 1). This culture is characterised by the eponymous terramaresites, defined by the presence of features surrounding the settlements, often an earthen embankment accompanied by a ditch.

The inhabitants of the terramare sites in the floodplain constructed irrigation and drainage networks, most likely for agricultural practices, frequently involving the re-routing of natural watercourses towards the peripheral ditch enclosing the sites (Balista *et al.*, 2004; Cremaschi *et al.*, 2007).

It is worth noting that the water network of the Po Plain has significantly evolved since the Bronze Age, notably its orientation (south/north today, south/north-east during the Bronze Age), mainly for tectonic reasons(Cremaschi *et al.*, 1980; Castaldini, 1989).

The Terramare culture shows highly contrasted demographic trends. From the Middle Bronze Age 2 onwards (1550-1450 BC), one sees demographic growth that cannot be attributed to natural conditions of demographic development and is therefore thought to have been triggered by the arrival of new populations (Cardarelli, 1988). Just as suddenly the terramare sites disappear from the Emilian plain between the end of the Late Bronze Age 2 and the beginning of the Final Bronze Age (around 1150 BC). Current research (Cardarelli, 2009; Cremaschi *et al.*, 2007; Cremaschi, 2009) focuses on a multifactor explanation of what appears to be a societal collapse. A drought is thought to have destabilised a system already affected by overpopulation and overexploitation of the territorial resources, thus implying poor management of this situation by the local elites. Moreover, these explanations have to be placed within a general context of instability.[1] This crisis is followed by the abandonment of the region south of the Po River for over two centuries until the emergence of the Villanovan culture.

1.2 Research questions and context of the study

The marked peculiarities of this culture were the starting point of our reflection. In fact the incoming populations

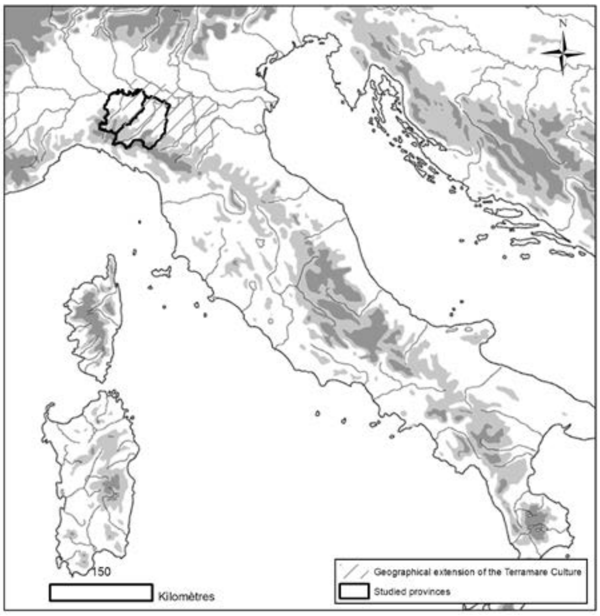


Figure 1: Geographical frame and studied provinces.

[1] At about 1200 BC the Mycenaean culture collapses, together with the Hittite empire and the large palatial centres in the Levant. In Egypt the arrival of the sea people is reported (De Marinis, 1997).

that established themselves in the Emilian plain developed within a territory that was previously very sparsely populated (Bernabò Brea, in press). They therefore had multiple choices with regard to settlement location in a new landscape, to territorial and societal strategies, as well as to internal site organisation. We therefore decided to focus the study presented here on the strategies of settlement foundation and territorial management of the terramare sites, takinginto consideration the geomorphology of the Po Plain and especially the reconstruction of its water network. This is indeed an essential aspect, given not only the strong impact of the construction of irrigation and drainage networks on the water network, but also the major changes in this network since the Bronze Age.

The study thus focuses on a spatial and geomorphological approach within a geographical area corresponding to the administrative provinces of Parma and Piacenza (Figure 1). These provinces were chosen as they have rarely been subject to studies of this type. The phenomenon of the terramaresettlements in the strict sense is related to the plain,as they are not found outside this area. However, sites with the same material culture are known in other geomorphological areas, for example the mountains and the hills. The entire surface area of the two provinces has therefore been analysed, as it includes various geomorphological zones (plain, hill and mountain, with the northern Apennines). The aim was to analyse the whole system and to understand the interactions between these different zones.

The archaeological corpus is comprised of 212 sites.[2] From a temporal perspective, the Early Bronze Age (2200-1650 BC) and the Final Bronze Age (1150-850 BC), corresponding to the beginning and the end of the phenomenon of terramaresettlements, were taken into consideration. The aim wasto gain a better understandingof processes of occupation and abandonment that took place within the area occupied by this culture.

2 Methodology

2.1 The reconstruction of a Bronze Age water network

Nearly 900 aerial photographs – of different periods, scales and treatments – and as many satellite images were studied. Several hundred traces were recorded and interpreted (Ferri, 1996; Piccarreta, Ceraudo, 2000).Out of these, 347 were retained as they were considered to be features related to paleowater networks. Only features with a south/north-east orientation[3], or directly linked to a terramare site (figure 8), were used for reconstruction of the water network. These traces were edited in .shp format using ArcGIS based on orthophotographs or georeferenced aerial photographs.If traces were observed with Google Earththeywere imported after conversion of the .kml file to a .shp file. These imported elements enabled us to run cross-analyses in the GIS, using data and notably georeferenced maps from geomorphological publications.

[2] 111 sites are located on the plain, for 101 mountain sites. The plain sites are grouped into various types according to their characteristics and evidence from archaeological fieldwork (terramare, open settlement, occupation, surface find, cemetery, hoard, basin, oven). Altogether, 45 % of sites are well dated and 20% have been excavated (Boudry, forthcoming). 3 The orientation attested for the Bronze Age water network.

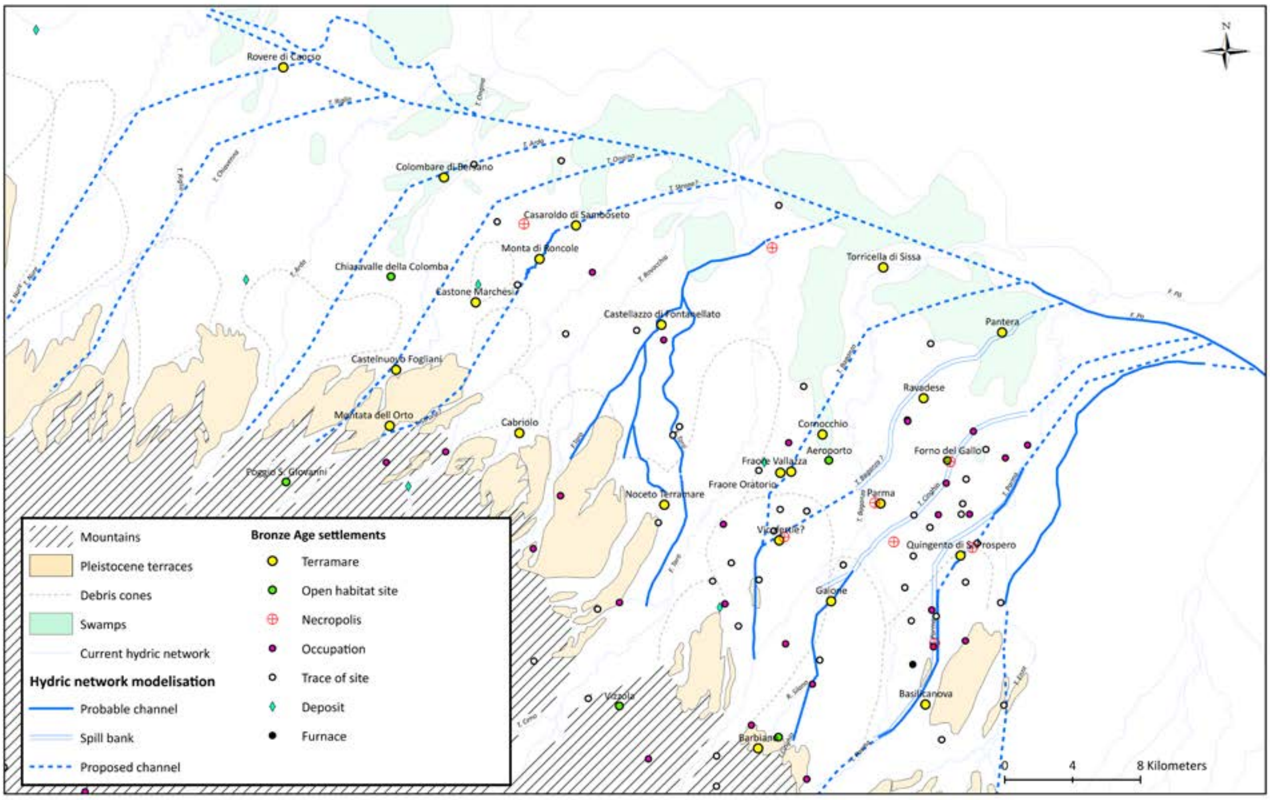


Figure 2: Bronze Age settlement locationsplotted against the geomorphology of the study area and the reconstruction of the water network.

The proposed reconstruction of the water network (Figure 2) thus includes Bronze Age paleochannels known frompublications,features observed to be in direct relationship with terramaresites, together with published Bronze Age levees , and lastly reconstructions based on some more generalpropositions for this period from the geomorphological literature, represented as dotted lines on the map.

This reconstruction of the water network was then transferred to a base map that was used to carry out a series of spatial and statistical analyses that enabled us to address issues of spatial and social organisation.

2.2 Territorial analyses

2.2.1 Spatial and temporal evolution of the occupation

Standard deviation ellipseand barycentre

The barycentre method, widely used in archaeology (e.g. Saligny *et al.*, 2008), makes it possible to identify possible phases of expansion, of retreat or of displacement of the archaeological sites over the course of time.

This method involves calculating the mean point of a series of points in a given space.The mean point corresponds to the mean x and y coordinates of the specimen taken into account. Around this central point the standard deviation ellipse develops as the method calculates the standard deviations in the x and y direction from the mean centre in order to define the ellipse axes. The ellipse makes it possible to show the orientation of the point cluster as well as its distribution; the bigger and wider the ellipse, the looser the site distribution; the smaller and more concentrated the ellipse around the barycentre, the more clustered the organisation of the sites. The two analyses were carried out using ArcGIS software.

In order to examine the spatial pattern of Terramare sites over the long term, this method was applied to the whole of our corpus, to provide a broad picture of spatial trends. Thus one sees a loose distribution of sites in Early Bronze 1, with a concentration on the plain from Early Bronze 2 until the end of the Late Bronze Age (with a slight dispersion in Late Bronze 1 however), reflecting a certain stability. Lastly, in the Final Bronze Age, there is apparently a withdrawal to the mountain zones (figure 3). These general results were confirmed by density calculations.

Ripley'sK-function and Besag'sL-function

Ripley's K-function turned out to be a suitable tool for the study of the spatial and temporal development of the Terramare settlement. We are dealing here with a method of exploratory spatial statistics based on the distance between the points, as this is the better known nearest neighbour method. These two analyses enable us to study the global pattern of point distribution (here the archaeological sites) in order to determine whether it is clustered, uniform or

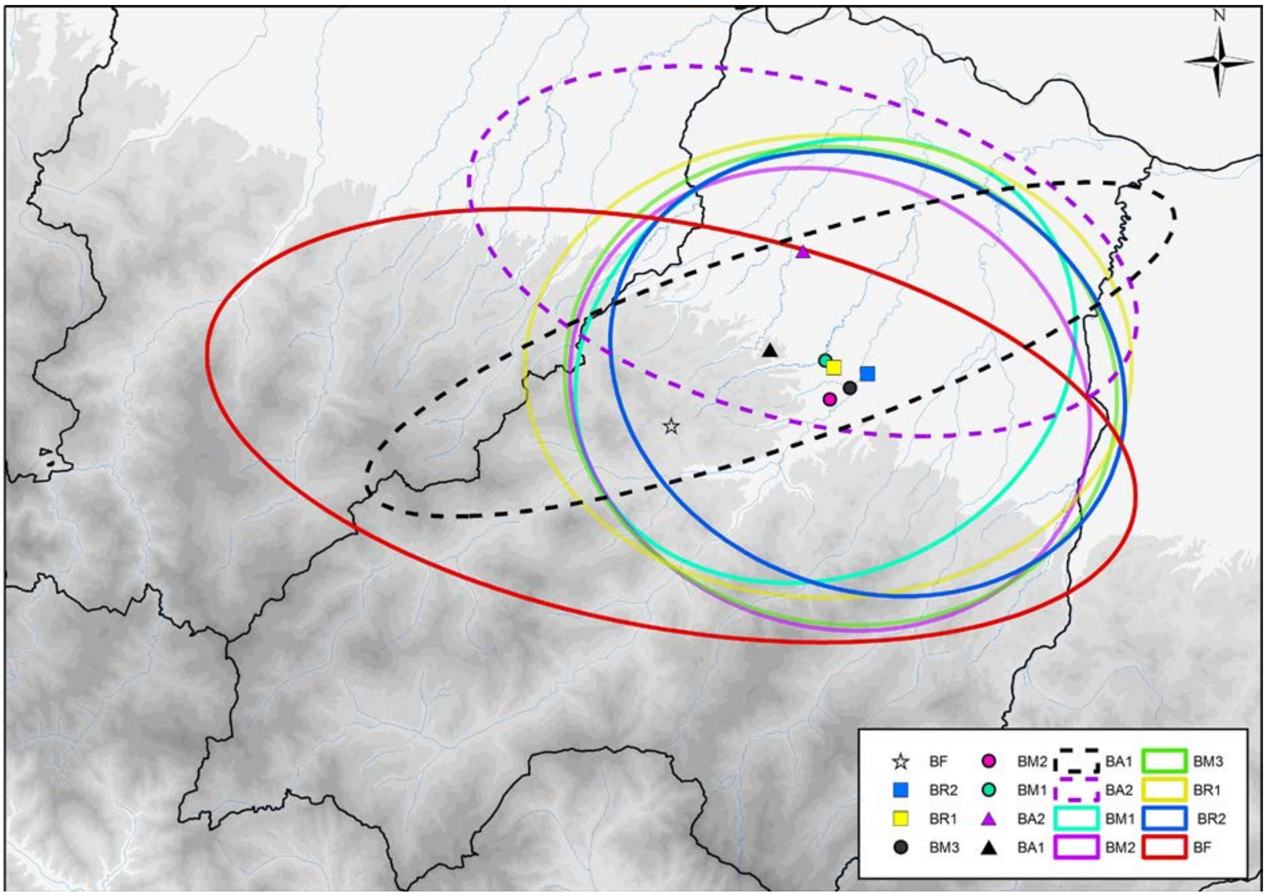


Figure 3: Barycentres and standard deviation ellipses from Early Bronze Age to Final Bronze Age.

rather random (Zaninetti, 2005). Ripley's K-function was preferred to nearest neighbour analysis as it enables a second-order analysis of neighbourhood. That means that the observation of the point distribution is made at several scales whereas the nearest neighbour method is closer to a first-order analysis of the distribution enabling aninterpretation of the phenomenon on a single scale only, which may introduce interpretation errors. The K-function therefore makes it possible to show more complex patterns of organisation. Thus on a certain scale, it will be possible to identify clusters of sites whilst on a different scale these clusters may be part of a uniform organisation of the sites. This analysis was carried out using the software program CrimeStat 3.2a

In fact, the method consists of the recording of a number of point locations (sites) within a circle with a research radius around a given site. This process is carried out for each site. The number of site locations by circle are added to those of the preceding circle. Once the exercise has been carried out for each site, the radius of the circle is increased and the whole process is repeated. The radiuses of the circles are slightly increased. Thus, on average, between 50 and 100 radiuses are taken into account. The K-function ‘ approximately follows an increasing exponential distribution and its expected value is increasing ‘ (Zaninetti, 2005). In 1977 Julian Besag proposed a modified version, and an easier reading as the expected value is zero, named the L-function.

This is tested by Monte-Carlo simulation of random distribution of the points within the study area. Each simulation may randomly generate a uniform or clustered model, etc. It is therefore necessary to repeat these simulations many times - at least a hundred times. They produce a ‘confidence interval under the null hypothesis of random distribution’ (Zaninetti, 2005) corresponding to the minimum and maximum values stemming from the simulations carried out for each distance. These results can be visualised on a correlogram, which makes it possible to put in relationship a variable with itself in time. The correlogram can be interpreted according to the position of the curve of the L-function with regard to the limits of the confidence interval, i.e. L MIN and L MAX. If the L-function is comprised within the interval, it can be accepted that the site distribution is random; if it sits above the interval, the distribution is clustered; and finally if it sits below, the distribution within the study area is uniform (figure 4).

This statistical method was tested for each chronological sub-phase initially on the whole corpus. It then became clear that it would be more relevant to analyse the sites by geomorphological zone – mountains and hills, plains – for reasons of different natural constraints which thus imply different manners of human management of space and landscape.

We were able to observe, for the sites on the plain, that the concentration of sites was only visible from a minimum scale of about 6kilometers, depending on the chronological period (figure 5). We could also see more marked peaks of concentration, the average spacing of which was calculated with a view to defining territories of influence around the sites (Boudry, 2013). In the mountains, the organisation of the sites is mostly random. This reflects the impossibility

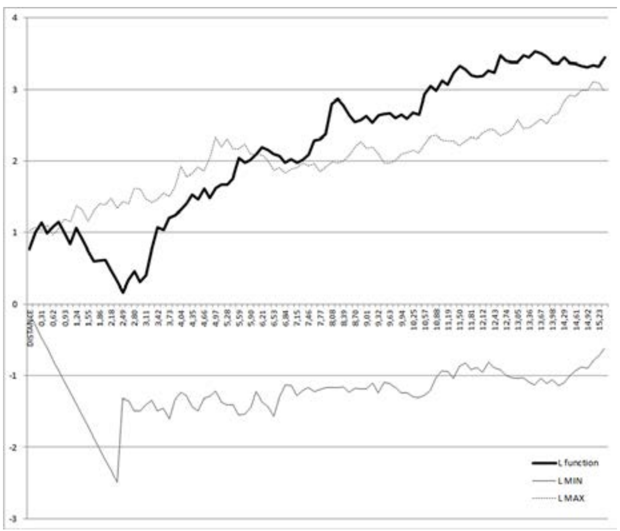


Figure 4: Correlogram based on Besag'sL-function of the Middle Bronze Age 3 sites established in the plain.

	Sites distribution on the whole area	Sites distribution on the plain area	Sites distribution on the hill and mountain area
Early Bronze	Random	Random	Random
Middle Bronze 1	Random	Random	Random
Middle Bronze 2	Random until 7 kilometers then peaks of aggregation each 5,77 kilometers	Random until 7 kilometers then peaks of aggregation each 4,44 kilometers	Random
Middle Bronze 3	Peaks of aggregation each 4,25 kilometers	Random until 6 kilometers then peaks of aggregation each 3,06 kilometers	Mostly random with little peaks of aggregation each 3,42 kilometers
Recent Bronze 1	Random until 3 kilometers then peaks of aggregation each 5,04 kilometers	Random until 5,36 kilometers then peaks of aggregation each 2,68 kilometers	Random
Recent Bronze 2 + Recent Bronze	Random until 3 kilometers then peaks of aggregation each 5 kilometers	Random until 7,19 kilometers then peaks of aggregation each 2,98 kilometers	Random until 3,58 kilometers then peaks of aggregation each 4,69 kilometers
Final Bronze	Random	Random	Mostly random

Figure 5:Results from Besag's L-function analysis.

of occupying this zone in a systematic manner due to physical constraints, except during key periods like Middle Bronze 3 and Late Bronze 2 with continuity into the Final Bronze Age.

The Kernel density

Smoothing by the Kernel method is an interpolation technique that makes it possible to generalise a local event to a whole area. It also makes it possible to create density maps with the aim of visualisingthe spatial distribution of archaeological sites and more particularly of areas of high density. Thus one can reveal zones that were preferentially occupied and see whether they evolved in time and space (Saligny *et al.* 2008; Di Salvo 2005).

This method was implemented using ArcGIS 9.2. The width of the radius of influence corresponds to a fixed interval defined by the mean distance between the peaks of aggregation stemming from the Ripley correlograms (Boudry, forthcoming).

The software program imposes the choice of the function, in this case a quadratic function. This function has the ability to attribute greater weight to the nearest points compared to the most distant. Its decrease is gradual but is interrupted when the total width of the radius of influence is reached. The density therefore equals zero for any point located out of the radius of influence of each point of the distribution (Zaninetti, 2005).

The spatial density was also analysed here separately for the two geomorphological zones. We were thus able to observe that throughout the Middle and Late Bronze Ages there were two settlement poles: in the north-west and the south-east of our study zone. An elevated position where the Apennine valleys open out onto the plain is also favoured. A period of change is notable in Late Bronze 1, when the eastern zone of settlements occupied for centuries are abandoned and new sites are founded in previously unoccupied zones. The sites located on openings onto the plain are also deserted in Late Bronze 1. In the Final Bronze Age, over 90% of plains sites have disappeared. In the mountains, occupation is less dense but remains stable until late Bronze 2, when one sees abandonments and new settlements. However, occupation becomes more durable in the Final Bronze Age.

2.2.2 What was the hierarchically oriented organisation between the sites?

In order to ascertain the hierarchically oriented organisation of this society, the Kernel density was again implemented but this time it was weighted by an index corresponding to the number of outstanding artefacts (themselves weighted according to their typology and their raw materials) subdivided by the area in m² of the site where they were discovered.

The results of this analysis were then interpreted and weighted in the light of the type of investigation carried

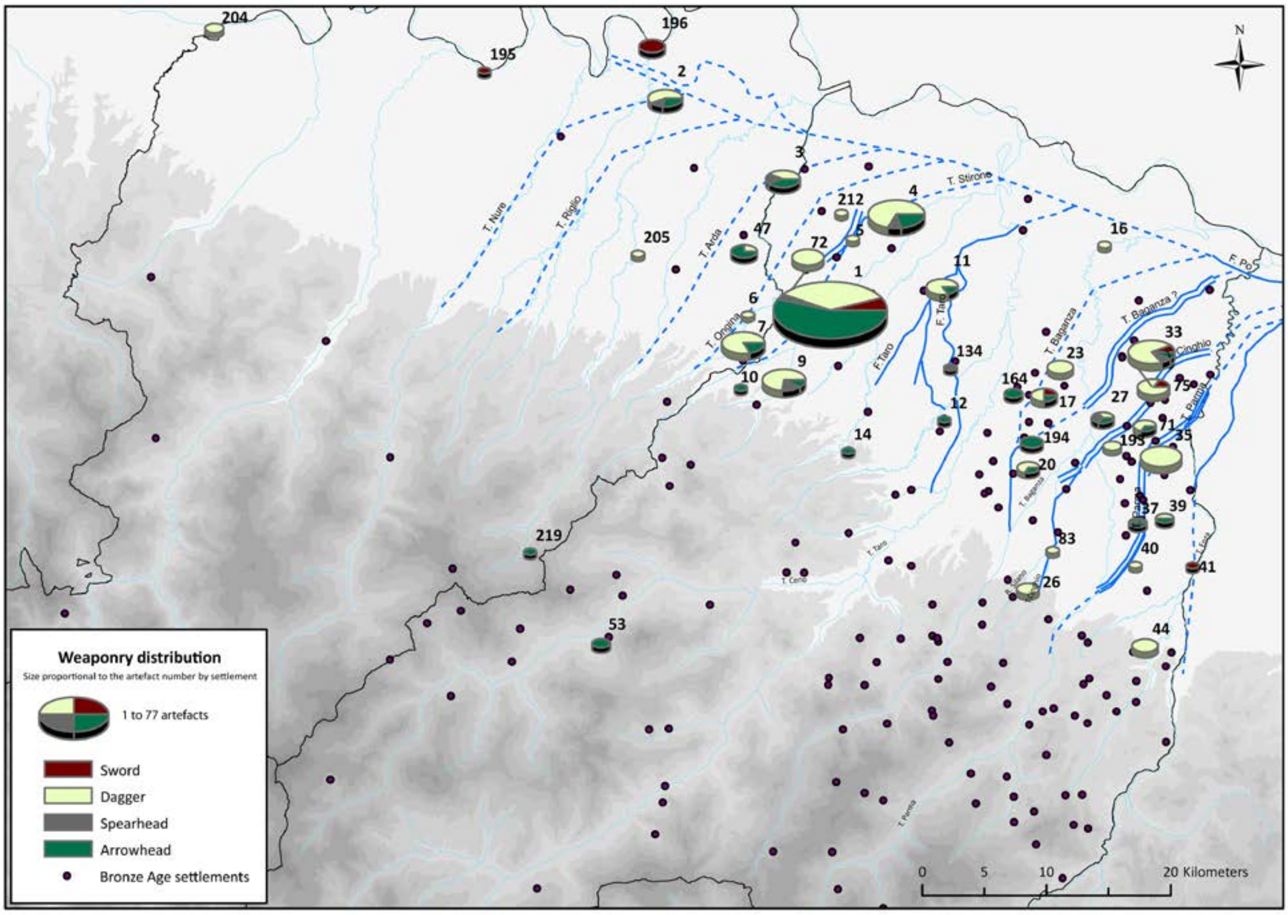


Figure 6: Spatial distribution of the weaponry in the terramare sites (the numbers refer to the site inventory).

out at the sites and of the period when they took place: excavations of the 20th or 19th centuries, trial trenches, surveys, exploitation as chalk mines, etc.

We were also interested in the spatial distribution of particular artefacts: tools linked with metallurgical production, tools related to textile crafts, weapons, personal ornaments, clothing accessories, votive or cult objects, imports as well as the spatial distribution of the settlement phases over the entire length of occupationand architectural characteristics (peripheral features and internal organisation, as for example houses on piles, and lastly the presence of redirected water courses). These elements were visualized using GIS thanks to a symbology by sectors. The choice of thissymbology enabled us to make a proportional estimation of the representativeness of distinct categories of objects by site (for example weapons) while also assessing the precise number of daggers or swords at the same site (Figure 6).

2.2.3 The particular case of mountain and hill sites

Selection of criteria

Most of the sites in mountain areas have been investigated by surface surveys only and they have yielded little information except for their cultural attribution and some details concerning their dating. The aim of this analysis therefore was to establish a new typology of the mountain sites looking at the settlement choices from an environmental perspective only. Several procedures using ArcGIS (unless otherwise specified) were applied to a digital terrain model of 10metres' resolution . It was possible to recover distinct informative raster data. The sites were projected onto these grids in order to gather the information related to their precise location. The criteria investigated are:

– Slope: This, combined with the exposure, will have a major effect on the thermic conditions of a given place. Two territories with differing slopes will receive different amounts of solar energy as a result of the angle at which the radiation reaches the soil.
– Accessibility: Most of the sites in the mountains were established on slight or moderate slopes; however, this does not exclude the choice of steep locations and those with reduced accessibility. In order to measure this parameter a buffer of 200 metres' radius was traced around the mountain sites on the slope raster. The number of slopes exceeding 35 % included in these buffer zones were then counted.
– Visibility: The viewshedis a tool that makes it possible, based on a DTM, to determine the raster surface visible at a given point. The analysis was carried out using Global Mapper v12.00. The height of the transmitter was fixed at 2 metres and the observation radius at 4kilometers . The visibility was calculated around each site.
– Exposure of the slopes: In the Apennine mountains the slopes which are the most exposed to the

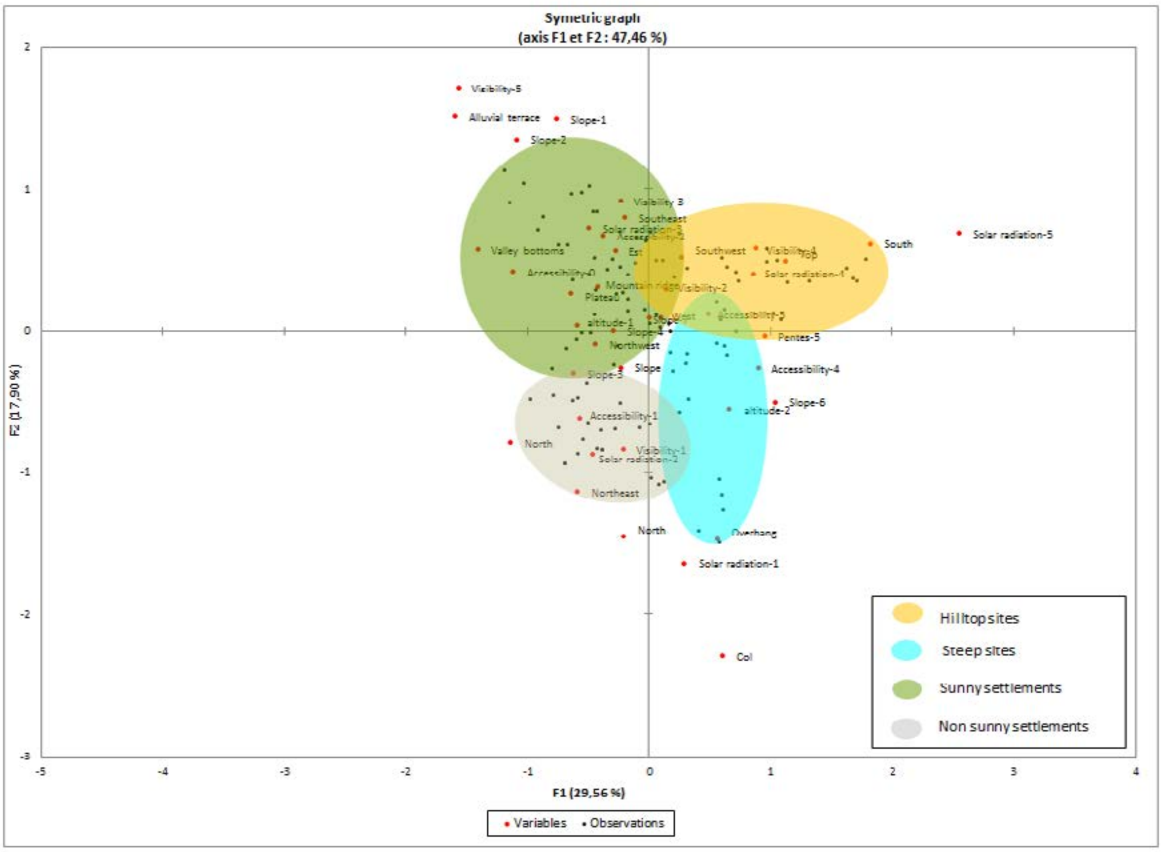


Figure 7: Factorial correspondence analysis of the mountain sites and the four groups defined.

sun are located in the south as they receive, for the most part,the most intense radiation, i.e. at midday,whereas slopesfacing north will receive radiation only in the morning and the evening (Lacoste, Salanon, 1999). This produces a major difference between south-facing and north-facing slopes, notably with regard to the type of vegetation on the slopes or to the lower limit of snowfall and its duration.

- Altitude: It was considered as zonation of mountain vegetation . Altitude indeed strongly influences fauna and, by extension, mountain economy. Three altitudinal zones were distinguished: the foothill zone between 200 and 800 metres, the mountain zone between 800 and 1750 metres, and the subalpine zone between 1750 and 2200 metres (Lacoste, Salanon, 1999).
- Solar radiation: The solar radiation was measured using the software program Saga and the tool ' Potential Incoming Solar Radiation '. A decision was taken to restrict the measurement to 21 March rather than to measure the radiation over the whole year in order to limit the period of analysis. The quality of sunshine recorded at this equinox can be interpreted as equalling the average value for a whole year.
- Geomorphological setting: This factor was recorded manually in the Access database according to the information available in the literature.

All these criteria were categorised into classes identified by numbers which made it possible to establish a contingency table required to carry out factorial analysis.

Factorial correspondence analysis

These different elements were then compared using multiple factorial correspondence analysis, a variant of factorial analysis , as this makes it possible to combine a large number of data. It thus offers a graphical display describing conjunction (or attraction), squaring (or independency) and opposition (or repulsion) (Cibois 2006) among the various characteristics of the elements under study.Widely used in archaeology (Chartier, 2005; Durand-Dastès et al., 1998; Gandini et al., 2008; Van Der Leeuw et al., 2003), the method was carried out here using XLSTAT. One hundred sites were analysed.

The MFCA shows a rather homogenous point cloud concentrated around the crossing point of the two main axes. The graphical display of the F1 and F2 axes, additionally interpreted in the light of a hierarchical ascending classification, highlights four groups bringing together a series of observations according to their close relationship with the variables (figure 7):

- The first group contains sites mostly established on a hill top, moderately steep, enjoying good visibility, exposed to the south, south-west or south-east and with good or very good sunshine exposure. Twenty-five sites belong to this group.
- The second group contains a few hilltop sites but consists mainly of spur sites. In these cases secondary significance only is accorded to sunshine and visibility, whereas the difficulty of access is key for these sites. These are sites in very steep locations, surrounded by or established on very steep slopes. Seventeen sites can be attributed to this type.
- The third group, in contrast to the preceding one, brings together sites of all geomorphological situations but first and foremost plateau sites and sites on mountain ridges with varying exposures that provide them with moderate sunshine. Within this group, a subgroup consists of sites located on alluvial terraces and in valley bottoms with rather good visibility. Thirty-two sites are included in this group.
- Finally, the last group, in contrast to the preceding types of sites, contains mainly sites established on north-facing or north-east-facing slopes with very low sunshine intensity and visibility. Twenty-five sites match this configuration.

On the basis of their characteristics these four groups can be interpreted as follows: the first group are hilltop sites ensuring territorial control, the second group are defensive sites;and the last two groups may are sites with functions related to agricultural practices and/or pasture.

3. Reading the results

The combined interpretation of all these analyses has provided us with a great deal of evidence for choices of settlement location in the Terramare culture, as well as for its territorial management. Furthermore, some propositions can be outlined with regard to societal organisation.

From this perspective the incorporation of the most recently discovered Early Bronze Age sites into this study enabled us to examine the emergence of this culture. It was thus possible to challenge the idea of territories that would have been void of initial occupation and to assume a local contribution to the development of the Terramare culture, even if the arrival of new populations seems to be undeniable with regard to demographic growth.

Three areas could be highlighted with differing territorial management (figure 10). These areas were distinguished by varying settlement densities and were apparently initially conditioned by distinct geomorphological elements, but not only by these. Thus the wide Taro River forms a boundary in the plain as well as in the mountains. The transition between the plain and the mountains was apparently not a limiting factor and depends on the configuration of the water courses(Boudry, forthcoming). These zones communicate along fluvial and terrestrial axes.

For the first area, called the eastern territory, a regular distribution of the hilltop sites along the valleys could be shown, using the established typology. These hilltop

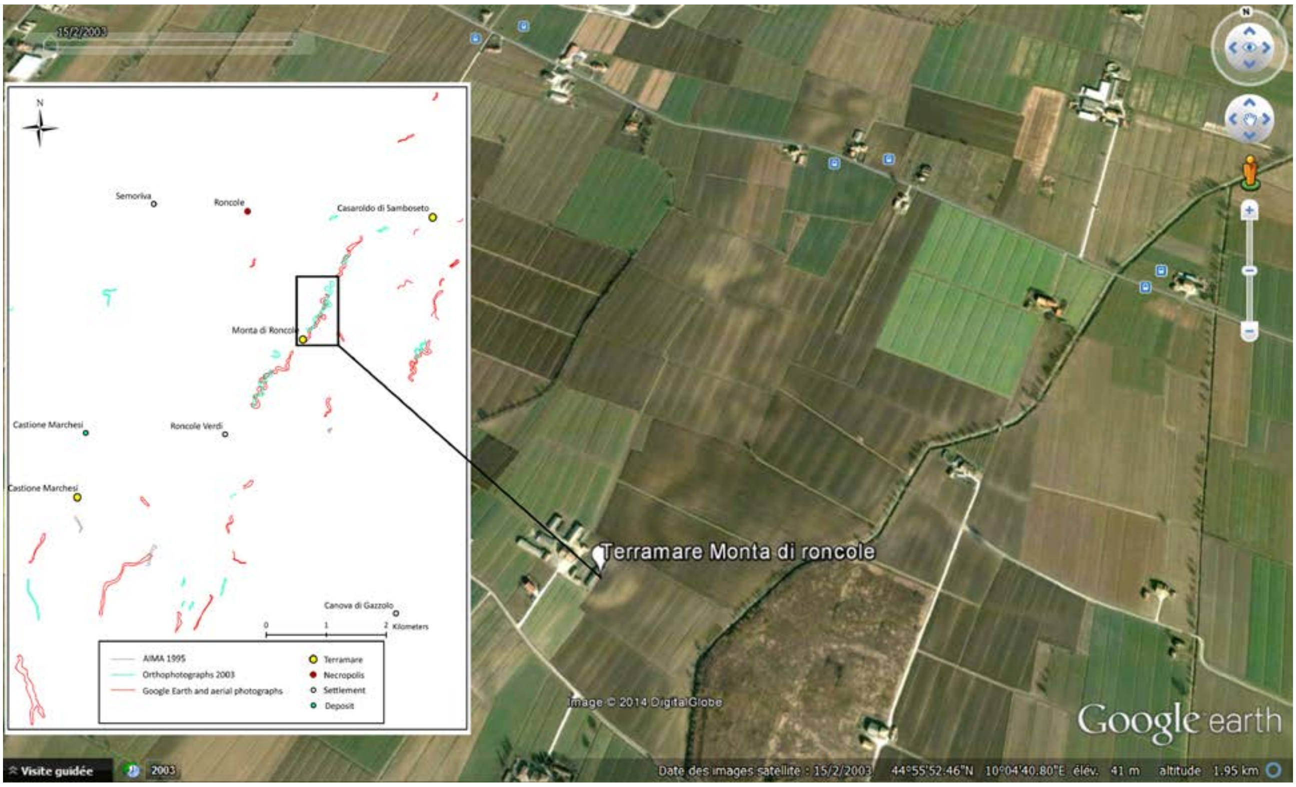


Figure 8: Trace of paleochannel, supposed to have been active during the Bronze Age, in relationship with the terramare site of Monta di Roncole (Google Earth V 7.1.2.2041. (15 February 2003). Monta di Roncole, Italie. 44°55'52.46'N, 10°04'40.80'E, altitude: 1, 95 kilometer. DigitalGlobe 2014).

sites have a very wide view over the fluvial axes and their valleys. The distribution of sites along the rivers continues in the plain with the terramare settlements. In different geomorphological spaces we thus have sites functioning within a network along the fluvial axes.

The second area, the western territory, presents the same organisation as that of the eastern plain, with a lower density. A terramare network functioning concomitantly to a paleochannel (belonging most probably to the Stirone River) has very clearly revealed here (figure 8). This channel is thought to have also connected the terramare sites of Monta di Roncole and Casaroldo di Samboseto as well as the Roncole Verdi site.

The existence of a route predating the Via Emilia, the Roman route connecting Piacenza with Rimini on the Adriatic coast, has been assumed based on the observation of importsof Adriatic origin. It may have connected the Adriatic coast with the terramare region but also the two zones of plain between them. Other terrestrial axes have also been proposed, again based on the distribution of imports, possibly connecting the Via Emilia to the Po River, another major fluvial axis.

Finally, a third zone, called no man's land, differs in its organisation from that of the eastern mountain area. It is not connected with the plain area, a factor that could be explained by the absence of a fluvial axis connecting these two areas. The western water courses indeed originate in the hills or are tributaries of the Taro River and there are no valleys connecting the mountains with the plain.

The interpretation of our hierarchisation index (number of ostentatious artefacts / surface of the site in m²) associated with the duration of occupation of the sites enabled us to highlight eight centres with concentrations of wealth (figure 10). Except for a single case, none of the terramare are located within the territories of influence proposed for these sites, due to the results from Ripley'sK-analysis and the calculation of Kernel density. Each terramare site therefore seems to be independent, albeit exhibiting varying levels of wealth. As we have seen, the terramare sites are frequently in relation with several important exchange axes, both fluvial and terrestrial, which suggests the existence of social elites. The centres are particularly closely located, on average 6 kilometresapart in the eastern territory and 7.5 kilometres apart in the western

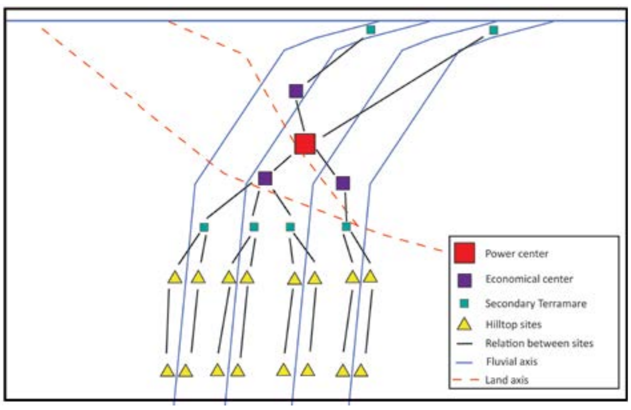


Figure 9: Drawing of a confederation applied to the Terramare culture.

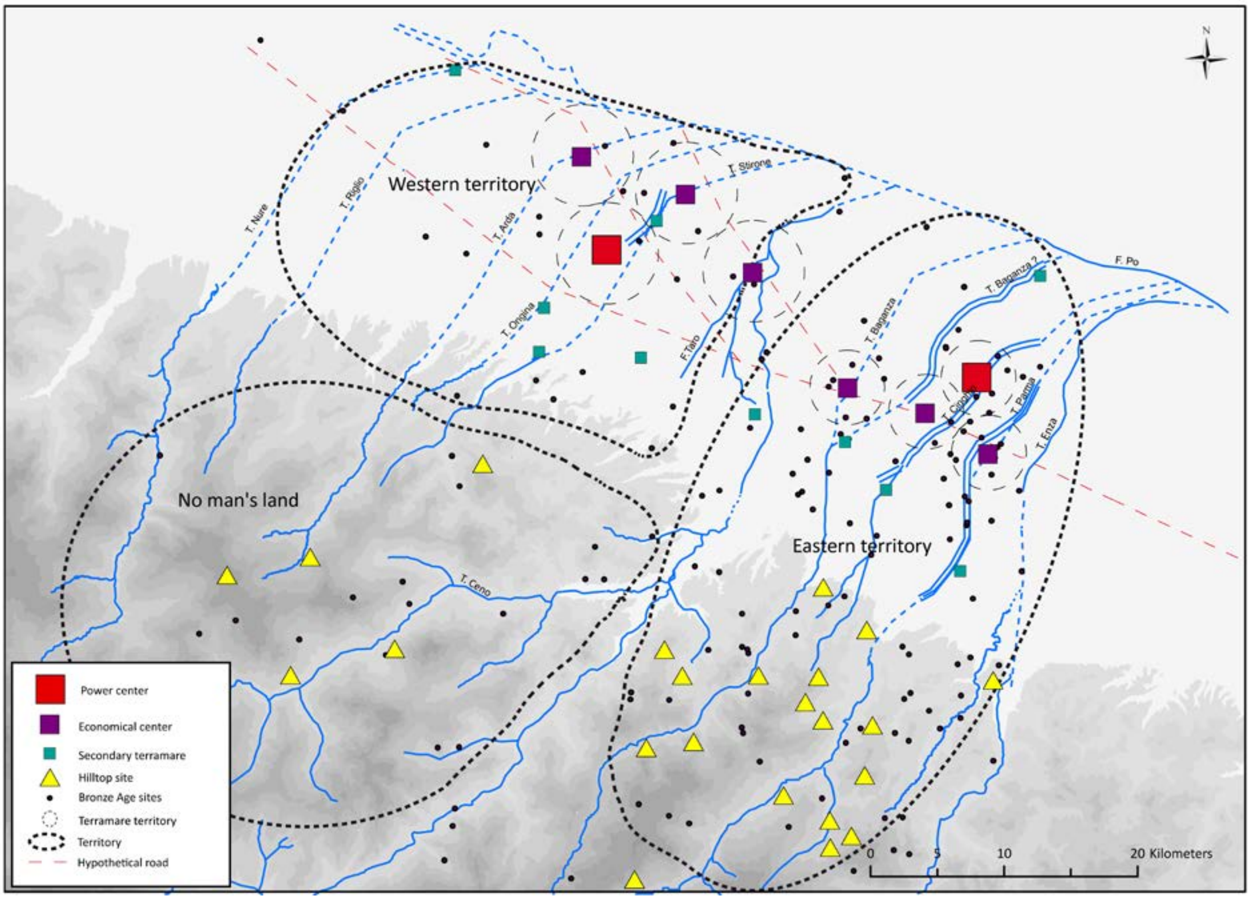


Figure 10: Confederative organisation in confederation of the terramarewith the three territories highlighted as well as the influence territory for each terramare.

territory. This proximity necessarily implies good relations between these elites. They may have been organised in a form of confederation (figure 9). The term confederation would suppose that each terramare site preserved its independence with regard to a large number of fields, and more particularly in terms of economy. It is also possible that the sites were subjected to one central site concentrating political and/or religious power.

On either side of the boundary formed by the Taro River, two sites are distinguished by their exceptional level of wealth (CastioneMarchesi and FornodelGallo) (figure 10). They may have served as a centre, even if in the current state of research, only the signs of greater wealth supports this hypothesis. It may be possible that the secondary terramare sites played the role of outposts within a complex network maintaining relationships with the mountain sites for the terramare located in the upper plain or controlling the Po River for the sites in the lower plain.

Ultimately, a confederation system would not exclude large-scale communal construction work, as for example the enclosing features (Cardarelli, 1988) or the construction of irrigation and drainage systems.

The analyses carried out for the mountain sites did not reveal a hierarchy between them, most certainly because of major environmental constraints. Apparently each site had a precise and well-defined role (control sites, agricultural sites, livestock enclosures).

With regard to the demographic crises that the Terramare culture underwent in the mid-12th century BC, the first signs indicating disturbance of the established system appear as early as the Late Bronze Age1 (1340/1330 – 1230/1200 BC): the number of sites in the plain begins to decrease; new sites are founded in new areas. Our interpretation is a less sudden crisis than previously suggested by Italian researchers. Thuswe would see a crisis emerging at the beginning of the Late Bronze Age - in other words one century prior to the generalised collapse. It is interesting to note that the beginning of the Late Bronze Age 1 (1340/1330 BC) coincides with the return of less humid conditions in Italy (Magny *et al.*, 2007).

At the end of the Late Bronze Age, surviving settlements are present in the eastern territory and more particularly in the Stirone and Baganza valleys. This collapse is therefore uneven and unequally affects the territories defined here. We may deduce that the crisis wasmanaged in different ways by the elites of these two territories.

4. Conclusion

This comprehensive approach (figure 11) was intended, firstly, to reconstruct the Bronze Age water network. This reconstruction provides a dynamic and systemic vision of the Terramare culture. In fact previous archaeological analyses have too often been based on the current water network and this can introduce errors in interpretation, at least on the plain.

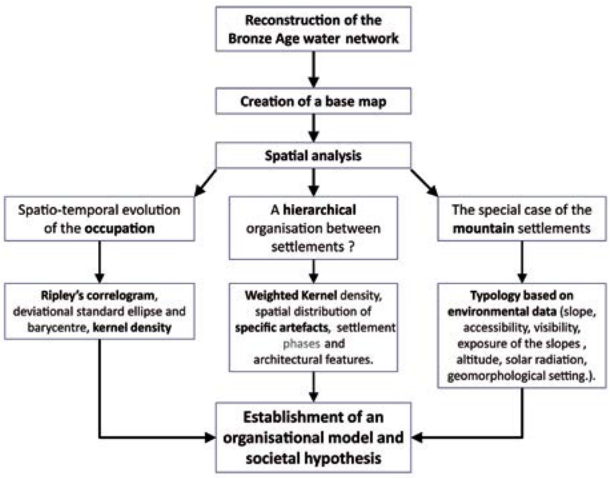


Figure 11: Diagram of the comprehensive approach of the study.

The water network was then incorporated into a base map serving as a support for all the spatial analyses carried out on the archaeological corpus. Lastly, the spatial and geomorphological analyses were interpreted together, leading both to the creation of an organisational model and to propositions about society.

This original approach, combining statistics and spatial analyses of the archaeological corpus in its environmental context, has shed new light on the Terramare culture, notably by:

- underlining the importance of local contributions in its formation
- revealing new cases of direct relations between terramare sites and water courses, as well as showing fluvial or terrestrial circulation routes that structure and connect several distinct geomorphological zones;
- identifying several territories, interconnected yet managed in different ways, thus indicating decisions taken locally;
- proposing a model of the Terramare phenomenon, with societal interpretations - reconsidering the suddenness of the crisis and its supposed homogeneity.

Translation: Karoline Mazurié de Keroualin and Michael Ilett.

References

BALISTA, C., DE GUIO, A., VANZETTI, A., BETTO, A., DE ANGELI, G., SARTOR, F., NICOSIA, C., MARCHESINI, M., and MARVELLI, S. (2004). Paleoidrografie, impianti terramaricoli e strade su argine : evoluzione paleoambientale, dinamiche insediative e organizzazione territoriale nelle Valli Grandi Veronesi alla fine dell'età del Bronzo. *Padusa*. XL. p. 97-138.

BENZECRI, J.-P. (1992). *Correspondence Analysis Handbook*. New York: M. Dekker.

BERNABÒ BREA, M. (in press) L'Età del bronzo. In: *Preistoria e Protostoria dell'Emilia-Romagna. Atti della XLV Riunione Scientifica dell' Istituto Italiano di Preistoria e Protostoria*, Modena, 27-31 Ottobre 2010. Firenze: Istituto Italiano di Preistoria e Protostoria.

BERNABÒ BREA, M., CARDARELLI, A., CREMASCHI, M. (eds) (1997). *Le Terramare: la piùa nticaciviltàpadana:esposizione, ForoBoario, Modena, 15 marzo – 1 giugno 1997*. Milano: Electa.

BESAG, J. (1977). Contribution to the discussion of Dr Ripley's paper. *Journal of the royal statistical society.* B39. pp. 193-195.

BEVAN, A., CONOLLY, J. (2006). Multiscalar Approaches to Settlement Pattern Analysis. In: Lock, G.R., Molyneaux, B.L. (eds). *Confronting scale in archaeology : issues of theory and practice*, New York: Springer, pp. 217-234.

BOUDRY, J. (2013). Vers une organisation hiérarchisée des sites de la culture des Terramares (Italie, XVIe-XIIe siècles av. n. è.). In: Brunet, O., Sauvin, C., Al Halabi, T. (eds) *Les marqueurs archéologiques du pouvoir.* Actes de la 4e Journée doctorale d'archéologie, Paris, 27 Mai 2009. Archéo.doct, 4. Paris : Publications de la Sorbonne. pp. 243-258.

BOUDRY, J. (forthcoming). *Implantation territoriale des terramares (Italie, provinces de Parme et Plaisance, XVIIe-XIIe siècles av. n. ère). Analyses géomorphologiques et spatiales*. Oxford: Archaeopress.

CARDARELLI, A. (1988) L'età del Bronzo: organizzazione del territorio, forme economiche, strutturesociali. In: *Modena dalleoriginiall'anno mille: studi di archeologia e storia: esposizione*, Galleria Civica, Modena, gennaio-giugno 1989. Volume I. Modena: Edizioni Panini. p. 86-127.

CARDARELLI, A. (2009). The collapse of the Terramare culture and growth of new economic and social systems during the late Bronze Age in Italy. In: *Le ragioni del cambiamento.* Atti del convegno internazionale, Roma, 15-17 Giugno 2006. Scienze dell'Antichità, 15. Rome: Quasar. pp. 449-520.

CASTALDINI, D. (1989). Evoluzione della rete idrografia centropadana in epoca protostorica e storica. In: *Insediamenti e viabilità nell'alto ferrarese dall'età Romana al Medioevo. Atti del convegno Nazionale di Studi, Cento, 8-9 Maggio 1987.* Atti dell'Accademia delle Scienze di Ferrara, 64, supplementi. p. 115-134.

CASTIGLIONI, G.B., BIANCOTTI, A., BONDESAN, M., CORTEMIGLIA, G.C., ELMI, C., FAVERO, V., GASPERI, G., MARCHETII, G., OROMBELLI, G., PELLEGRINI, G.B., TELLINI, C. (1999)

.Geomorphological map of the Po plain, Italy, at a scale of 1:250 000. *Earth Surface Processes and Landforms*. 24, pp. 1115-1120.

CHARTIER, M. (2005). L'analyse statistique des implantations des sites du Néolithique Ancien et Moyen dans le Bassin Parisien : l'apport de quelques critères spatiaux. *Bulletin de la Société Préhistorique Française*. t.102, n°2. p.269-280.

CIBOIS, P. (2006). Principe de l'analyse factorielle. Available from: http://cibois.pagespersoorange.fr/PrincipeAnalyseFactorielle.pdf. [Accessed: 13th December 2011].

CREMASCHI, M. (2009). Ambiente, clima ed uso del suolo nella crisi della culture delle terramare. In: *Le ragioni del cambiamento*. Atti del convegno internazionale, Roma, 15-17 Giugno 2006. Scienze dell'Antichità, 15. Rome: Quasar. p. 521-534.

CREMASCHI, M., BERNABÒ BREA, M., TIRABASSI, J., D'AGOSTINI, A., DALL'AGLIO, P.L., MAGRI, S., BARICCHI, W., MARCHESINI, A., NEPOTI, S. (1980). L'evoluzione della pianura emiliana durante l'età del bronzo, l'etàromana e l'alto medioevo: geomorfologia ed insediamenti. *Padusa*. XVI. pp. 53-161.

CREMASCHI, M., PIZZI, C., VALSECCHI, V. (2007) .Gestion des eaux et du territoire et éventualité d'un changement environnemental dans les terramares à la fin du Bronze récent en Italie du Nord. Le cas de la terramare de Santa Rosa (plaine du Pô, Italie du Nord). In: Richard, H., Magny, M., Mordant, C. (eds) *Environnements et cultures à l'âge du Bronze en Europe occidentale*. Actes du 129e Congrès national des sociétés historiques et scientifiques.Besançon, 2004. Documents Préhistoriques, 21. Paris : édition du Comité des travaux historiques et scientifiques, pp. 327-342.

DE MARINIS, R.C. (1997). L'età el Bronzo nella regione benacense e nella pianura padana a nord del Po. In : Bernabò Brea, M., Cardarelli, A., Cremaschi, M. (eds) *Le Terramare: la più anticaciviltàpadana: esposizione,* ForoBoario, Modena, 15 marzo – 1 giugno 1997. Milano: Electa. p. 405-422.

DI SALVO, M., GADAIS, M., ROCHE-WOILLEZ, G. (2005). Programme ACTEUR : L'Estimation du noyau de densité : méthode et outils. Ministère de l'Equipement, Certu. Availablefrom: http://lara.inist.fr/bitstream:2332/879/4/0949-T1.pdf [Accessed: 15th April 2010].

DURAND-DASTES, F., FAVORY, F., FICHES, J., MATHIAN, H., PUMAIN, D., RAYNAUD, C., SANDERS, L., VAN DER LEEUW, S. (1998). *Des oppida aux métropoles, Archéologues et géographes en vallée du Rhône.* Collection 'Villes'. Paris : Anthropos.

FERRI, R. (1996). La fotografiaaerea in zone di bassa pianura : iconografiaanalitica delle traccefluvialiedevoluzioneidrografica delle Valli Grandi Veronesi. In: La ricercaarcheologica di superficie in area padana. Atti del workshop. Villadose 1 ottobre 1994. Padova: AGS Ed., p. 25-33.

GANDINI, C., F. BERTONCELLO, E. GAUTHIER, L. NUNINGER AND F. TRÉMENT (2008). Hierarchical Typology and Settlement Patterns Modelling at Interregional Scale. In: Posluschny, A., K. Lambers & I. Herzog (eds.) *Layers of Perception. Proceedings of the 35th International Conference on Computer Applications and Quantitative Methods in Archaeology (CAA)*, Berlin, Germany, April 2–6, 2007. KolloquienzurVor- und Frühgeschichte. Vol. 10. [CD-ROM]. Bonn :Habelt.

LACOSTE, A., SALANON, R. (1999). *Éléments de biogéographie et d'écologie.* 2nd Edition. Paris: Nathan.

MAGNY, M., BOSSUET, G., GAUTHIER, E., RICHARD, H., VANNIÈRE, B., BILLAUD, Y., MARGUET, A., MOUTHON, J. (2007) Variations du climat pendant l'âge du Bronze au centre-ouest de l'Europe : vers l'établissement d'une chronologie à haute résolution. In: Richard, H., Magny, M., Mordant, C. (eds) *Environnements et cultures à l'âge du Bronze en Europe occidentale*. Actes du 129e Congrès national des sociétés historiques et scientifiques, Besançon, 2004. Documents Préhistoriques, 21. Paris : édition du Comité des travaux historiques et scientifiques. p. 13-28.

PICCARRETA, F., CERAUDO, G. (2000). *Manuale di aerofotografia archeologica*. Metodologia, tecniche e applicazioni. Bari: Edipuglia.

SALIGNY, L., NUNINGER, L., OSTIR, K., POIRIER, N., FOVET, E., GANDINI, C., GAUTHIER, E., KORALI, Z., TOLLE, F. (2008) Models and tools for territorial dynamic studies. In: *ArchaeDyn, 7 Millennia of territorial dynamics, settlement pattern, production and trades from Neolithic to Middle Ages*. ACI 'Spaces and territories'. Final conference, Dijon, 23-25 june 2008, Preprints. pp.25-44.

VAN DER LEEUW, S., FAVORY, F., FICHES, J. (eds) (2003). *Archéologieetsystèmessocioenvironnementaux. Etudes multiscalaires sur la vallée du Rhône dans le programme ARCHAEOMEDES*. Monographie du CRA, 27. Paris: Editions du CNRS.

ZANINETTI, J.-M. (2005). *Statistique spatiale, méthodes et applications géomatiques*. Collection application des SIG. Paris: Hermès science publ.

From the Excavation to the Territory: Contributions of GIS Tools to the Study of the Spatial Organization of the Archaeological Site of Argentomagus (France, Indre, Saint-Marcel/Argenton-sur-Creuse)

Emeline LE GOFF,[1] Laure LAÜT,[2] Yoann RABASTÉ,[3] Françoise DUMASY[4]

[1]Association pour la Sauvegarde du Site archéologique d'Argentomagus et Amis du Musée, Saint-Marcel, France
[2]Université Paris 1 Panthéon Sorbonne, Paris, France,
[3]INRAP - Institut National de Recherches Archéologiques Préventives - Base Grand Est Nord, Reims, France
[4]Université Paris 1 Panthéon Sorbonne, Paris, France
legoff.emeline@laposte.net, laure.laut-taccoen@univ-paris1.fr, yoann.rabaste@inrap.fr, francoise.dumasy@wanadoo.fr

Abstract

The site of Argentomagus is located in the Midwest of France, where a total of 120 archaeological sites have been identified, distributed on a surface of 20km². Over centuries, the territory of Argentomagus had been the place of an oppidum, of an antique secondary urban area, and then of a medieval and a modern village. A Geographical Information System (GIS) was used to homogenise, to compile and to analyse all the sites identified since the 1960s.

This paper presents the methods used for the study of this antique city within all its spatial extension and also within its chronological thickness. Among others, this approach aimed at defining the plan of the Celtic rampart more precisely. Moreover, it also concerns the scale of the whole territory and the analyses related to the implantation of sites on the relief: analyses of visibility, especially for the construction of buildings, but also analyses of distances.

Keywords: Argentomagus, geodatabase, spatial organization, digital terrain models, GIS tools.

1. Introduction. Spatial approach of the ancient site of *Argentomagus*

1.1. Geographical and historical contexts

Argentomagus is located in what is now the Region Centre in France, and extends across two little towns of the *Indre* department – *Saint-Marcel* and *Argenton-sur-Creuse*, about 300 kilometers south of Paris (Fig. 1). It is an important archaeological site, known since the 16th century and explored by archaeologists since the 19th century; nowadays, regular surveys and excavations are carried out and a museum has been created to showcase part of the remains.

The site is on a sharp bend of the *Creuse*, a river which cuts deeply into the limestone plateau. The hilly terrain and proximity to water have made it attractive for human occupation (Fig. 2). During prehistoric times, the presence of Magdalenian hunter-gatherers in the cave of *La Garenne* is the first important episode of the history of the site. Later, occupation was concentrated slightly further east, on the plateau of *Les Mersans*, a naturally protected area which provided a valuable vantage point for the valley. A Celtic *oppidum* and then a Roman town thus developed on this plateau. At the start of the Roman period, *Argentomagus* was a town situated in the north of the province of Aquitaine, in the territory of the *civitas* of the *Bituriges Cubi*, whose capital was Bourges (*Avaricum*). Covering an area of 18,000 km², this territory was one of the largest in Gaul, with a dense network of towns and of river and land routes. *Argentomagus* was one of the largest secondary towns, in the south-west part of this ancient *civitas* (Laüt, 2013). During the Middle Ages, the settlement moved west and south, to the sites of the present-day village of

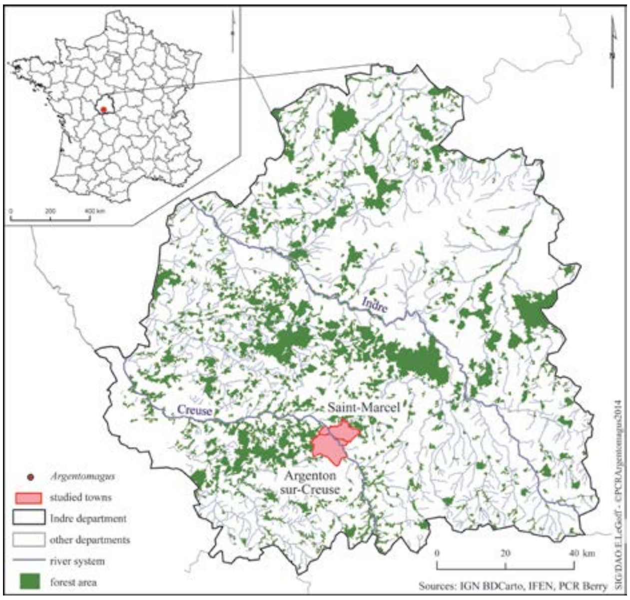


Figure 1: The archaeological site of *Argentomagus* at its present state.

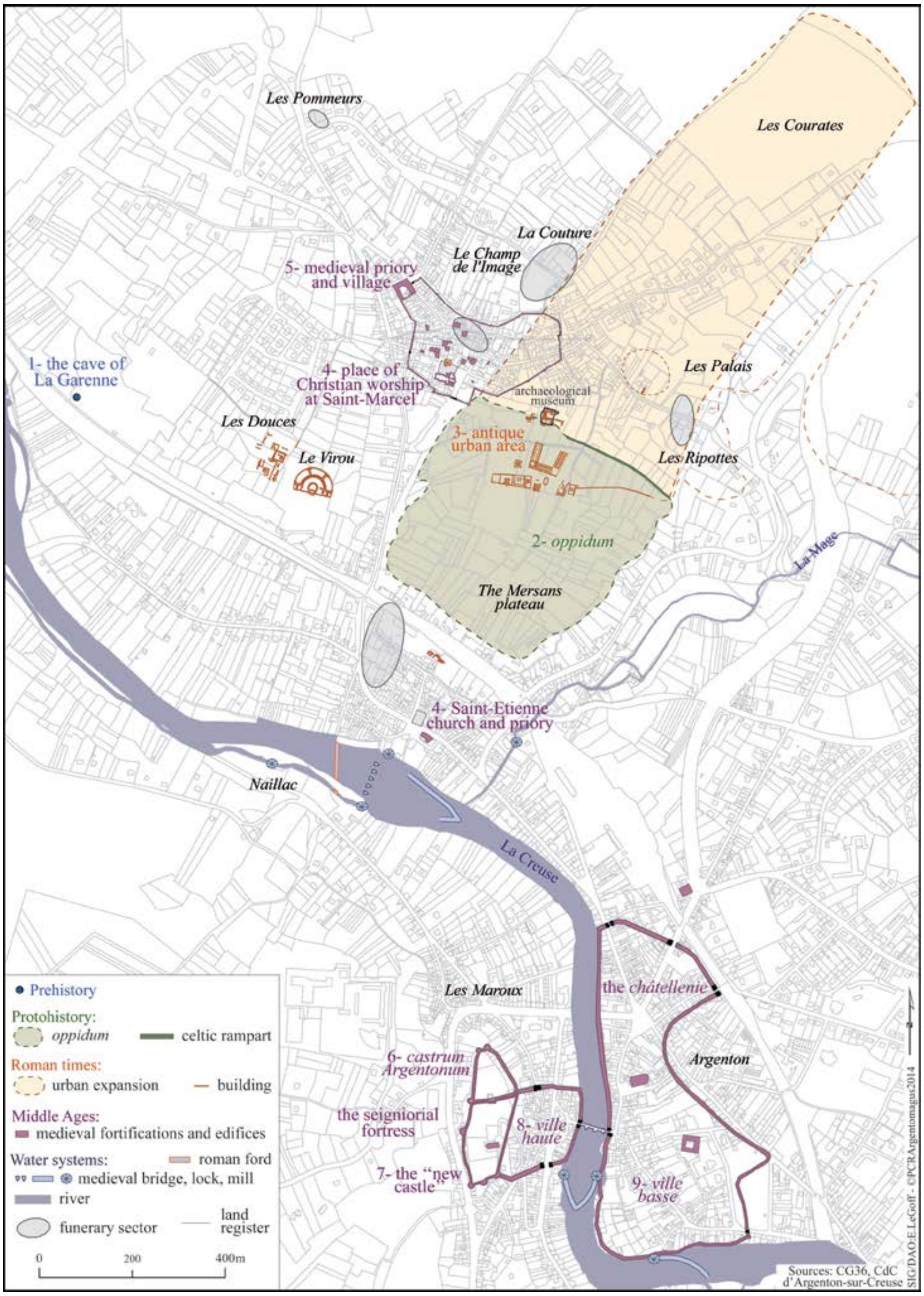


Figure 2: Chronological map of *Argentomagus*: a site located on a hill overlooking the *Creuse* valley.

Saint-Marcel and the small town of *Argenton-sur-Creuse*. This sector is particularly rich in heritage and has been the subject of a large number of archaeological operations, particularly in the last 30 years (Dumasy, 2013).

1.2. Concerns of the Collective Research Program 'Argentomagus'

In 2004, a Collaborative Research Project on the *Argentomagus* site was set up, with the principle aim of investigating the Gallo-Roman period of the site, bringing together nearly 30 researchers from different fields. To deal with the large quantity of data to be processed, the members of the project decided to create in 2010 an archaeological map using GIS, incorporating geo-referenced data from various sources and increasing possibilities of spatial analysis on site (Le Goff, 2013).

The Collective Research Program is organised around two themes: the spatial distribution and the functional characterisation of archaeological remains on one hand; the relationships between a site and its environment on the other hand. The first research theme focuses on a diachronic perspective at scale of the intra-site. Requests from the spatial database permit the elaboration of time-phase maps offering a renewed vision of the successive transformations of this territory, especially about the evolution of the antique urban organisation. The second axis deals with the scale of the whole territory and the analyses related to the implantation of sites on the relief: analyses of visibility, especially for the construction of buildings, but also analyses of distances. Among others, this approach aimed at defining the plan of the Celtic rampart more precisely. Moreover, hydro-morphological

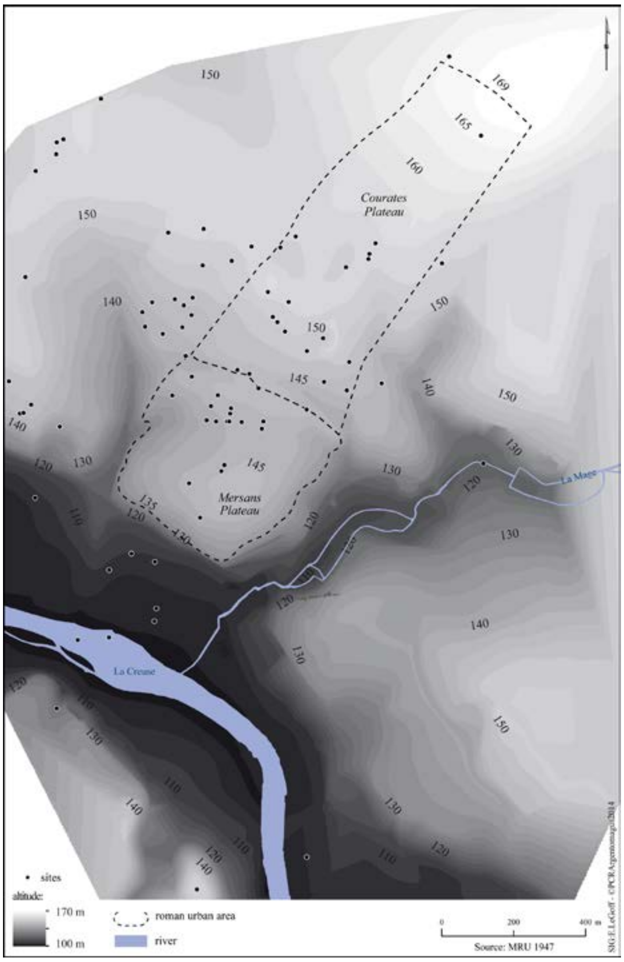


Figure 3: For local scale (from 1/10,000 to 1/1,000): the DTM MRU 1947.

analyses were realized in order to validate the hypothesis of the anthropogenic planning of a former thalweg.

2. GIS architecture

2.1. Integration of the mapping data sources

The archaeological map of *Argentomagus* was designed first and foremost as an on-site research aid, but it is also a valuable tool for local heritage management and for setting up urban development projects in the short, medium and long term. All of which involves specific presentations of archaeological remains on different base maps.

Currently, three main types of document are used: the digital land registry map of the *Argenton-sur-Creuse* district, which shows the borders, ownership and place names of the current plots of land; The digitized file of the river network of the district, which is more precise than the network of the IGN (the national geographic institute in France) topographic map; several Digital Terrain Models (DTM) to gain a good understanding of the relief of the sector.

In fact, at the regional scale, this is the IGN® BD Alti, with a resolution of 50m, obtained during a previous Collaborative Research Project, for the whole *Biturige* territory. At the micro-regional scale, we use a DTM with 1m resolution (Fig. 3), created from a 1947 topographic map produced by the Ministry for Reconstruction and Urbanism (MRU); it covers the whole area of the Roman town and provides a fairly accurate view of the overall topograhy of the site.

At the scale of the excavation site, we have a DTM with 20cm resolution (Fig. 4), created during a microtopography operation on the *Mersans* plateau in 2013 (Le Goff 2014); it covers only 10.6 hectares, in other words 1/6 of the maximum area of the ancient town, but for that zone, it provides details about the layout of remains within their topographic and hydrographic environment.

2.2. Geo-referencing of archaeological data

For the geo-referencing of archaeological data, we took the results of all the studies carried out on the site over at least the last fifty years. Various methods were used in order to cover all the area (Dumasy, 2013: 23). Since the 1960s, planned excavations focused on the *Mersans* plateau and the theater, with a total surface area of 2 hectares. Preventive archaeological operations (diagnosis and excavations) carried out by INRAP mainly concerned the funerary sectors on the outskirts of the ancient town, with a total surface area of 9 ha. Ground prospecting was carried out since 1988 on arable land in the parish of *Saint-Marcel* and along the D927 road. This revealed the extension of the town to the north, in the sector of *Les Courates*, and allowed surface exploration of an area of about 63 hectares. Geophysical surveys carried out by quad bike since 2006 covered 10 hectares of ground, revealing the network of streets delimiting housing blocks, a public square with a basilica and shops, and districts for housing or artisanal activities. These different sources and the details that they provide are clearly complementary and occasionally overlap. It is indeed GIS cartography that makes it possible to visualize these different features simultaneously, improving our capacity to interpret archaeological data. In 2004, the first archaeological map was produced using CAD, with Adobe Illustrator. We started by georeferencing this data with GIS, creating a spatial database. Further information was fed into the base, regarding old digs and historical data from mediaeval and modern periods. The data were homogenized by defining common class attributes and by creating lists of values to characterize remains, facts, structures, sites and periods. To date, 145 archaeological entities have been identified in a 20 km² area, covering all periods.

2.3. Architecture of the Argentomagus *GIS*

The *Argentomagus* spatial database can be divided into three broad categories of information:

- ‘Framework data’, bringing together layers of information linked to the current land register, to place names, to hydrography, to relief, and to the area covered by the archaeological interventions.
- ‘Field data’, involving files linked to former, recent and current excavations, now recorded in real time

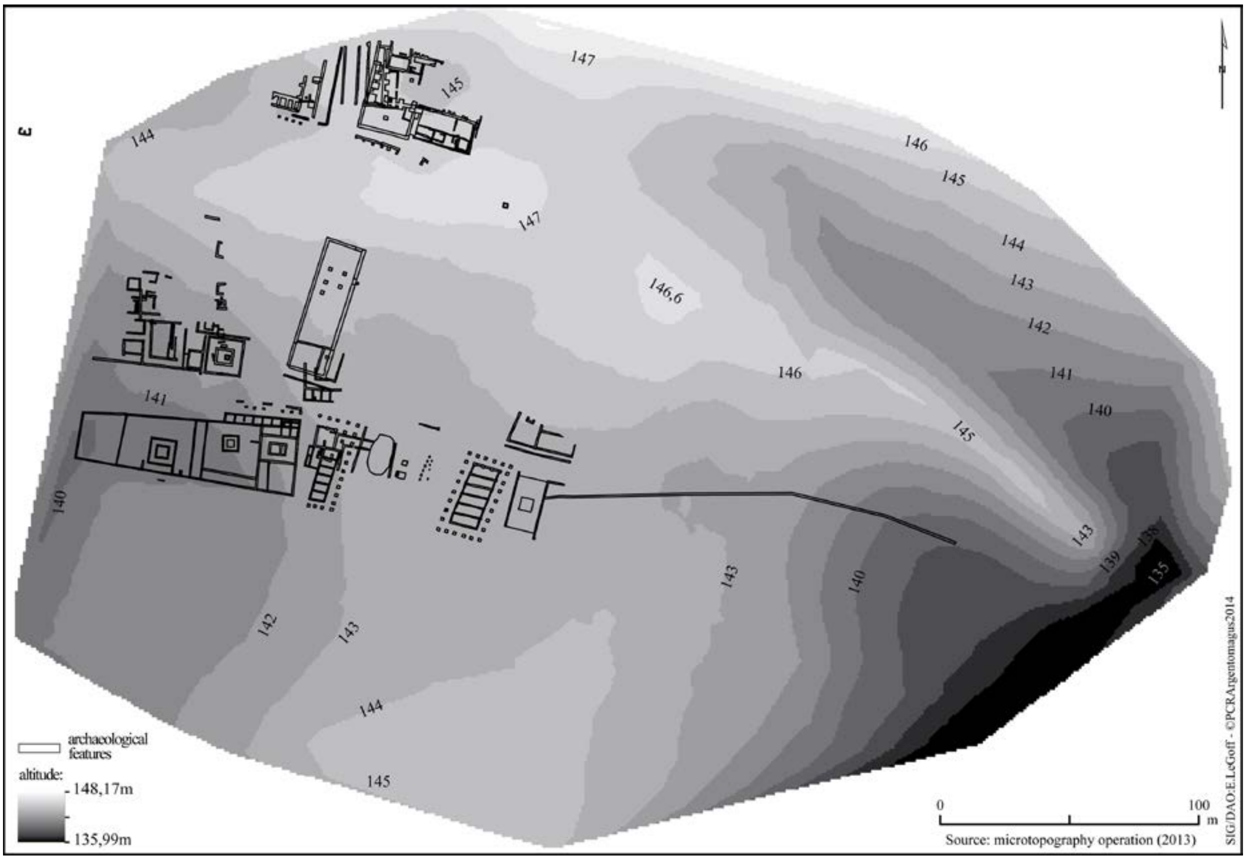


Figure 4: For excavation scale (from 1/1,000 to 1/100): the DTM 2013.

on the site. This could involve the delimitations of the excavation sector, features, or isolates. We thus use plans of the excavations sites on a scale of 1:100 to monitor how the site develops. Archaeological structures are drawn directly in the GIS from topographical points recorded by tacheometry.

– 'Interpretation of structures' data involve all the archaeological entities recorded by the SRA (regional archaeology department) and series showing functional coherence: buildings, hydraulic structures, fortifications, funerary sectors and structures, and road systems. In this way, plans of the site on a scale of 1:10,000 can be produced rapidly.

3. Chronological Maps

3.1. From 1st c. B.C. to 3rd c. A.D.

GIS mapping of the remains identified in *Saint-Marcel* and *Argenton-sur-Creuse* has enabled us to identify the spatio-temporal evolution of occupation in this sector, from Prehistory to the present (Fig. 2). In this paper, just key moments will be presented, on which the following spatial analyses focused.

In the years 100 to 80 B.C., a Celtic *oppidum* was established on the *Mersans* plateau (Krausz, 1998; Dumasy, 2013: 30-39). This site is 35 m above the *Creuse* valley, protected naturally on three sides and by a rampart on the north. On the western flank, the wide cut in the relief which extends into the plateau could have provided one means of access to the site. The internal area of this *oppidum* is estimated to be 40 hectares. Life within this stronghold is still relatively unknown, but the artifacts that have been collected indicate that *Argentomagus* was a production and trade centre, with a well-developed road network.

The Roman town developed during the 1st century A.D. (Dumasy, 2013: 38-41). All the buildings and funerary enclosures lined a network of streets that were oriented North-South/East-West. From 50 A.D., the earth and timber houses were replaced by masonry and ashlar buildings. Public buildings appeared: temples in the town centre, the *domus* of a wealthy person, porticos, and, to the west of the former *oppidum*, a theatre, probably already dominated by a sanctuary. At the beginning of the 2nd century, the town centre was largely modified. The orientation of the streets shifted 20 to 30° eastward, delimiting about ten blocks of variable size. The workshops on the East-West street were demolished to make way for a public square surrounded by residential houses. To the south, next to the town sanctuary, a monumental fountain and shops were built. Around this 4 hectares centre were districts devoted to artisanal production and housing. Occupation extended to the north-east in the *Courates* sector, where prospecting and excavations have revealed a number of buildings and artisanal structures. The town then reached its maximum extent, covering about sixty hectares. The necropolises and the suburban buildings were added on the outskirts, including a sanctuary and completely rebuilt theatre to the west, and a public bath to the south.

Figure 5: The imposing buttress wall.

3.2. From the 4th c. to the Middle Ages

In the 4th century, the urban space went through a final transformation phase. Some sectors appear to have been abandoned, others completely reorganized. Several works were started to repair the water-distribution system or to convert shops into workshops. We also know from the *Notitia Dignitatum* that *Argentomagus* became a centre of arms production, which could explain this shift towards artisanal activities. However, the Roman town was definitively abandoned at the end of the 4th century. New settlements can be observed in the early Middle Ages (Dumasy, 2013: 154-157), initially close to the ruins of the Roman town where building material was reclaimed. The first places of Christian worship were built in some necropolises that were still in use. New centres of occupation developed in the sector of the church of *Saint-Marcel*, on the western edge of the Roman town, and in the sector of *Saint-Etienne* at the foot of the *Mersans* plateau. A topographical and toponymic shift occurred towards the left bank of the *Creuse*, where a seigniorial fortress was built, named *Argentonum* in an 8th century text. In the Middle Ages (Dumasy, 2013: 164-167), the communities settled in the two current towns: in the north, the village of Saint-Marcel, with its church, priory, and the three defensive walls built between the 12th and 16th centuries; in the south, the upper and lower town of *Argenton*, which became a *châtellenie* and the larger town of the two. We now have very detailed and accurate topographical and archaeological maps. They are not the result of automatic data display, but of an in-depth analysis by the researchers to characterize and prioritize the data entered in the GIS.

4. Spatial Analyses

4.1. Exploitation of the digital terrain models

A series of spatial analyses were carried out, using the available DTMs in order to clarify certain aspects of the development of the site in relation to its environment - inspired by the researches made on the *Epomanduodurum* site (Thivet 2009). One current research issue concerns how and why the rampart of the *oppidum* was built and the purpose of the imposing buttress walls built in the same sector during the Roman period (Fig. 5).

The relief of the rampart protecting the *oppidum* can still be seen along a distance of 300 meters. At the northern end they are completed by a wide ditch, and the difference in level between the top of the rampart and the bottom of the ditch is 10m. These structures were only partially revealed during one-off interventions in 1967, 1996 and 2012. Thus, to identify better the topography of this northern limit of the *oppidum*, we worked on the two DTMs. First, surface analysis tools were used to quantify and visualize the relief: the ‘slope exposure’ tool emphasized the direction of the greatest downward slope; the ‘‘hillshading’’ tool was used with the recommended default parameters (azimuth of 315° and altitude of 45°). The DTM derived from microtopography (Fig. 6) gives in particular a clearer view of the reliefs of this area, making it possible to reproduce an E-W rectilinear outline of the rampart. We can thus hypothesize that the buttress walls were built in the ditch to the north of the rampart. With regard to the ditch itself, after the construction of the walls, the southern half is likely to have been filled in and the northern half left hollowed out.

4.2. Hydromorphological analyses

Following the results obtained from the two DTMs, we decided to use hydrology tools to look in greater detail at the ditch bordering the rampart of the *oppidum*. According to diagnostic studies carried out by Sophie Krausz (Krausz, 1996) then by Jean-Philippe Chimier (Chimier, 2012), we can postulate that it was a natural depression and could have been the bed of a former river.

As shown on the previous maps, and as confirmed by the profiles carried out on the 2013 DTM, this depression is 80 centimeters wide, which seems excessive for a simple ditch. To support the hypothesis of a former river bed, we used two tools of the SAGA-GIS software. First,

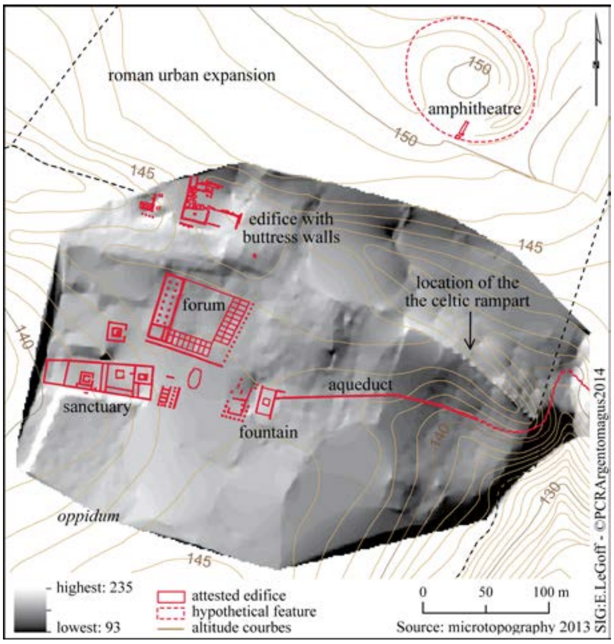


Figure 6: Some surface analyses on the DTM 2013: result of the hillshading tool.

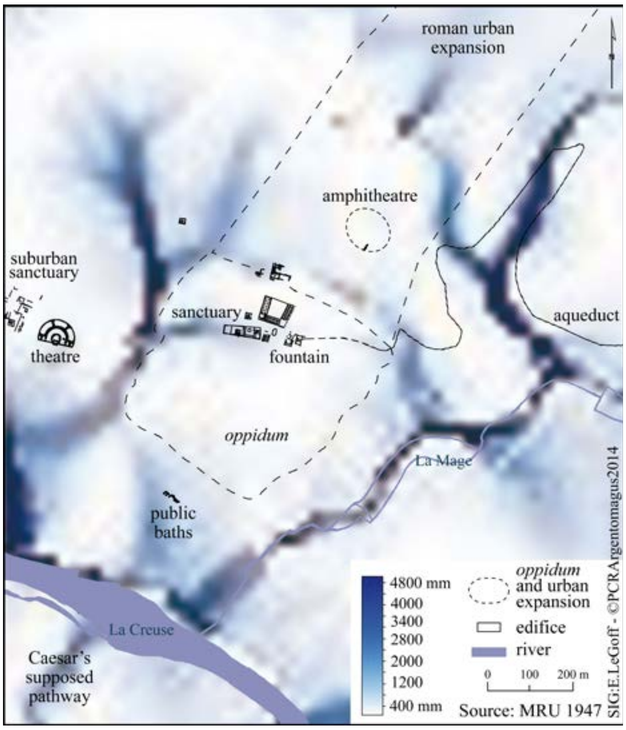


Figure 7: Some hydromorphological analyses on the DTM 2013: result of the flow accumulation tool.

the 'Topographic Wetness Index' tool is used to model hydrological flow paths on a digital terrain model; the drainage surface area and the network through which the water flows to an outlet are called the drainage system. It is also used by J. Andresen team's in a previous CAA (Andresen 2008). Secondly, the Flow Accumulation tool is used to calculate the accumulated flow as the accumulated weight of all the cells flowing into each zone.

The accumulation flow chart thus obtained shows a high rate of accumulated flow in the assumed ditch (Fig. 7), which terminates to the east by a small ravine where a tributary of the *Mage* probably flowed. It is now fairly certain that the ditch of the *oppidum* did indeed make use of a former talweg. It can even be deduced that the location of the rampart was determined by this natural depression, facilitating the task of the builders of the *oppidum* during the late Iron Age. Core drillings, recently realized in the location of the ditch of the *oppidum* (Rabasté, *in progress*), confirmed these hypotheses proposed by GIS analyses, of a thalweg fitted out as defensive system in the North of the *oppidum*.

4.3. Viewshed analyses

Finally, the third series of spatial analyses deals with the field of visibility in the Roman town. The *Argentomagus* site is in an elevated topographical position, on a promontory overlooking the *Creuse* valley. Which districts and monuments were the most visible from this advantageous position?

To investigate this issue, we used the ArcGIS 'viewshed' tool, as well as R. Zaplata team's in a previous CAA (Zaplata, 2008), which shows the areas that can be seen from one or more points, using a DTM. Note that an altitude value can be set in order to take into account the elevation of buildings. At the time, trees or other buildings could have created visual barriers that we cannot assess, but it is nevertheless possible to identify the most prominent buildings in the urban landscape.

In this way, we tested the field of view of several series of points in *Argentomagus*: the entries to the town *via* known roads, the forum, the temples, the entertainment buildings (theatre, amphitheatre) and the necropolises. Among these, the temples were undoubtedly located on the most visible position. This is particularly the case for the suburban sanctuary to the west of the town and which is illustrated here (Fig. 8). The temples built above the theatre were visible from the town centre on the *Mersans* plateau, and also from as far as the northern end of the town and the *Creuse* valley. This monumental complex thus appears to be a major visual landmark, very close to the town. We also observed that there was inter-visibility between all the places of worship in *Argentomagus*. If we extend this field-of-view test, we notice that the suburban sanctuary was theoretically visible from a large part of the land in the south-east within a radius of about 10 kilometers. Thus, the divinities who were worshipped in these temples watched over not only the town of *Argentomagus* but also a large part of the surrounding countryside.

4.4. Distance and least-cost path analyses

As the whole road and street network of the town is not yet known, we used the analysis possibilities of the GIS to try to identify more clearly the most topographically logical pathways between various points of the urban and suburban space. Inspired by the research led on an Alpine

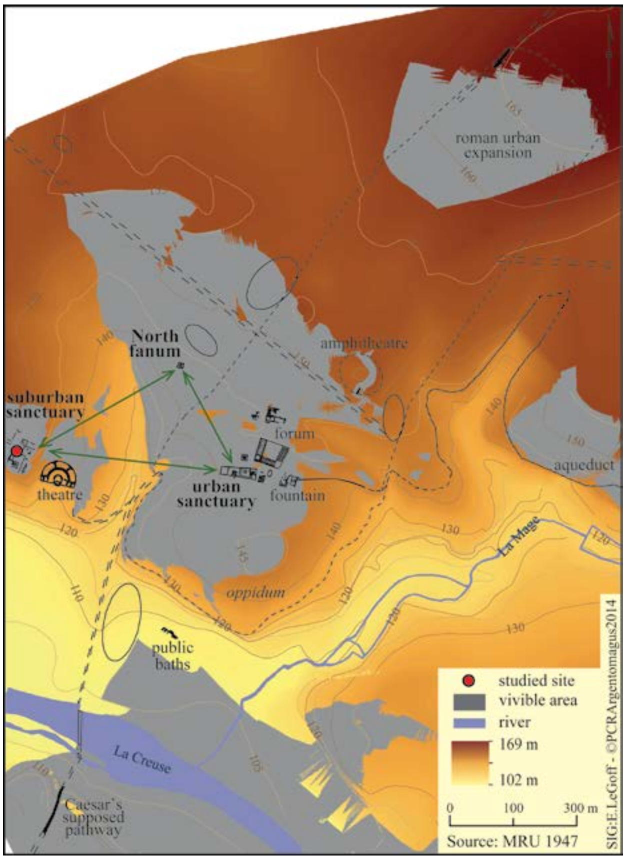


Figure 8: The importance of 'to see and to be seen': test realized from the suburban sanctuary

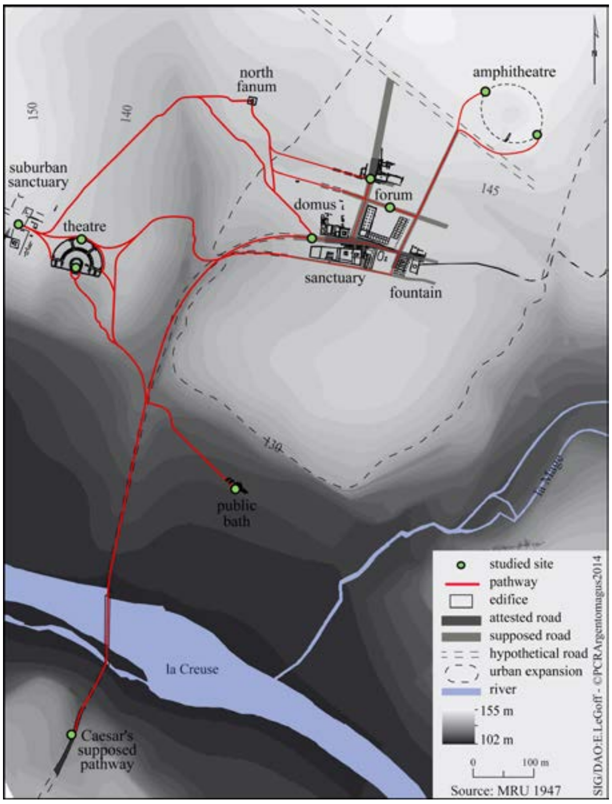


Figure 9: Reinterpreted pathways, from existing archaeological features.

Environment (Gietl, Doneus & Fera, 2008), the ArcGIS Cost-path tool was used to calculate the least-cost path between two points, taking into account not only distance but also slope values to find the most energy efficient route. The analysis involved five stages: calculation of the friction surface, of the cost allocation, of the cost path, of the transit time and of the least-cost path (the time of the journey is calculated on the basis of an average walking speed of 5 km/hour). Several routes were tested (Fig. 9): from the southern entrance to the urban centre and the suburban sanctuary, from the suburban sanctuary to the urban centre, and from the urban centre to the amphitheatre and the north sanctuary.

This modeling of the most energy-efficient route supported the as-yet hypothetical outline of certain roads, for example, that of the path between the left bank of the *Creuse* and the town centre. The most rational route was shown to be *via* the ford and the western entrance to the *oppidum*. However, other suggestions for routes can be postulated, for example, between the suburban sanctuary in the west and the town centre in the east. While it was previously thought that there was only one entrance from the south, two routes to the theatre or the temples *via* the north are now being studied, leading us to reconsider the logic of routes that could have been used for processions, from the centre to the periphery, and from one place of worship to the other.

5. Conclusion. Results and prospects

The use of GIS is thus beginning to produce the first tangible results in understanding the *Argentomagus* site. To extend the analyses that have already been carried out on relief and hydromorphology, core drilling operations, associated to sampling and dynamic penetrometer analyses (PANDA) were realized in order to confirm the location of the Celtic rampart. Combining these two methods allowed us to observe the stratigraphy of the ditch filling, providing details of the nature of the different layers and identifying precisely changes in the strata. Furthermore, it is intended to lead a new micro-topographic survey in the sector of *Les Courates*, in the northern extension of the Roman town. Analysis of man-made reliefs will be particularly useful to plan future excavation operations in this still little known sector of the town. Apart from the ditch of the *oppidum*, hydromorphological analyses revealed other paleochannel traces on the *Mersans* plateau. We will thus work on this new information, and more generally on the issue of water resources in *Argentomagus*.

Finally, some of the analyses which have been carried out on the *Argentomagus* site could be used at the scale of the surrounding area. The rural area around *Argentomagus* has already been the subject of numerous studies (Dumasy *& al.,* 2010), and we can make use of the available data for new tests, with the aid of the GIS. For example, to study the most energy-efficient routes for transporting building materials from the local stone quarries; or to analyze the field of view of rural sanctuaries, as well as funerary monuments or certain large *villae*. From the level of the site to the level of the area, GIS is now the prime medium for the spatio-temporal representation and analysis of archaeological remains in this sector.

Acknowledgements

We would like to stress that this research is supported by the ASSAAM association, the *communauté de communes d'Argenton-sur-Creuse* (grouping of local towns directed by the town of *Argenton-sur-Creuse*), and the *DRAC Centre* (regional institute for cultural research); so, we sincerely thank them for their contribution. Finally, we would like to thank Jean-Baptiste Rigot (CNRS, CITERES - LAT) for sharing his experience of using Cost Distance Analysis tools, as well as the members of the ASSAAM association and the staff of the archaeological museum of *Argentomagus* for their support and their friendship.

References

ANDRESEN, J. (2008). Topographic Wetness Index and Prehistoric Land Use. In: Posluschny, A., Lambers, K. and Herzog, I. (eds). *Layers of Perception. Papers from the 35th Annual Conference of Computer Applications and Quantitative Methods in Archaeology (CAA), Berlin, 2-6 April 2007,* Bonn: Dr. Rudolf Habelt GmbH*, p*. 405-410. Available from: http://proceedings.caaconference.org/paper/128_andresen_caa2007/

CHIMIER, J.-Ph. (2012). *' Les Ripottes ‹, 24 rue du Rio à Saint-Marcel. Rapport Final d'Opération de diagnostic archéologique*, Tours : Inrap CIF.

DUMASY, F. (dir) (2013). *Argentomagus* : La ville se dévoile. 25 années de recherches archéologiques. Catalogue de l'exposition 5 juillet - 15 décembre 2013, Saint-Marcel : Musée archéologique d'Argentomagus.

DUMASY, F., DIEUDONNE-GLAD, N. and LAÜT, L. (2010). *Travail de la terre, travail du fer : l'espace rural autour d'Argentomagus (Saint-Marcel, Indre)*, collection Mémoires 23, Bordeaux: Ausonius.

GIETL, R., DONEUS, M. and FERA, M. (2008). Cost Distance Analysis in an Alpine Environment: Comparison of Different Cost Surface Modules. In: Posluschny, A., Lambers, K. and Herzog, I. (eds). *Layers of Perception. Papers from the 35th Annual Conference of Computer Applications and Quantitative Methods in Archaeology (CAA), Berlin, 2-6 April 2007, Bonn :* Dr. Rudolf Habelt GmbH*, p*. 336-341. Available from: http://proceedings.caaconference.org/paper/106_gietl_et_al_caa2007/

KRAUSZ, S. (dir) (1996). *Le rempart protohistorique du plateau des Mersans à Saint-Marcel (Indre). Document Final de Synthèse d'évaluation archéologique*, Orléans : Direction des Affaires Culturelles de la Région Centre, Service Régional de l'Archéologie.

KRAUSZ, S. (1998). Un murus gallicus à Saint-Marcel (Indre) : nouvelle approche de l'oppidum biturige d'Argentomagus dans le contexte européen. *Revue Archéologique du Centre de la France* 37. Available from: http://www.persee.fr/web/revues/home/prescript/article/racf_0220-6617_1998_num_37_1_2789

LAÜT, L. (2013). Argentomagus dans son territoire. In : Dumasy, F. (dir.). *Argentomagus* : La ville se dévoile. 25 années de recherches archéologiques. Catalogue de l'exposition 5 juillet - 15 décembre 2013, Saint-Marcel : Musée archéologique d'Argentomagus. p. 14-19.

LE GOFF, E. (2013). La reprise du chantier archéologique de l'ouvrage d'art : l'occasion d'un enregistrement géoréférencé des structures et d'une intégration dans un SIG au temps de la fouille. In: Rabaste, Y. (dir.). *Poursuite de la fouille de l'ensemble monumental aux abords immédiats est du musée archéologique d'Argentomagus. Rapport Final d'Opération de fouille programmée*, *campagne de juillet 2012*, Tours : Inrap CIF, ASSAAM. p. 70-80.

LE GOFF, E. (2014) Relevés topographiques 2013 : prémices d'une étude portant sur le relief du plateau des Mersans. In: RABASTE, Y. (dir.) (2014) *L'ensemble monumental à l'est du musée archéologique d'Argentomagus. Rapport intermédiaire d'opération de fouille programmée, campagne de Juillet 2013*, Tours : Inrap CIF, ASSAAM. p. 95-101.

RABASTE, Y. (dir.) (in progress) *L'ensemble monumental à l'est du musée archéologique d'Argentomagus. Rapport intermédiaire d'opération de fouille programmée, Saint-Marcel, campagne de Juillet 2014*, Tours : Inrap CIF, ASSAAM.

THIVET, M. (2009). Mise en évidence de l'impact de l'occupation antique dans la morphogenèse d'un paysage de fond de vallée : le cas du site d'*Epomanduodurum* (Mandeure-Mathay, Doubs). *Revue archéologique de l'Est* 58. Available from: http://rae.revues.org/5877 [Accessed: 27 November 2009].

ZAPLATA, R. (2008). Viewshed Analysis, Regional Studies and Cultural Perception. In: POSLUSCHNY, A., LAMBERS, K. and HERZOG, I. (eds) (2008) *Layers of Perception. Papers from the 35th Annual Conference of Computer Applications and Quantitative Methods in Archaeology (CAA), Berlin, 2-6 April 2007, Bonn :* Dr. Rudolf Habelt GmbH*, p*. 343 + CD-ROM. Available from: http://proceedings.caaconference.org/paper/108_zaplata_caa2007/

The Integration of Landscape Processes in Archaeological Site Prediction in the Mugello Basin (Tuscany/Italy)

Elmar SCHMALTZ,[1,2] Michael MÄRKER,[1,3] Hans-Joachim ROSNER,[2] Adrew-Williams KANDEL[1]

[1] Heidelberg Academy of Sciences and Humanities, Germany
[2] Institute of Geography / Eberhard Karls University Tübingen, Germany
[3] Dipartimento di Scienze della Terra, Università di Firenze, Italy
elmar.schmaltz@student.uni-tuebingen.de, michael.maerker@unifi.it,
hans-joachim.rosner@tuebingen.de, andrew-william.kandel@uni-tuebingen.de

Abstract

In this study we detect and discuss possible find locations of Palaeolithic artefacts in the intra-montane Mugello basin (Italy), investigating landforms and the respective landscape forming processes based on soil erosion dynamics. The recent fluvial terrace landscape of the basin shows three main Pleistocene terraces with several Middle and Upper Palaeolithic artefacts. Fresh artefact edges on the higher terraces indicate short transport distances. On the contrary, the more rounded material from the lower terrace point to water related transport and erosion processes. Beside the application of a hydro-erosive model to understand the present day, as well as the ancient landscape dynamics and the respective distribution of artefacts, we performed a detailed terrain analysis. Static and dynamic model information was utilized to explain the present day spatial distribution of artefact assemblages. Finally, we trained a stochastic model based on boosted regression trees to predict potential artefact sites in the Mugello valley. The distribution of potential sites reflects Paleo-topography and areas characterized by geomorphic stability.

Keywords: landforming processes, Palaeolithic artefacts, hydro-erosion modelling, boosted regression trees, Mugello

1. Introduction

The northern Apennine in Italy is characterized by several intra-montane basins with a quite similar lithostraticraphic situation. The recent terraced landscape of the Mugello basin shows three main fluvial, Pleistocene terraces. The oldest and highest terrace indicates one of the paleo surfaces and is exposed at the locality Trebbiolo, where mainly Middle Palaeolithic stone artefacts (20%) were found. The middle terrace is the second oldest terrace and visible at the locality Lucigliano. Stone artefacts from both the Middle (55%) and Upper Palaeolithic (13%) were collected and examined. The youngest terrace apparent at the locality Toro contains rearranged artefacts from the Middle (35%) and Upper Palaeolithic (13%, see Fig 1). In general, the fresh edges of the artefacts collected on the higher terraces suggest short transport distances. On the contrary the archaeological material from the lower terraces is more rounded, indicating water related transport and erosion processes. A remarkable fact is that gathered artefacts from Lucigliano show only minimal evidence of transportation (~1%) and that especially the archaeological finds of the youngest (Toro) and oldest (Trebbiolo)

terraces show a quite high ratio of indets compared to these on the Lucigliano terrace. Even though, the percentages differ a lot, the relations are more or less 2/3 to 1/3 between Middle and Upper Palaeolithic find locations. Middle Palaeolithic artefacts can be related to classic Neanderthals using the Levallois technique. Predominantly cores were found, while the smaller flakes are hard to spot in the landscape. Most cores consist of local flint and silex that is available closeby, at the original geologic outcrop or in the alluvial deposits of the Sieve River or its tributaries. The

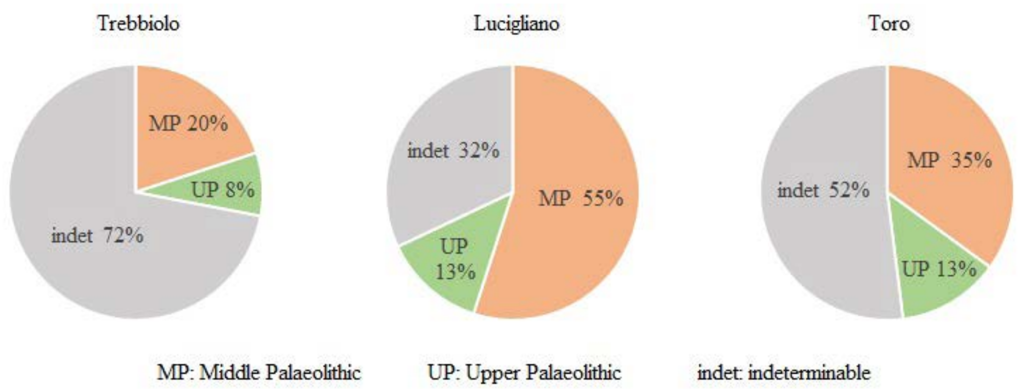


Figure 1: Percentages of Upper and Middle Palaeolithic artefacts and indets on the several terrace systems.

cortex on the bases on most of the cores suggests that river pebbles were used as raw material. The patination and the size of the artefacts decrease from the highest to the lowest terrace. In general the Middle Palaeolithic assemblages outbalance the Upper Palaeolithic. The latter consist basically of platform cores that modern humans used to produce linear flakes, which were smaller than the Levallois cores. The study yields basic information on Paleo-landscape features like terraces, characterized by artefacts. Thus, allowing a reconstruction of Paleo-landscape fragments. The assessment of present day processes will provide a deeper understanding of landscape dynamics and thus, on artefact transport and deposition conditions. Lastly, the prediction of Paleo-landscape features by underlying archaeological find locations with landforming processes can also be derived for the similar intra-montane basins in the northern Apennine.

2. Study area

The study area of the upper Sieve catchment in the Mugello basin is shown in Fig. 2. The basin is part of the Northern Apennine in the administrative Province of Florence, Tuscany. With a length of 29 km and a maximum width of 20 km, the Mugello valley has an area of about 578 km². The Sieve River drains from its source near the Futa Pass with an elevation of 930m south-eastwards through the basin and exits the basin at Dicomano, where it merges with the San Godenzo River from the homonymous valley in north-eastern direction. In Pontassieve the Sieve River converges with the Arno River, which keeps flowing into the Tyrrhenian Sea.

The climate is characterized by cold winters and mildly warm to hot summers and a concentration of medium to intense rainfalls in the autumn (Garfagnoli *et al.,* 2013). The average temperature values vary from 11-12°C in January to 22-23°C in August. The cumulative annual precipitation is around 1100mm (ibid.). The Mugello basin is one of the peripheral basins (Martini and Sagri, 1993) at the western flank of the northern Apennine. The basin developed during the blocking stage (Bousquet and Philip, 1986) in the uplift phase of the Northern Apennines and fully developed during the Pleistocene (Bousquet and Philip, 1986; Martini and Sagri, 1993). It is filled with Plio-Pleistocene alluvial and lacustrine sediments (Benvenuti, 2003) with at least three main fluvial terrace systems formed during the subsequent alluvial phase in the latest early Pliocene-Holocene, by three major episodes of baselevel fall, which also caused a facies and alluvial deposition tilting of the basin (Sanesi, 1965; Benvenuti, 2003). According to soil taxonomy the soil types most commonly occurring are chromic luvisols, cambisols and leptosols.

Other characteristics are the high content of clay, decalcified A-horizons and B-horizons where carbonate is accumulated. The chromic cambisols are colored dark yellow and slightly red and has a plastic texture which causes shrinking and swelling processes due to its clay content. The acidic soil is carbonate-free but developed out of a subsoil containing clay and carbonate. Vegetation in the Mugello basin is characterized by large areas of mixed oak or beech forests. Macchia shrubland consisting of low growing trees and spiny bushes is also prevalent. Common conifers are cypresses and trees of the genus Pinus. Agricultural areas are characterized by olive trees, wine production, corn, crop, meadow or pasture.

3. Methods

To get an adequate prediction of the potential Palaeolithic site locations in the basin, we followed a three steps approach (Fig. 3):

– Field work and pre-processing of the data set by measuring infiltrability and permeability in the field to derive the soil texture for the erosion modelling and topographic parameters. Moreover, a detection of the recent land use change was investigated, to estimate its effects on the Palaeolithic sites. Due to the fact that soil erosion is highly dependent of precipitation, we

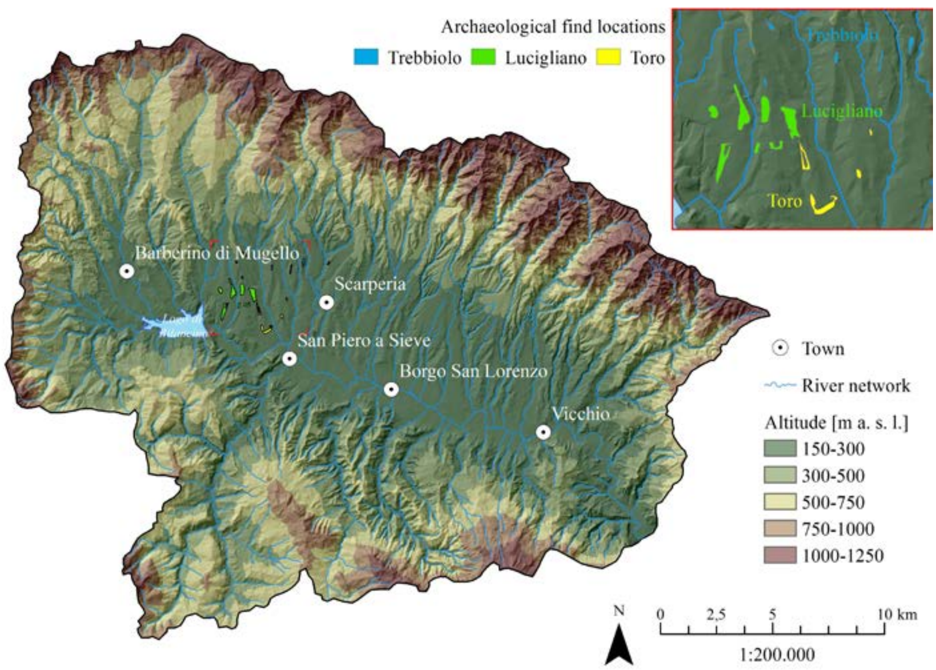


Figure 2: The Palaeolithic sites in the Mugello basin, Tuscany, Italy.

analysed as well heavy rainfall events of the last 5 decades.

- Statistical analyses of several topographic parameters.
- Prediction modelling using a stochastic approach based on Boosted Regression Trees. As predictor variable topographic parameters and the results of the soil erosion modelling was used and as dependent variable, the archaeological find locations.

3.1. Field work, pre-processing and statistical analyses

The elevation points and contour lines were collected from the open source geodata service of the region of Tuscany. With the topographic data we generated a Digital Elevation Model (DEM) after Hutchinson‘s (1989) Topo2Raster algorithm implemented in ArcGIS with 10 m resolution. To derive the soil texture class, we took infiltrability and permeability measurements with Amoozemeter and Hood-Infiltrometer (Amoozegar 1989, Schoeneberger and Amoozegar 1990) on 33 transects on slopes near the archaeological find locations.

The archaeological find locations show a significant correlation with certain topographic parameters. We used topographic indices later on as prediction variables for the boosted regression model. Therefore, we calculated complex topographic parameters like the Topographic Wetness Index (TWI, Beven and Kirkby 1979), Topographic Capacity Index (TCI), Topographic Positioning Index (TPI) and the Topographic Roughness Index (TRI), which are the parameters with the highest importance for the model. Moreover, we derived basic topographic parameters from the DEM, such as gradient slope, aspect, channel network base level, distance to the stream channel network, catchment area, plan and profile curvature, convergence index and the stream power index. We created a point grid from the shaped archaeological find locations and skipped all points with extreme values below and above the whiskers. 500 random points for each classified archaeological find location on the specific terrace system were selected. These points were attributed with the values of the topographic parameter values at the point location.

To indicate the soil erosion potential of the whole basin, we used the Revised Universal Soil Loss Equation (RUSLE) after Renard *et al.,* (1991, 1997) and the USPED according to Mitasova *et al.,* (1996) and Mitas and Mitasova (1998) for an indication of erosion, transportation and deposition processes. Five elements are implemented in the RUSLE model: rainfall erosivity (R-Factor), soil erodibility (K-Factor), slope length and steepness substituting the topography (LS-Factor), cover management (C-Factor) and support practice (P-Factor). Multiplying these parameters after Equ. 1 yields the annual soil loss per tons and hectare (Wischmeier 1959, Wischmeier and Smith 1987):

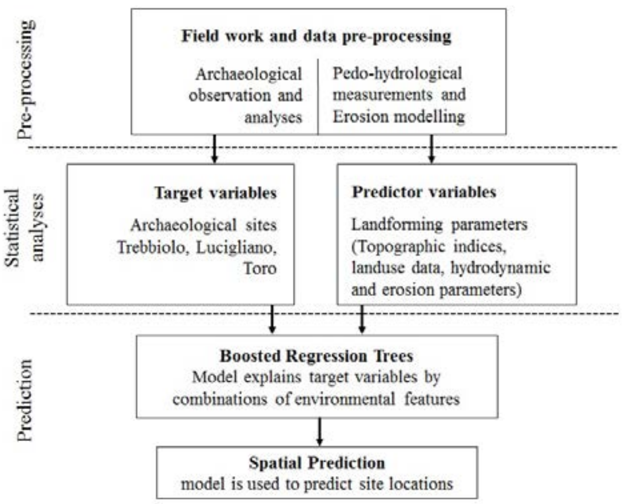


Figure 3: Workflow of the comprehensive three step predictive modelling approach.

$$A=R*K*LS*C*P \quad (1)$$

The R-Factor represents the rainfall-runoff erosivity which is basically the erosion force of intense rainfall. According to Märker *et al.,* (2008) the rainfall erosivity or R-factor was calculated with a linear model after Rufino *et al.,* (1993). The model uses monthly precipitation data (M) for calculation:

$$R=19.55+(4.2\times M) \quad (2)$$

The value of each pixel was calculated with monthly precipitation data with a resolution of 30 arc seconds and generic format. With the lowest RSME of the tested models and a R^2 of 0.61, it seems to fit best for the Tuscany region and also the Mugello basin (Märker *et al.,* 2008).

The C-factor is the vegetation cover factor constructed with remote sensing data (see Land use change detection). Low values represent high vegetation cover rates and therefore low erosion, high values stand for high erosion due to low vegetation cover rates. For RUSLE the LS-factor (slope length and steepness) was calculated after McCool (1989) with the following equation:

$$LS=(L/22)^{m}\times(slope/9)\times(slope/9)^{(1/2)} \quad (3)$$

$$m=1.2\times\sin(1/3 slope) \quad (4)$$

with: slope [%] < 4°

With increasing slope steepness and slope length, the transport capacity of the surface runoff rises. The LS-factor describes the influence of reliefs on soil loss. With the LS-calculation after McCool it‘s possible to consider slope curvatures over normal default values of the standard LS-Factor calculation in SAGA GIS. The K-factor is the soil erodibility factor which was derived using a shapefile with soil texture data of the Tuscany region. The data was regionalized for use in TreeNet and only shows two classes: sandy loam and sandy clay loam. To estimate the K-factor values, the equation after Schwertmann *et al.,* (1990) was applied:

$$K=2.77\times10\text{-}6\times M^{1.14}\times(12-OS)+0.043\times(A-2)+0.033\times(4-D) \quad (5)$$

with:

$M=(silt[\%]+sand[\%])\times(silt[\%]+sand[\%])$

OS = organic substances (for OS > 4%: OS = 4)

A = aggregate class

D = infiltribility class

After Van der Knijff, Jones and Montanarella (1999):

Sandy loam: 9 % clay, 8 % silt, 83 % sand

Sandy clay loam: 27 % clay, 15 % silt, 58 % sand

All factors (R, C, LS, K) were multiplied in a grid calculation. Three different results were acquired based on C-factors from satellite images recorded in three different years (1987, 2000, 2007). Due to the fact that the Lago di Bilancino as an artificial lake was built after 1987, the oldest satellite image doesn‘t include the lake shape. For that reason, the lake is displayed but lacks proper C-factor values. The data set which was used to create the DEM contains the lake. In this area the slope is 0. The support practice factor (P-Factor) was kept constant with the value 1.

A vital factor concerning soil erosion (besides a closer examination of precipitation per unit of time as the R-factor) is the potential of intense/heavy precipitation. A distinction is made here between heavy precipitation in periods of constant rainfall and after longer dry periods, because the latter is characterized by a higher erosion potential, which is however hard to quantify. In addition, rates of infiltration change depending on length of the dry period and the length of the more intense precipitation. Therefore, this fact is only mentioned in passing, and they are not considered in the calculation of the average soil erosion rate. Precipitation data was provided by the Servizio Idrologico Regionale of Tuscany for the stations of Borgo San Lorenzo and San Piero a Sieve. These have been filtered as following (Fig. 4 and 5): >30 mm/d independent from dryness of preceding days and >5 mm/d for dry periods lasting longer than 14 days. Dry days are considered days with <1 mm precipitation. In addition average annual precipitation was used to analyse effects of the current climate change.

An event of heavy rainfall on the 31/8/1966 in Borgo San Lorenzo can be seen as particularly noticeable, providing 134.4 mm/m² precipitation, and in San Piero a Sieve nevertheless 98 mm/m² (Fig. 5). It can be assumed that this event was a heavy summer thunderstorm, because it was no part of a longer precipitation period. However, the preceding days were not continuously dry, so that e. g. a second event on the 31/8/1962 brought 59.2 mm/m² rainfall in San Piero a Sieve and 23.2 mm/m² in Borgo San Lorenzo. Remarkable differences in precipitation and length of the dry period highlights the local influence, so we conclude that this was a summer thunderstorm too. These small-scale differences make it even more difficult to quantify the effects of such heavy precipitation events on the soil erosion processes. Hence, we assume that, the spatial density of the provided data for the Mugello basin is not sufficient. As an example for such a Vb-type weather condition is the event of the of 6/2/1951. This day had a rainfall of 135.1 mm/m², and 45.3 mm/m² on the next and 23.2 mm/m² on the following day in San Piero a Sieve. In Borgo San Lorenzo, the amount of precipitation was very similar with 40.8 mm/m², 111.6 mm/m² and 37 mm/m², so that at least the amounts seem to be characteristic for the western part of the Mugello basin. Overviewing the 60 years of precipitation data and using the filter mentioned above, Borgo San Lorenzo had 381 and San Piero a Sieve 366 events, that had a strong influence on soil erosion, temporarily and locally. These events were not considered in the R-factor because of their infiltration rate-affecting characteristics due to the fact that the R-factor uses average monthly- or annual precipitation as a basis of data.

In order to derive a soil erosion model based on the hydrology, it is necessary to figure out the land use classification of the study area. Therefore, a supervised classification of Landsat and Aster images from the U.S. Geological Survey was conducted. We defined eight land use signatures: i) water, ii) built-up areas (e.g. residential and industrial), iii) bare soil (fields), iv) meadow/pasture, v) field crops, vi) broad-leaved forest, vii) conifer forest, viii) mixed forest. The classification results were validated with the Corine land use dataset of 2007.

3.2. Predictive modelling using Boosted Regression Trees

In the last decades, geoinformatics, statistical and machine-learning methodologies made a huge progress (Märker *et al.,* 2011), especially combining these with archaeological datasets. Several methods such as logistic regression (Hosmer and Lemeshow, 2000), artificial neural networks (Kohonen 1984) and or classification and regression trees (Breiman *et al.,* 1984; De'ath and Fabricius, 2000) have been applied in a wide range of geomorphologic studies in the past (Moore *et al.,* 1993; Gessler *et al.,* 1995; Paruelo and Tomasel, 1997; Mertens *et al.,* 2002; Brenning, 2005; Grimm *et al.,* 2008). In this study, we applied a sophisticated method based on classification trees: TreeNet® (TN), which uses a stochastic gradient boosting technique (Salford Systems implementation: TN, cf. Friedman, 1999, also called boosted regression trees: Elith *et al.,* 2008). Gradient boosting constructs additive regression models by sequentially fitting

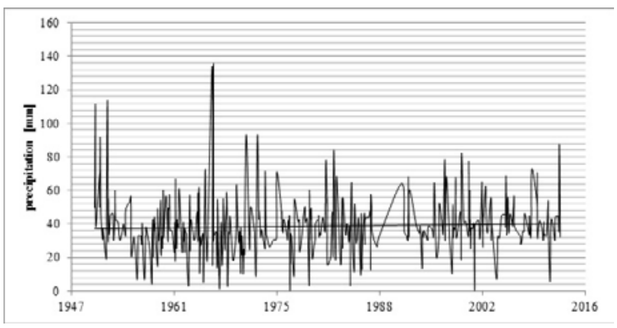


Figure 4: Heavy precipitation events 1951-2012 in Borgo San Lorenzo, no data for 1990 and 1991.

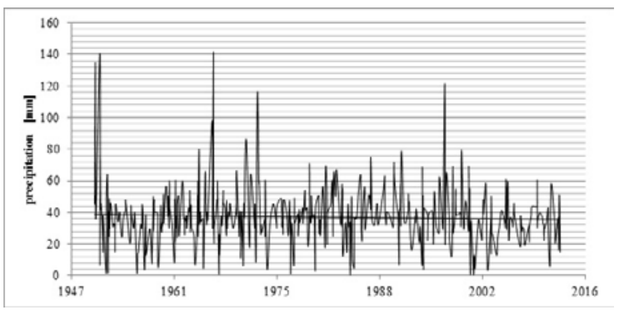


Figure 5: Heavy precipitation events 1951-2012 in San Piero a Sieve.

a simple parameterized function to current ‚pseudo'-residuals by least squares at each iteration (Märker *et al.,* 2011). The pseudo-residuals are the gradient of the loss function being minimized, with respect to the model values at each training data point, evaluated at the current step (Friedman, 1999).

The method derives several of small trees (in this study, we fixed to a maximum of six nodes), in which each tree is devoted to contributing a small portion of the overall model whereas the final model prediction is constructed by adding up each of the individual tree contributions. It's a big advantage of this methodology not being sensitive to data errors in the input variables. For our study, we run the two models with the following input parameters and settings: The entire dataset (N=1500) used for the modelling was separated into a training fraction and a test fraction (Ntrain=1196, 0.8 of N, and Ntest=304, 0.2 of N). The separation between train and test data was performed by random selection. The maximum number of trees to use was set to 2000. The TreeNet model uses a regression model with the Huber-M loss function.

4. Results

The result of the land use classification shows that for pixel values of built-up and bare soil there is a large overlap. It is difficult to differentiate these classes properly due to similar spectral information. In order to solve this problem for the classification of the year 2007 the prior probability values are specified to rather bare soil than built up. Since the result is still not satisfying the dimension of built-up areas are digitize and a new classification was generated. The result was different from the first classification but rather worse because of disproportional huge areas for field crops and meadow. For the evaluation of the training areas scatterplots are created using the spectral information from band 3 and 4. All scatterplots show a large overlap between field crops and meadow/pasture pixel values. Some problems occurred during the process of finding the best possible result for the land use distribution in the Mugello basin. Important for the classification for the year of 1987 is mainly, that the artificial lake did not exist at this time. Nevertheless the supervised classification for 1987 showed a lot of water cells. They were spread over the whole Sieve catchment area even though there are just a few waterholes which we know from the visual review of the remote sensed pictures. The Scatterplot of June 1987 shows a small overlap of the pixel values for water and mixed forest. In this value set it is difficult to distinguish and assign to a certain class. Subsequently the prior probability values were specified rather to mixed forest than water. However, the change in distribution was marginal. It is unclear if there are recently wet weather conditions that could have also influenced the radiation at the day of the recording. Therefore, the waterholes and the Sieve River were digitized separately as well as possible in their whole dimension. In forest areas the false-coloured infrared picture shows clearly that there are different types of forest whose values for reflected radiation differ from each other. These different types of forest reappear in all of the classifications. It seems that the different types of forest are broad-leaved, mixed and conifer forest. The dark spots are related to forest which shows high absorption of IR-radiation and therefore low reflection. These spots might be conifer forests with their small leaves. The CORINE3-land use file supports this assumption. On the other hand, there are bright red patches on the picture which present high reflection of IR-radiation. This could be broad-leaved forest. Darker parts in between would then be mixed forest. The Accuracy Assessment showed an overall accuracy of 55% which is rather poor. Looking at the individual classes, only the forest is classified accurate with a User Accuracy of 90%. For the other classes both the User's and the Producer's accuracy show poor values of less than 60%.

The statistical analyses by attributing the three main terrace systems with the whole amount of the dataset yielded some

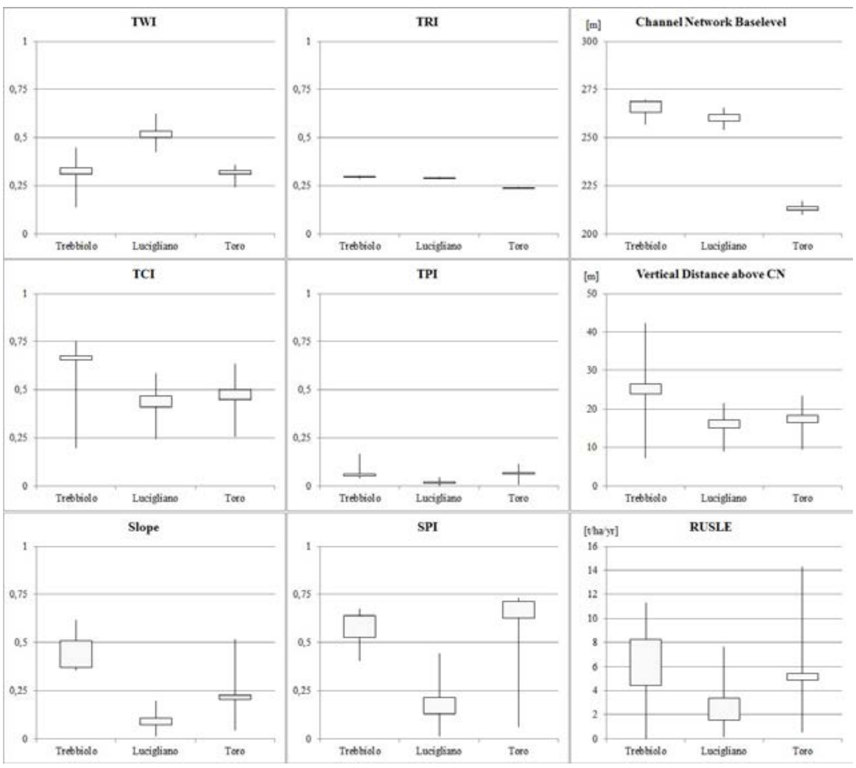


Figure 6: Relationships between topographic and landform parameters as well as soil erosion ratios with the three main terrace systems Trebbiolo, Lucigliano and Toro. For a better comparison of the topographic and landform parameters the y-axis values are shown in a normalized scale.

quite remarkable correlations between certain terraces and several input variables, such as topographic parameters like TWI and TCI, but also the soil erosion ratios and landform parameters (Fig. 6).

The results of the predictive modelling show that the model is quite robust with an accuracy of 0.98 of the training dataset and 0.97 of the test dataset, validated by the Receive Operator Curve Integral (ROC). Throughout all terrace systems, there was a high relevance of evidently important and statistically significant parameters, such as the vertical distances to the channel network, with a variable importance of 0.98 (Toro) to 1 (Lucigliano and Trebbiolo). Fig. 7 illustrates the high importance of topographic parameters for all classes, especially TWI and TCI. However, TPI and TRI played a quite unimportant role in the model. Moreover, elevation, slope, length-steepness-factor and the channel network base level were the most important predictor variables for all classes.

On the basis of the model results we regionalized the predicted areas for the whole Mugello catchment. Fig. 8 illustrates the distribution of the predicted terrace areas in the valley and furthermore possible archaeological find locations on the terraces.

5. Discussion and Outlook

In general, the patination of the artefacts decreases from the highest to the lower terraces as well as their size showing that the artefacts collected on the lower terraces were exposed and have been transported over long distances. Similarly the preservation decreases with the exception of the middle terrace at Lucigliano. These show only minimal evidence of transportation (1%) which leads to the assumption that this terrace possibly was not significantly eroded after deposition of the artefacts. Considering the several terrace niveaus, by calculating the vertical distances of the channel network to the surface and the DEM, it's possible to allocate the not assigned finds to the fitting terrace niveau. At steep slopes less energy is required to move soil. Find sites are solely located in exposed areas with an average slope ratio (not considering the fund area) of about 5.8° which causes low intensity erosion. Actually, this average ratio can be lowered to avoid too high values. Very interesting are the slope ratio differences between the three terraces regarding their find locations. The Trebbiolo and Lucigliano finds have an average slope ratio of 3.3° respectively 3.8° and are just found on very flat and low sloped plateaus. In contrast, the Toro finds are located in areas with an average slope ratio of 7.2° with two extreme values of 13.6° and 25.8°. High C-factor values represent low vegetation cover and thus, unprotected soil surfaces. For this reason soil loss is higher in areas with high slope and C-factor values. Marginal differences between the years are due to seasonal and daily changes in insolation. The higher LS-factor of the mountain areas is hardly distinguishable in current data. The same problem occurs at the calculation of the difference images. As expected, we can see that most of the finds – especially these lying in the Lucigliano terrace – are located in exposed areas with very low soil erosion and runoff ratios, respectively very low transport capacities, consequently the find areas are not highly affected by runoff, extreme precipitation events or soil erosion. We clearly show that nearly all finds are located in areas where we have low slope-, LS-Factor-, soil-erosion- and runoff-ratios, hence stable conditions.

The results of the predictive modelling with Boosted Regression Trees have shown that in first line elevation differences between the drainage network and the surface – represented by vertical distance to the channel network – and elevation are the most important parameters regarding the modelling of terrace systems in certain catchments. Secondly, especially topographic parameters like TWI and TCI not just representing slope length and steepness differences but also soil moisture content, areas of accumulation and the effect of drainage and runoff dynamics for each grid cell. Attributing these parameters to points with archaeological information on certain terrace systems yielded a high importance in the model. In order to be able to transfer determined find sites from the Mugello basin onto other intra-montane basins, the methods described in this work was tested successfully and can easily be adapted.

The use of remote sensing, terrain analysis and GIS based process models together with properly collected physiographic and archeologic field data provides the basis to get insights in the spatial distribution of archaeological find locations and techno complexes. Moreover, the use of stochastic models based on the correlation of topographic indices, satellite

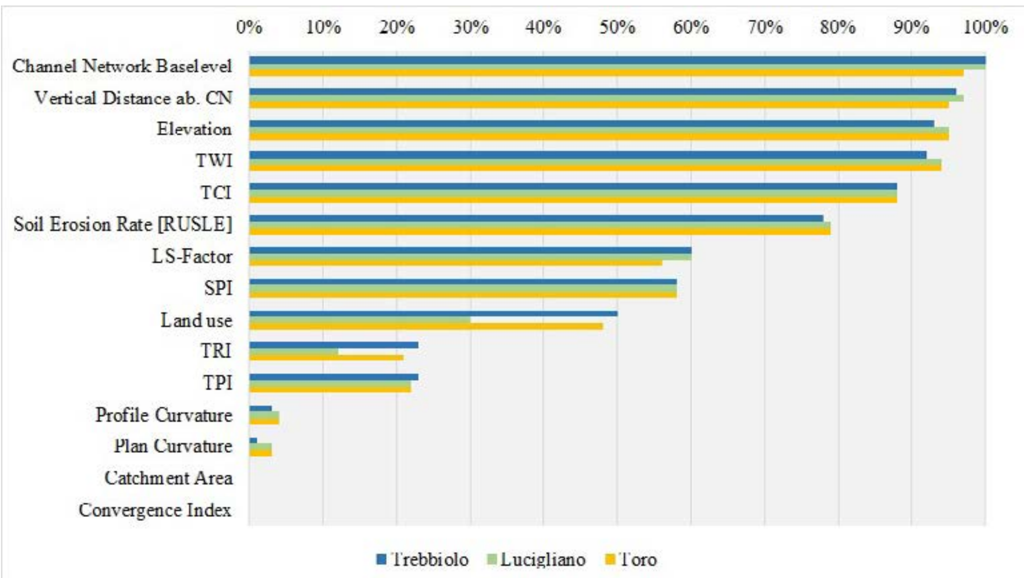


Figure 7: Variable importance scaled in relation to the most important predictor by class.

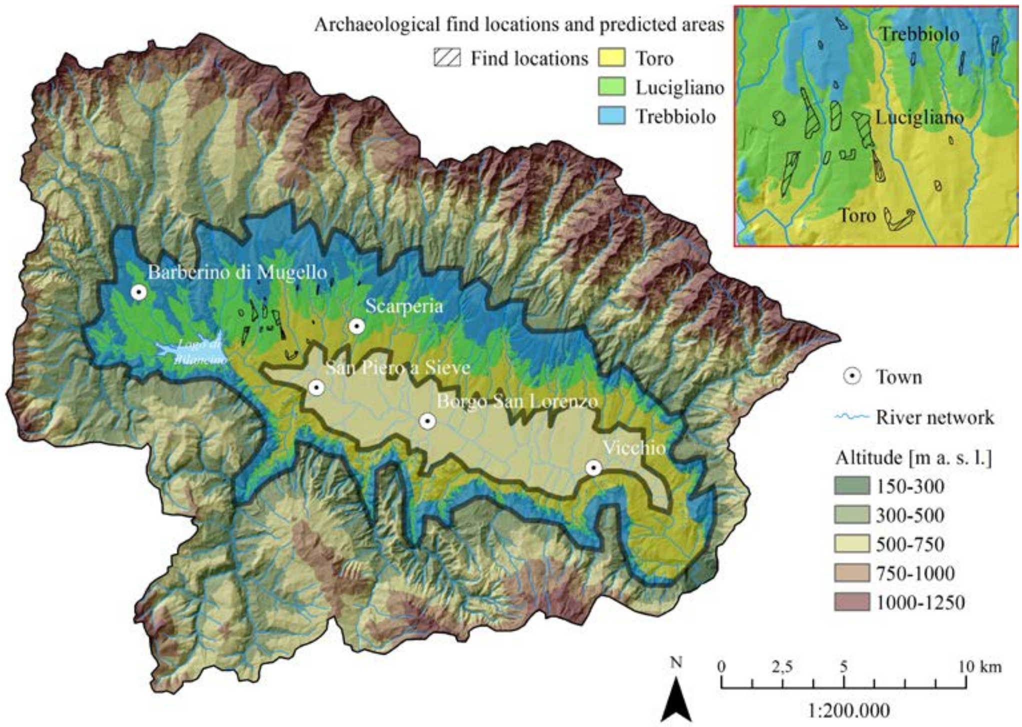


Figure 8: Distribution of the predicted terrace areas and possible locations of archaeological finds.

data and model results allow also the prediction of potential archeological sites of the Upper and Middle Palaeolithic in areas with similar physiography. Using GIS methods and geomorphological models allows also to derive a first approximation. Since no pollen profiles for the particular study area in this work were provided, further assessment of such data might be useful as additional to the reconstruction rainfall erosivities and erosion rates in the past. A spot where mentioned pollen profiles could be possibly found are cut-off meanders of the sieve, where constant sedimentation under anaerobic conditions could provide valuable data.

Summing up, high quality and quantity of specific archaeological site and technocomplex information allow stochastic modeling approaches. We show that stochastic model taking into account landscape forming processes yield valuable results. However, the quality of the model has to be validated in the field. Further integration of additional geomorphologic models, like landslides, may improve the spatial prediction. This approach can be used in similar or related areas of the northern Apennine like the Casentino basin south-east of the Mugello valley.

Acknowledgements

We would like to thank the participants of the Field School 2012, namely Sina Kaiser, Jule Maute, Rico März, Tobias Rentschler, Patrick Schwarz, Vincenz Spießhofer, Anja Willmann and Andrea Ziegler for helping us collecting and pre-processing the data. Special thanks to Paolo Gucci and Lorenzo Cammelli of the archaeological group of St. Agata. Moreover we would like to thank the Heidelberg Academy of Sciences and Humanities and the research centre: The Role of Culture in Early Expansions of Human (ROCEEH) for supporting this research.

References

AMOOZEGAR, A. (1989). A compact constant-head permeameter for measuring saturated hydraulic conductivity of the vadose zone. *Soil Science Society of America Journal* 53, pp. 1356-1361.

BENVENUTI, M. (1994). *L'evoluzione sedimentaria dei bacini intermontani dell'Appennino Setten-trionale: il bacino fluvio-lacustre del Mugello (Firenze)*. Doctoral dissertation, Universita` di Firenze, Italy.

BENVENUTI, M. (1995). Il controllo strutturale nei bacini intermontani plio-pleistocenici dell'Appennino Settentrionale: l'esempio della successone fluvio-lacustre del Mugello. *Il Quaternario* 8, pp. 53-60.

BENVENUTI, M. (1997). Physical stratigraphy of the fluvial– lacustrine Mugello Basin (Pleistocene, Northern Apennines, Italy) – *Giornale di Geologia* 59, pp. 91-111.

BENVENUTI, M. (2003). Facies analysis and tectonic significance of lacustrine fan-deltaic successions in the Pliocene–Pleistocene Mugello Basin, Central Italy. *Sedimentary Geology* 157, pp. 197-234.

BEVEN, K.J. and KIRKBY, M.J. (1979). A physically-based variable contributing area model of basin hydrology. *Hydrological Science Bulletin* 24, pp. 43-69.

BOUSQUET, J.C. and PHILIP, H. (1986). Neotectonics of the Calabrian arc and Apennines (Italy): an example of Plio-Quaternary evolution from island arcs to collisional stages. In: F.C. Wezel (ed.). *The Origin of Arcs*. Elsevier, Amsterdam, pp. 305-326.

BREIMAN, L., FRIEDMAN, J.H., OLSHEN, R.A., STONE, C.J. (1984). *Classification and Regression Trees*. Chapman and Hall, Boca Raton.

BRENNING, A. (2005). Spatial prediction models for landslide hazards: review, comparison and evaluation. *Natural Hazards and Earth System Sciences* 5, pp. 853–862.

DE'ATH, G. (2007). Boosted trees for ecological modeling and prediction. *Ecology* 88, pp. 243–251.

DE'ATH, G., FABRICIUS, K.E. (2000). Classification and regression trees: a powerful yet simple technique for ecological data analysis. *Ecology* 81, pp. 3178–3192.

ELITH, J., LEATHWICK, J.R., HASTIE, T. (2008). A working guide to boosted regression trees. *Journal of Animal Ecolology* 77, pp. 802-813.

FRIEDMAN, J.H. (1999). *Stochastic gradient boosting*. – Technical Report. Department of Statistics, Stanford University, USA. http://www.salford-systems.com/treenet.html.

GARFAGNOLI, F., CIAMPALINI, J., MORETTI, S., CHIARANTINI, L., VETTORI, S. (2013).Quantitative mapping of clay minerals using airborne imaging spectroscopy: new data on Mugello (Italy) from SIMGA prototypal sensor. *European Journal of Remote Sensing* 46, pp. 1-17.

GESSLER, P.E., MOORE, I.D., MCKENZIE, N.J., RYAN, P.J. (1995). Soil-landscape modelling and spatial prediction of soil attributes. *International Journal of Geographical Information System*s 9, pp. 421–432.

GRIMM, R., BEHRENS, T., MÄRKER, M., ELSENBEER, H. (2008). Soil organic carbon concentrations and stocks on Barro Colorado Island — digital soil mapping using Random Forest analysis. *Geoderma* 146, pp. 102–113.

HOSMER, D.W. and LEMESHOW, S. (2000): *Applied Logistic Regression*, 2nd ed. Wiley, New York.

HUTCHINSON, M.F. (1989): A new procedure for gridding elevation and stream line data with automatic removal of spurious pits. *Journal of Hydrology* 106, pp. 211-232.

KOHONEN, T. (1984): *Self-organization and Associative Memory*. Springer, Berlin.

MÄRKER, M., ANGELI, L., BOTTAI, L., COSTANTINI, R., FERRARI, R., INNOCENTI, L., SICILIANO, G. (2008). Assessment of land degradation susceptibility by scenario analysis: A case study in Southern Tuscany, Italy. *Geomorphology* 93, pp. 120-129.

MÄRKER, M., PELACANI, S., SCHRÖDER, B. (2011). A functional entity approach to predict soil erosion processes in a small Plio-Pleistocene Mediterranean catchment in Northern Chianti, Italy. *Geomorphology* 125, pp. 530-540.

MARTINI, I.P. and SAGRI, M. (1993). Tectono-sedimentary characteristics of Late Miocene– Quaternary extensional basins of the Northern Apennines, Italy. *Earth Science Reviews* 34, pp. 197- 233.

MCCOOL, D., FOSTER, G., MUTCHLER C. AND MEYER, L. (1989). Revised slope length factor for the Universal Soil Loss Equation. *American Society of Agricultural Engineers* 32, pp. 1571-1576.

MERTENS, M., NESTLER, I., HUWE, B. (2002). GIS-based regionalization of soil profiles with Classification and Regression Trees (CART). *Journal of Plant Nutrition and Soil Science* 165, pp. 39-44.

MITAS, L. and MITASOVA, H. (1998). Distributed soil erosion simulation for effective erosion prevention. *Water Resources Research* 34(3), p. 505-516.

MITASOVA, H., HOFIERKA, J., ZLOCHA, M., IVERSON, L.R. (1996). Modeling topographic potential for erosion and deposition using GIS. *International Journal of GIS* 10(5), p. 629-641.

MOORE, I.D., GARSON, R.B., LADSON, A.R. (1991). Digital terrain modelling: a review of hydrological, geomorphological and biological applications. *Hydrological Processes* 5, pp. 3–30.

MOORE, I.D., GARSON, R.B., LADSON, A.R. (1997). Digital terrain modelling: a review of hydrological, geomorphological and biological applications. – *Soil Science Society of America Journal* 57, pp. 443–452.

MOORE, I.D., GESSLER, P.E., NIELSEN, G.A., PETERSON, G.A. (1993) Soil attribute prediction using terrain analysis. *Soil Science Society of America Journal* 57, pp. 443–452.

PARUELO, J.M. and TOMASEL, F. (1997). Prediction of functional characteristics of ecosystems: a comparison of artificial neural networks and regression models. – *Ecological Modelling* 98, 173–186.

RENARD, K.G., FOSTER, G.R., WEESIES, G.A., MCCOOL, D.K., YODER, D.C. (1997). Predicting Soil Erosion by Water: A Guide to Conservation Planning with the Revised Universal Soil Loss Equation (RUSLE). – *Agriculture Handbook* 703. U.S. Department of Agriculture Research Service, Washington, District of Columbia, USA, 404 pp.

RENARD, KG., FOSTER, G.R., WEESIES, G.A., PORTER, J.P. (1991). RUSLE – Revised universal soil loss equation. *Journal of Soil and Water Conservation* 46 (1), pp. 30-34.

RUFINO, R.L., BISCAIA, R.C.M., HERTEN, G.H. (1993): Determinacão do potencial erosivo da chuva do estado do Parana. *Revista Brasileira de Ciência do Solo* 17, p. 439-444.

SANESI, G. (1965). Geologia e Morfologia dell'antico bacino lacustre del Mugello – Firenze. *Bollettino della Societa` Geologica d'Italia* 84, pp. 170-252.

SCHOENEBERGER, P. and AMOOZEGAR, A. (1990): Directional saturated hydraulic conductivity and macropore morphology of a soil-saprolite sequence. – *Geoderma* 46, p. 31-49.

SCHWERTMANN, U., VOGEL, W., KAINZ, M. (1990). *Bodenerosion durch Wasser - Vorhersage des Abtrags und Bewertung von Gegenmaßnahmen*. Stuttgart.

VAN DER KNIJFF, J. M., JONES, R. J. A., MONTANARELLA, L. (1999). *Soil Erosion Risk*. Assessment in Italy. European Comission. Joint Research Centre. Space Applications Institute. European Soil Bureau.

WISCHMEIER, W.H. and SMITH, D.D. (1978). Predicting Rainfall Erosion Losses – A Guide to Conservation Planning. *Agricultural Handbook* 537. Sci. and Educ. Admin., U.S. Dept. Agr., Washington, D.C.

WISCHMEIER, W.H. (1959). A rainfall erosion index for a universal soil loss equation. *Soil Science Society Proceedings* 23, pp. 246-249.

The Use of Burgundy Stone from Ancient Times to the Present Day

Stéphane BÜTTNER,[1] Delphine MONTANGE,[1 & 2] Laure SALIGNY, [2] Marion FOUCHER[1]
With the collaboration of Eric LECLERCQ,[3] Marinette SAVONNET[3]

[1] Archéologie, Terre, Histoire, Sociétés (ArTeHiS), UMR 6298, CNRS, Université de Bourgogne, Dijon, France
[2] Maison des Sciences de l'Homme de Dijon (MSH Dijon), USR 3516, CNRS, Université de Bourgogne, Dijon, France
[3] Laboratoire Le2I, UMR CNRS 6306, Université de Bourgogne, Dijon, France
stephane.buttner@cem-auxerre.fr, laure.saligny@u-bourgogne.fr, montagnedelphine@gmail.com
marionfoucher@hotmail.fr, Eric.Leclercq@u-bourgogne.fr, Marinette.Savonnet@u-bourgogne.fr

Abstract
Burgundy stone is widely reputed for its quality and diversity. Because it is suitable for the various requirements of construction and sculpture, it has been used continuously since ancient times. Inventories drawn up so far by researchers have been piecemeal. A new comprehensive inventory has now been undertaken with a view to renewing the involvement of various academic disciplines (archaeology and history of architecture, socio-economic history and history of art). A partnership has been formed among various laboratories at the University of Burgundy in order to ensure the validity of archaeological and historical data, develop an open platform in the form of a Wiki collaborative space and design a GIS with which to analyse the diffusion of Burgundy stone since the beginning of our era.

Keywords: Semantic Wiki, GIS, Spatial modelling, Spatial Analyses, Burgundy stone

1. Introduction

Stone is the central feature of Burgundy architecture for the professionals involved in its conservation and restoration and for historians and archaeologists of the built environment. The varied physical and aesthetic qualities of the material make it a marker of local geography. The fact that it has sometimes been transported over long distances makes it a potential vector for techniques. This is essential to a better understanding of the socio-cultural and technical influences that have evolved over time in Burgundy and further afield. Burgundy's cultural heritage and stone have been the subject of many recent studies of the history of construction in Burgundy (Büttner, 2008; Büttner, 2010). On the basis of that work, in 2013 the European Union (European Regional Development Funds) and Burgundy Regional Council financed the Corpus Lapidum Burgundiae project to develop 'a shared digital platform on the nature and use of Burgundy stone' as a tool for collating and analysing data about the quarrying and use of stone from ancient times until the present day.

The objectives of the platform are:

– to provide the historical sciences (archaeologists and historians of the built environment, art historians) with a tool for making an inventory and a corpus of geo-referenced data related to the quarrying and use of various types of stone employed as building material or as objects (statues, sarcophagi, millstones, etc.) in order to provide fresh insight into the workings and the socio-economic and cultural evolution of ancient societies;
– to provide a value-enhancing tool for the industry in a fiercely competitive international context, as Burgundy with more than 70% of national output is France's leading region for exporting stone;
– to serve as an analytical tool and decision-making support for professionals tasked with the conservation and restoration of our cultural heritage.

Part 1 presents the tool, which is a semantic wiki associated with a spatial data base and a GIS. Part 2 describes the corpus and examples of summaries of the history of techniques derived from it. Part 3 sets out the diachronic spatial analyses implemented to provide an understanding of the circulation and diffusion of Burgundy stone. The development of several analytical models provides a means to evaluate and understand the evolution of the areas of diffusion of certain types of stone over the course of time. It also provides a way to trace the routes of commercialisation and so to understand how a city or major construction site could secure a source of building materials.

2. The *Corpus Lapidum Burgundiae* as a continually evolving tool: a semantic wiki linked to a spatial data base

The information in the corpus comes from multiple sources (archaeological excavations, textual data, iconographic data, documents, etc.) across a range of disciplines (earth sciences, history and archaeology). To process and disseminate these multi-model data to meet the needs of all involved, we propose a collaborative platform using a

FIGURE 1: WIKI INTERFACE INCLUDING MASTER DATA AND DOCUMENTATION.

semantic wiki in conjunction with a spatial data base and relying on a formalisation of knowledge.

The semantic wiki with its document-oriented approach meets the requirements for consultation by open access, collaboration, user role management and access to multimedia data. The spatial data base meets the needs for analysing and exporting data in a structured format to other tools such as geographical information systems.

2.1. Creation of the semantic wiki

We have used the MediaWiki (http://www.mediawiki.org/) engine as the basis for developing the semantic wiki. A wiki is organised in the form of articles that can be classified by category. An article is composed essentially of textual data to which documents (images, plans, pdf documents, etc.) and links (to other articles in the wiki or to external resources) are added (fig. 1). The semantic part has been developed with Semantic Forms (http://www.mediawiki.org/wiki/Extension:Semantic_Forms/fr), which is an extension of MediaWiki (http://semantic-mediawiki.org/). This allows users to add and modify data by using forms and models whose various fields may be automatically associated with a category via an annotation mechanism, thereby controlling the information entered (Leclercq *et al.*, 2013).

The annotations are based on an ontology specific to the Corpus Lapidum Burgundiae extended from the ontology of the CIDOC CRM which has emerged as a standard domain ontology for the cultural heritage domain (http://www.cidoc-crm.org). It deals with concepts at a high level of generality. The Corpus Lapidum Burgundiae ontology has several parts: (a) concepts related to buildings, their spatial relationships and characteristics; (b) types of stone and their characteristics; (c) stone cutting tools (chisels, bush hammers, picks, etc.) and construction techniques; and (d) quarries. Figure 2 illustrates all these concepts (concepts in blue with EXX are CIDOC-CRM concepts).

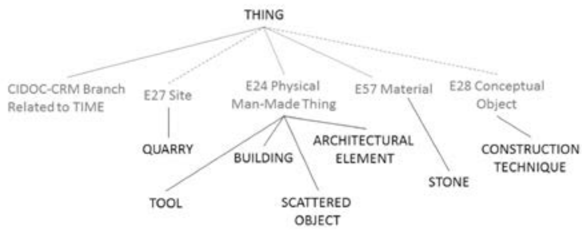


FIGURE 2: STRUCTURE OF CORPUS LAPIDUM BURGUNDIAE ONTOLOGY (LECLERCQ ET AL., 2013).

Navigation in the wiki is via a search module and an index-based input system suitable for each type of user of the platform. The index is created directly by semantic requests: buildings, stones used in construction, lapidary objects, stone in quarries, communes and quarries currently in use.

2.2. The spatial data base

From the categories associated with the items of the wiki and its own data base, a mechanism dynamically feeds a data base, the scheme for which is a subset of the wiki's semantic properties. Table 1 shows how the semantic wiki items match with the Postgre data base. For example, a wiki item describing a quarry corresponds to a line of the quarry relation of the spatial data base. That line is composed of descriptors from the text of the wiki article. Figure 3 summarises the architecture developed. PostGIS 2.1 interfaces directly with this base and supports mapmaking and analysis using QGIS and GRASS software.

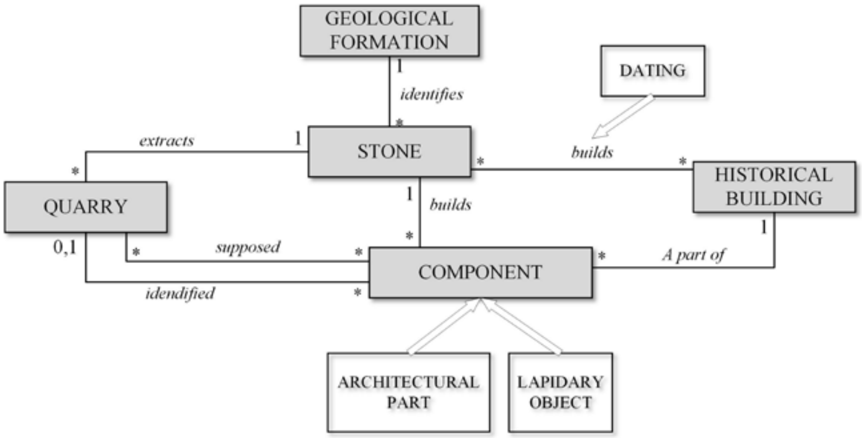


Figure 3: Conceptual model for the project.

3. A powerful data synthesis tool

3.1. The corpus

The collection of data reflects the multi-partner and interdisciplinary character of the project. Besides integrating many data already published, scientists, building specialists, archaeologists, geologists and art historians can feed their respective works into the corpus. They look primarily to supplement the inventory of built sites and stone objects, to clarify their dating and identify the materials on site and in quarries. They are supported by extraction industry specialists for the contemporary period and for quarries that are still operational.

The data in the inventory are (fig. 3): the quarries from which the stone is extracted (the stone is described by its geological formation), and the buildings in which the stone is used (as architectural elements such as walls, flooring, etc. or as lapidary objects).

The quarries are described by their name, geographical and administrative location, owner (private or public, etc.), operational status (operational, closed, periodic) and type (open cast, underground, mixed). The stone is described by its generic name, other names, colour, geological formation (fine sandstone, etc.), geological stage (Triassic sandstone, etc.). For each building the name, geographic

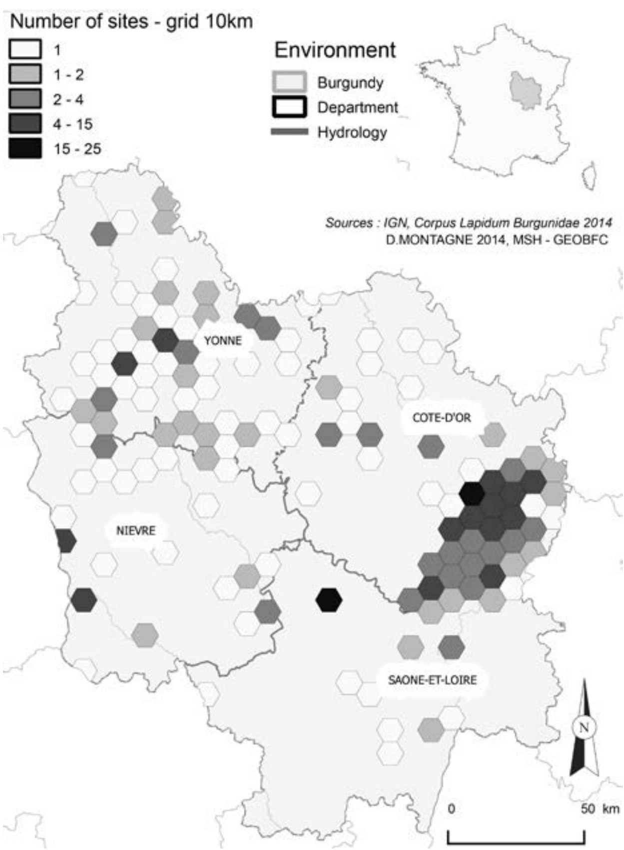


Figure 4: Density of studies sites.

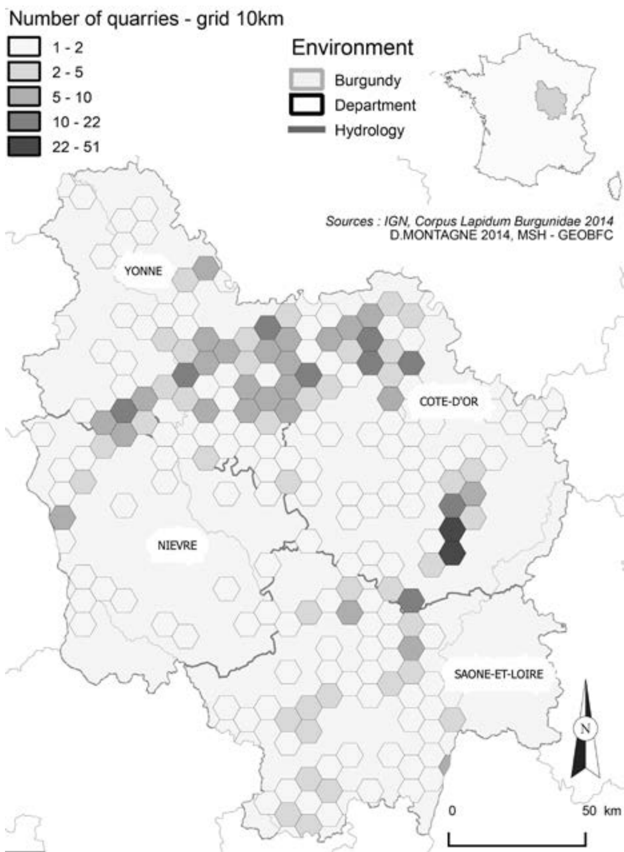


Figure 5: Density of quarries listed.

and administrative location, and type (listed, protected, etc.) are recorded in the data base. The link between the stone, the building and quarry is made via the element: the part of the monument or object (buttress, pillar, etc.).

In this way the wiki can be used to visualise the whole corpus of sites entered. These data can be openly consulted by any users (but data capture is restricted to authorised operators).

The corpus is still under construction. At present 353 sites have been listed for Burgundy with a very marked disparity (fig. 4). The data entered relate to the sectors in earlier or current studies: around Auxerre in the north (Büttner 2008), and between Dijon and Châlon-sur-Saône in the east (Foucher forthcoming).

For stones identified among the lapidary and building elements, the platform lists 1057 stones for 66 geological formations and 73 facies. Again the documentation exhibits disparities. The best represented geological facies are Pierre de Tonnerre (Yonne), Pierre de Dijon and Pierre d'Asnières (Côte d'Or).

Some 718 quarries have been included in the corpus, mostly from the Bureau de Recherches Géologiques et Minières (BRGM), supplemented by ancient quarries known from the literature and from prospecting studies. With the exception of the inventory of quarries for ancient periods which is not exhaustive, the data are representative of the true situation: heterogeneous distributions with voids in sparsely populated areas (the centre with the crystalline Morvan massif) and areas with no building stone (the clay floodplain of the Saône in the east) (fig. 5).

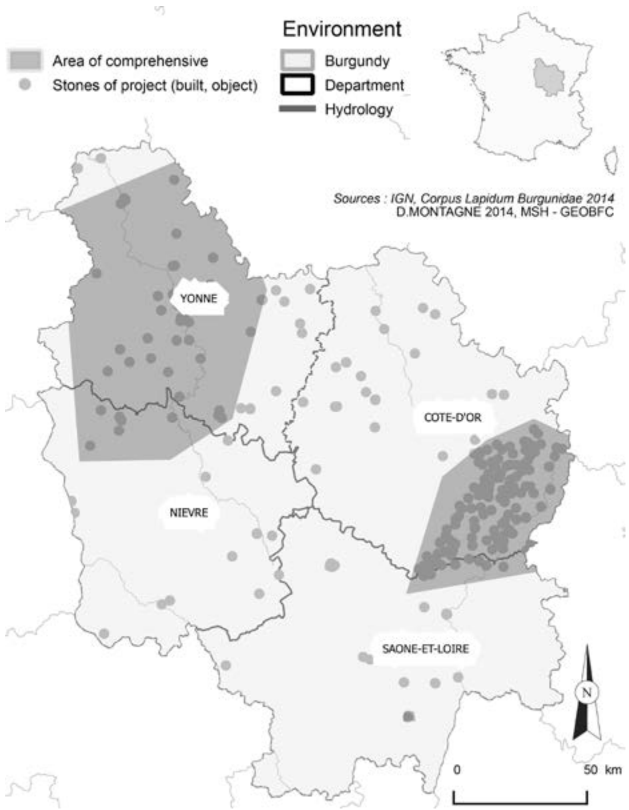


FIGURE 6: RELIABILITY MAP.

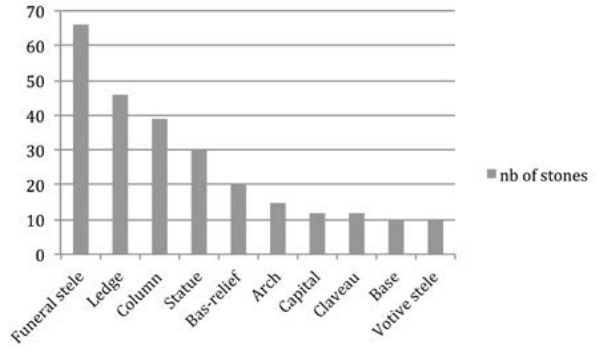


FIGURE 7: TOP TEN USES OF 'TONNERRE' LIMESTONE.

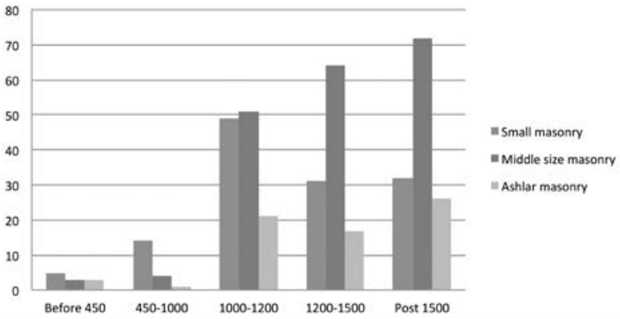


FIGURE 8: ARCHITECTURAL MODULE SIZE OVER TIME FOR 'TONNERRE' LIMESTONE.

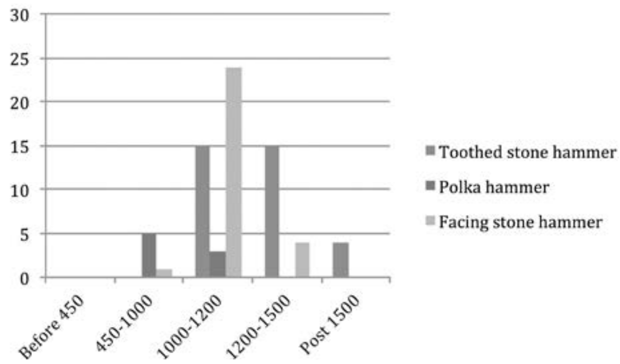


FIGURE 9: TOOL USE OVER TIME FOR 'TONNERRE' LIMESTONE.

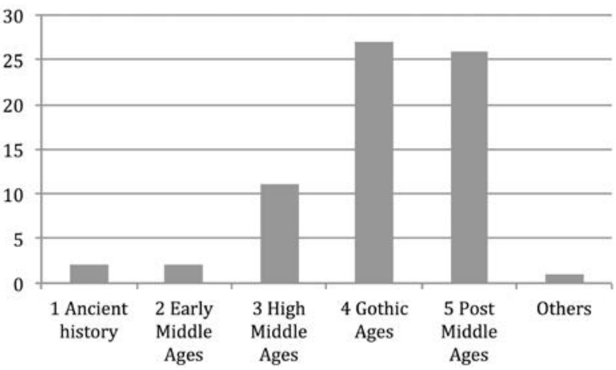


FIGURE 10: FREQUENCY OF USE OVER TIME FOR 'TONNERRE' LIMESTONE.

3.2. Corpus synthesis

Two fundamental principles have been imposed in structuring the data base. Each archaeological object referenced, whether a construction or a lapidary object (statue, sarcophagus, millstone, etc.) is systematically attributed to a site localised in space. The data are broken down into several thematic tables (site, object, quarry, stone, etc.) or cross-sectional tables (time data). This makes it possible:

– to extract syntheses but primarily to make complex queries by integrating spatial and chronological parameters;
– to take into account geographical and geological settings in addition to waterways or roadways, and even settlement patterns in considerations about the areas of diffusion and worksite supply.

As in the Archaedyn project (Ostir *et al.,* 2008), scientists were asked to record the geographical areas that were best covered or studied for stones in order to produce a state of research at a given point in time. The aim is to identify 'reliable zones' in which data may be mobilised for diachronic analysis and comparison (fig. 6).

Independently of data spatialisation, the processing of the corpus can be used to consider the evolution of certain parameters that appear to be information of interest not just for the history of the built environment but for the evolution of technical thinking and practices, and the development of quarrying centres.

As the architectural or decorative function of each stone is determined, it is easy to show that some stones were actually used preferentially for specific purposes depending on their physical and mechanical properties. Stone that was softer to cut (fig. 7) was chosen for sculptures or architectural elements with elaborate moulded features (capitals, cornices, etc.). Harder or less porous stone was selected for use in underpinnings or load-bearing elements (foundations, pillars, etc.).

The dimension of dressed stones is a parameter that interests historians of the built environment. Although it is not possible to extract large blocks from all geological facies, even so a gradual generalisation of medium modules is observed around the year 1000 (fig. 8). The success of some of these facies seems to coincide with this phenomenon, probably tied in with the more systematic use of stone vaulting requiring suitable load-bearing elements.

Tool marks, commonly recognised on building stone, provide evidence for dating that is often advanced by specialists of the ancient built environment (appearance and disappearance). However, this argument leaves room for improvement in all too many instances in terms of precision and in its relationship with the quality of materials. The corpus now provides input for consideration of this point but needs supplementing. For some facies, major trends do appear to be consistent with assessments already underway on this subject (fig. 9).

It is comparatively straightforward to determine fairly precisely the period of appearance or disappearance of a given geological facies (fig. 10). Again, this argument may appear to be crucial for historians of the built environment. While inventories are not exhaustive, this approach can

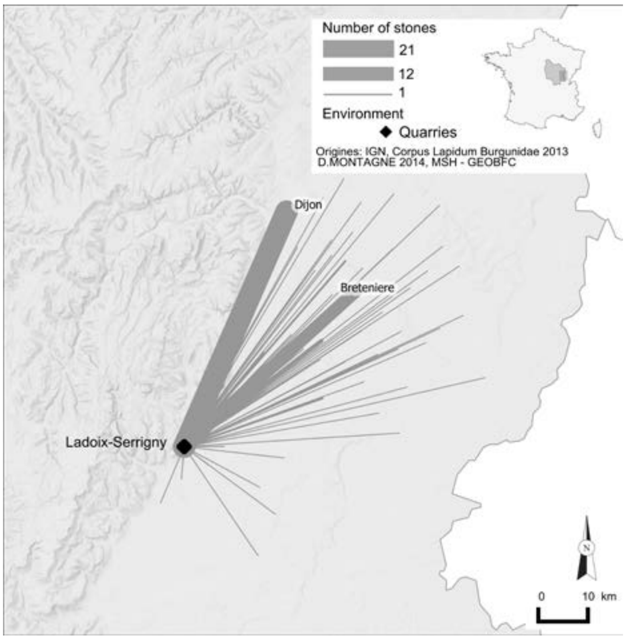


Figure 11: Potential area of diffusion from the Ladoix-Serrigny quarry

also be used to assess the intensity of use of a type of stone over time and so the evolution of activity in the quarries associated with it.

All these syntheses provide potentially essential information for understanding the history and development of building techniques and so are crucial to an understanding of how worksites operated in the past. A whole vista of the socio-economic organisation of ancient societies is reflected in this.

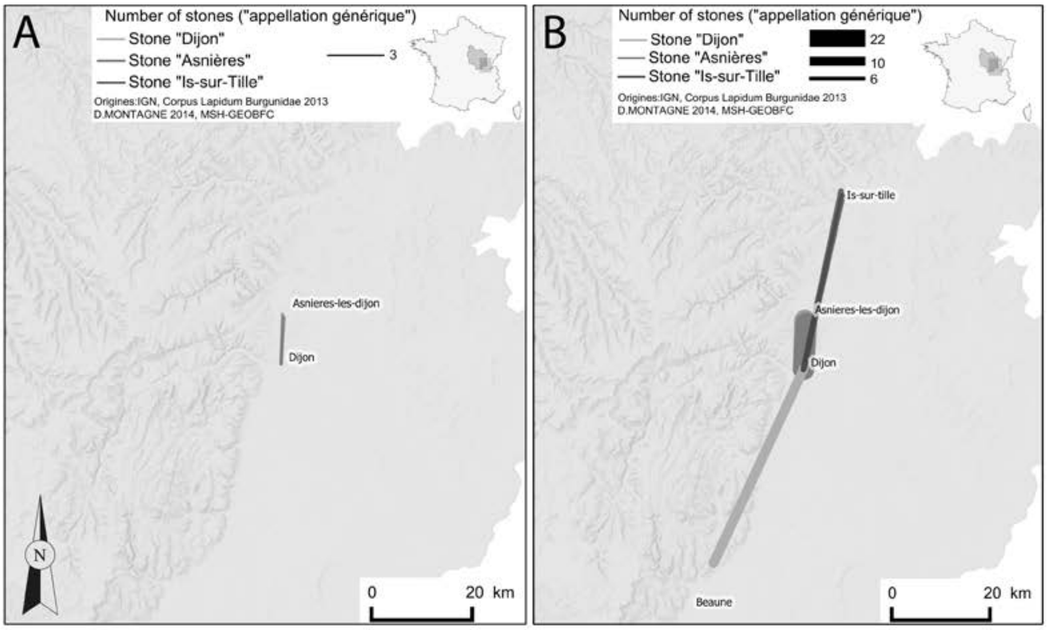
4. A tool for analysis and modelling

Because most of the data is geo-referenced, it is possible to study the potential areas of diffusion of a quarry's output. Conversely, it may also be possible to determine, for a given construction or administrative area (commune), the quarries that might have supplied the material. By modelling the routes by which building materials could be transported it is possible to contemplate the regional organisation of this market. In the absence of any historical data on the subject, this is the only way to approach the evolution of the market for stone over time and therefore the economic operation of building sites.

4.1. Models of diffusion

Sites and quarries are interrelated by the creation of flowmaps. This form of schematic representation is useful for highlighting trade between places or for 'visualising the attraction of certain places' (Brunet 1987; Cauvin 2007; Poidevin 1999). The scale of trade is represented by the size of flows which varies with the number of uses of a type of stone as a building element or an object (sculpture, sarcophagus, etc.) in a given place.

The analysis can be used to display the potential area of diffusion of a quarry's output (fig. 11). This example presents a quarry at Ladoix-Serrigny in Côte d'Or. This



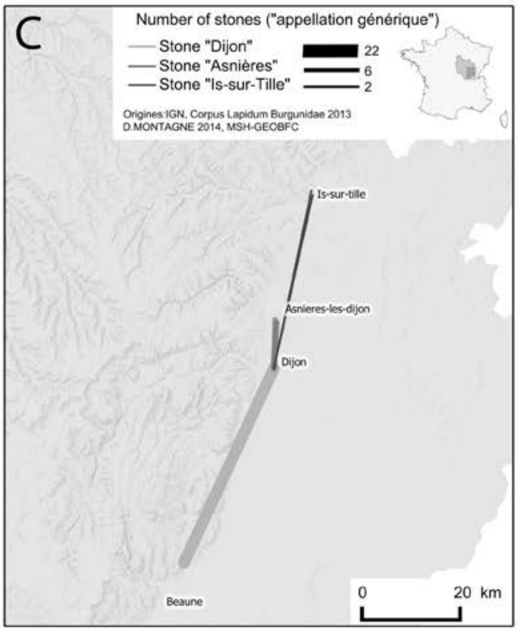


Figure 12: Potential origin of building stone for Dijon (A: before 1200, B: 1200-1500, C: after 1500).

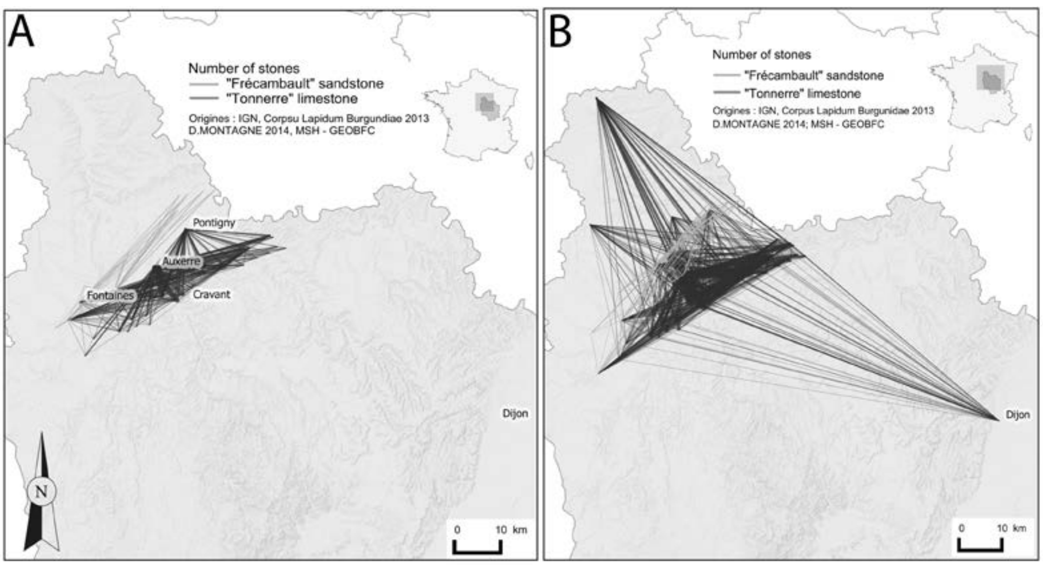


Figure 13: Potentials areas of competition between 'Tonnerre' Limestone and 'Frecambault' sansdstone (A: before 1200, B: after 1200).

quarry is still operational and supplied material for, among others, the construction sites of renowned religious edifices in the neighbourhood.

By adding the time factor, it becomes possible to display the evolution of the market for stone.

For Dijon:

– before 1200 most material came from the quarries at Asnières (fig. 12A)
– in the late Middle Ages, the sources of supply diversified. New types of stone appeared such as Dijon stone or Is-sur-Tille stone (fig. 12B).
– After 1500, Dijon stone (actually from quarries along the vineyard escarpment) apparently prevailed (fig. 12C).

In the case of building stone for Dijon, it is possible to see the progressive appearance of new geological facies, new types of stone with different qualities in terms of colour, grain and hardness. This is an essential parameter for an understanding of the changing architecture of Dijon in the Middle Ages.

It is also a crucial approach to understanding the major construction sites of the time by providing insight into the socio-economic aspect. Particularly because the urban centres and major worksites that developed there are to be viewed as catalysts for the local and neighbouring quarrying industry.

Another possible application is to compare and contrast areas of diffusion and therefore the relative success of various quarrying districts at a given time. This can be used for addressing the question of competition and for measuring its development over time.

Here the potential areas of diffusion of material from the Frécambault sandstone and Tonnerre limestone quarries are compared. In the Romanesque period (fig. 13A), use of both types of stone was largely confined to the areas where the corresponding geological strata cropped out. However, in the ensuing Gothic period (fig. 13B), Tonnerre limestone was in demand on a regional scale and seemed to corner the market. Unquestionably the quality of the stone, which is easy to cut and much sought after by sculptors and stone cutters, accounted for its success. By contrast, Frécambault sandstone, which is difficult to quarry and to dress, failed to extend beyond the local market. The routes for diffusing the output from the quarries probably also played a part in the contrasting success of the two facies.

4.2. Modelling the routes for transporting stone

Historical data and more specifically accounting records refer to transport by road and water. However, these records are essentially about worksites at the end of the Middle Ages. For earlier periods and for certain edifices of the end of the Middle Ages there are no accounting documents referring to this aspect. Modelling tests have been implemented to estimate potential pathways from the quarry to the work site (archaeological site).

The distance-cost method, used in archaeology (Chataigner and Barge, 2007; Fiz and Orengo, 2007; Gietl *et al.*, 2007) can be used to provide hypotheses about transport for known pairings of origins/destinations. These theoretical endeavours are based on various field criteria:

– an overland model, based on gradient class criteria. The digital elevation model has been revised to yield four classes with costs that increase with the increase in gradient.

Figure 14: The apse of Pontingy abbaye (Yonne).

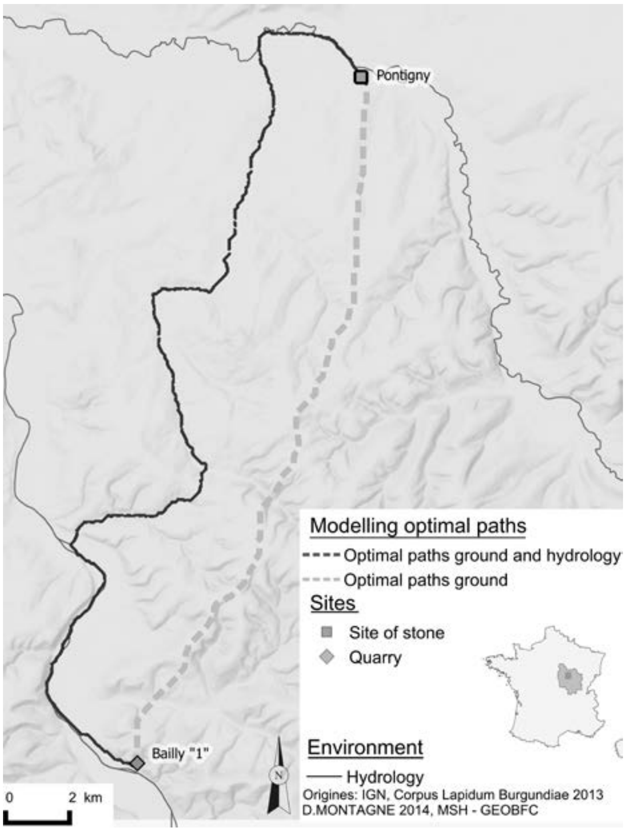


Figure 15: Optimal paths between quarry (Saint-Bris / Bailly) and site (Pontigny).

– a 'mixed' model has been chosen using gradient classes and the hydraulic factor. This model includes the navigability of waterways as the prevailing criterion with a lower cost than the gradient classes.

These two criteria are used to compute a cost raster with weighting for distance. This is then used to calculate an aggregate cost raster from some starting point. For example, we have attempted to model the cost for the supply of stone for the construction of the chevet of Pontigny Abbey (Yonne – fig. 14). While simple cost modelling with distance points to the quarries of the Chablis area (in the Serein valley), textual data refers to the quarries of Saint-Bris/Bailly in the Yonne valley.

In a second stage, the aggregate cost raster is used to analyse the shortest path between a starting point (the quarry) and the arrival point (the stone's destination, the site).

A first model, by an overland route (fig. 15), allows the material to be transported directly from the quarry to the worksite by a roadway with a relatively favourable gradient. However, a mixed model, combining transport down the Yonne valley to Auxerre and then overland from Auxerre to Pontigny, also seems feasible. Given that the same stone is known to have been carried for building in Auxerre and given the overlap of the path computed with a path shown on Cassini's map, it seems that the second model can be thought the more relevant.

These analyses open up above all new avenues of enquiry and new questions. For the overland mode, what is the actual influence of gradient on the cost of transport? How can changes in means of transport over time be contemplated? How can the cost of transport switching in the mixed model be evaluated? How can taxes and tolls that were common in the Middle Ages be taken into account?

These early analyses clearly lead to a redefinition of certain parameters of the model and prompt new collaboration with historians with specialist knowledge of road networks, potential for transport by waterways (geographers, watercraft specialists) and the associated duties.

In any event, many studies of transport of materials in the Middle Ages show that waterways were given precedence for practical reasons and above all for financial reasons. Accounting records from the late Middle Ages are there to attest to this and show that loading and unloading were crucial stages in the operation.

Even so, stone was often used outside the river valleys and mixed transport by boat and road is consequently an obvious model calling for facilities for switching modes of transport. The linear transformation model is a promising approach making it possible to consider such facilities that are still largely unknown (fig. 16).

This method of representation developed in 2004 by the team of geomaticians and archaeologists in Dijon has been used to represent circulation over time: linear cartographic

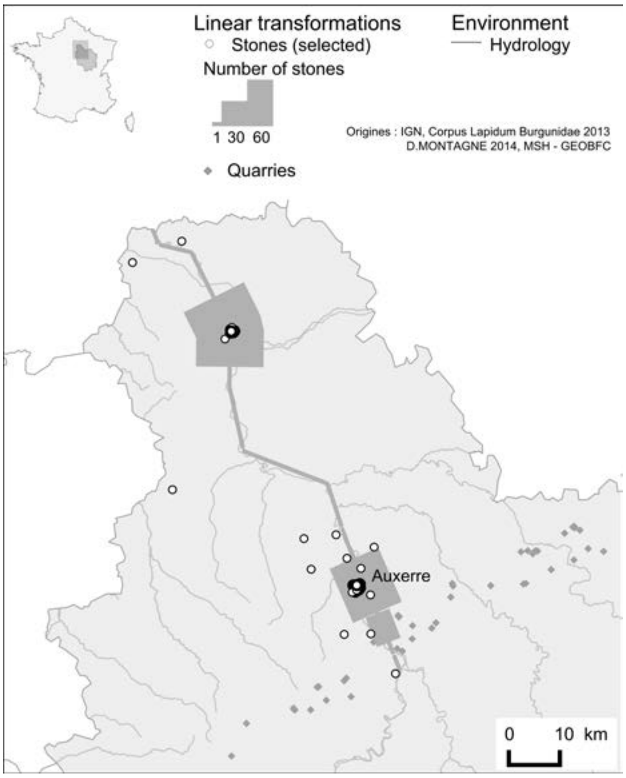


FIGURE 16: LINEAR PROJECTION ALONG YONNE RIVER.

transformations (Mordant *et al.,* 2005; Saligny *et al.*, 2006). Based on approaches of linear segmentation and geometrical projection, the aim is to transform a scatter of points cartographically into a schematic route the emplacement of which follows what is considered to be a circulation route such as a river valley. This method, which has proved its worth for studies of the circulation of materials and finished products in the Bronze Age (Gauthier, 2009; Tremblay-Cormier, 2014), has since been developed as a Python software package allowing it to be transposed to other thematic areas.

The corpus of data seems to be too small to consider any definitive conclusions at present, but it is a promising avenue for identifying potential disembarking areas. It will remain for archaeologists and historians to check whether port facilities are actually to be found at the focal points.

5. Conclusion

The initial objective of this project was to construct a collaborative and permanent tool for collating all types of data pertaining to Burgundy stone. The decision to develop a web platform on the basis of a wiki engine has come up to expectations by offering a flexible and developable tool. The platform also proposes possibilities for close interaction among contributors since data can be supplemented, criticised and modified at any time.

Analysis of the information captured, which is conducted using the data base and the GIS module, also makes it possible to discuss its validity. Processing of a large amount of data means that any diverging element can be identified and its validity queried. The models we have developed account for certain rationales about the supply of materials to urban agglomerations while emphasising the importance of technical and historical data. They can now be used to investigate means of transport and the routes used. Other analyses open up other avenues of enquiry especially into the evolution of the areas of diffusion of certain stones or the ensuing phenomena of competition. Similarly, art historians might reconsider certain stylistic affiliations in this new light. Finally, the tool is flexible enough to adapt to the issues involved (archaeology, art history, history of techniques, socio-economic history).

The census made of stones used in Burgundy clearly shows that present-day administrative boundaries bear little relation to history and that the geographical and geological context largely explains the emplacement of certain commercial networks. Again the issue of competition arises. It seems there was substantial trade with the Île-de-France in the north of the region (Paris stone) and with Allier stone on the Loire (Apremont stone).

Similar projects involving stone are currently being developed in other regions (a Franco-Belgian project for the Meuse, a project in Aquitaine, etc.) and the platform seeks to federate participants around the project which has proved its relevance and is sufficiently flexible to adapt to address new and potentially highly specific issues.

Acknowledgements

We are grateful to Association Pierre de Bourgogne, Laboratoire de Recherche des Monuments Historiques and Direction Régionale des Affaires Culturelles de Bourgogne for their investment in this project.

References

BRUNET, R. (1987). *La carte mode d'emploi*, Paris: Fayard-Reclus. 270 p.

BÜTTNER, S. (2010). Production et utilisation des pierres bourguignonnes de l'Antiquité à l'époque moderne. L'intérêt d'une base de données. In: Carvais, R., Guillerme, A. Negre, V., and Sakarovitch, J. (eds.). *Édifice et artifice Histoire Constructive. Proceeding of the 1th Colloque francophone sur l'histoire de la construction, Paris, 19-21 June 2008*. Paris: Picard. pp. 513-522.

BÜTTNER, S. (2008). Archéologie de la pierre à bâtir médiévale à Auxerre et dans l'Yonne. In: Blary, F., Gely, J.P. and Lorenz, J. (eds.). *Économie de la pierre de l'Antiquité à la fin des temps modernes. Proceedings of the International Conference ' Pierres du Patrimoine Européen '. Château-Thierry, 18-21 October 2005*. Paris: CTHS. pp. 143-158.

CHATAIGNER C. and BARGE O. (2008). Quantitative Approach to the Diffusion of Obsidian in the Ancient Northern Near East. In: Posluschny, A., Lambers, K., and Herzog, I. (eds.). *Layers of Perception. Proceedings of the 35th International Conference on Computer Applications and Quantitative Methods in Archaeology (CAA), Berlin, 2-6 April 2007*. Bonn: Kolloquien zur Vor- und Frühgeschichte Vol. 10. 375.

CAUVIN C., ESCOBAR F., and SERRADJ A. (eds.) (2007). *Cartographie Thématique ' Des transformations incontournables '*, Vol. 2, traité IGAT, série aspects fondamentaux de l'analyse spatiale, Paris: Hermès Sciences – Lavoisier. 269 p.

FIZ, I. and ORENGO, H. (2008). Simulating communication routes in Mediterranean alluvial plains, In: POSLUSCHNY, A., LAMBERS, K., and HERZOG, I. (eds.) (2008). *Layers of Perception. Proceedings of the 35th International Conference on Computer Applications and Quantitative Methods in Archaeology (CAA), Berlin, 2-6 April 2007*. Bonn: Kolloquien zur Vor- und Frühgeschichte Vol. 10. p. 316-321.

FOUCHER, M. (forthcoming). *L'économie de la pierre au Moyen Âge : approche pluridisciplinaire et comparative des ressources et de leur mise en œuvre, entre plaine de Saône et côte viticole bourguignonne*, PhD Thesis, Université de Bourgogne.

GAUTHIER E. (2009). La Consommation et la circulation du métal à l'âge du Bronze dans le Jura et les plaines de la Saône. In: Barral, Ph., Daubigney, A., Kaenal, G., Mordant, C., Piningre, J.F. and Richard, A. (eds.). *L'Isthme européen Rhin-Saône-Rhône dans la Protohistoire. Approches nouvelles en hommage à Jacques-Pierre Milotte, Besançon, 16 -18 octobre 2006*. Besançon: Presses universitaire de France-Comté. Annales Littéraires ; Série ' Environnement, sociétés et archéologie. pp. 145-159.

LECLERCQ, E., SAVONNET, M., TROYA, A.C. and BÜTTNER, S. (2013) Investigating a multi-paradigm system for the management of archaeological data : Corpus Lapidum Burgundiae. In: D*igital Heritage: Proceeding International Congress, Marseille, 28 October - 1 November 2013*. pp. 679-682.

MORDANT, C., SALIGNY, L., and POULET-CROVISIER, N. (2006). La circulation des objets métalliques à l'âge du Bronze entre Rhône et Saône: utilisation d'une transformation cartographique linéaire. @*Mappemonde* 83 (3). Avaible from: http://mappemonde.mgm.fr/num11/articles/art06304.html

OŠTIR, K., KOKALJ, Z., SALIGNY, L., TOLLE, F., NUNINGER, L., ZAKŠEK, K. and PENNORS. F. (2008). Confidence maps: a tool to evaluate archaeological data's relevance in spatial analysis. In: Posluschny, A., Lambers, K., and Herzog, I. (eds.). *Layers of Perception. Proceedings of the 35th International Conference on Computer Applications and Quantitative Methods in Archaeology (CAA), Berlin, 2-6 April 2007*. Bonn: Kolloquien zur Vor- und Frühgeschichte Vol. 10. p. 272-277.

POIDEVIN, D. (1999). *La carte, moyen d'action*. Paris: Ellipses.

GIETL, R., DONEUS, M., and FERA, M. (2008). Cost Distance Analysis in an Alpine Environment: Comparison of Different Cost Surface Modules. In: Posluschny, A., Lambers, K., and Herzog, I. (eds.). *Layers of Perception. Proceedings of the 35th International Conference on Computer Applications and Quantitative Methods in Archaeology (CAA), Berlin, 2-6 April 2007*. Bonn: Kolloquien zur Vor- und Frühgeschichte Vol. 10, pp. 336-341

SALIGNY, L., POULET-CROVISIER, N., and MORDANT, C. (2005). Les anamorphoses linéaires: une nouvelle approche pour l'analyse des flux de circulation des objets métalliques à l'Âge du Bronze. In: BERGER, J.F., BERTONCELLO, F., BRAEMER, F., DAVTIAN, G. and GAZENBEEK, M. (eds) (2005) *Temps et espaces de l'Homme en société: analyses et modèles spatiaux en archéologie, XXVe rencontres internationales d'archéologie et d'histoire d'Antibes, Antibes, 21-23 October 2004*. Antibes: APDCA, pp. 151-161.

TREMBLAY-CORMIER, L. (2014). Le mobilier métallique méditerranéen, témoin des échanges à longue distance entre Rhin et Rhône du Xe au Ve siècle avant notre ère. In: Barral, Ph., Guillaumet, J.P., Rouliere-Lambert, M.J., Saracino, M. and Vitali, D. (eds). *Les Celtes et le Nord de l'Italie (Premier et Second Ages du fer). Proceedings of the XXXVIth colloque international de l'A.F.E.A.F., Vérone, 17-20 May 2012*. Dijon: 36e supplément à la R.A.E. pp. 24-37.

Exploring Intervisibility Networks: A Case Study From Bronze and Iron Age Istria (Croatia and Slovenia)

Zoran ČUČKOVIĆ

Laboratoire Chrono-environnement (UMR 6249), Université de Franche-Comté, Besançon, France
cuckovic.zoran@gmail.com

Abstract

Using the rich archaeological record of 480 hillfort sites during the Bronze and Iron Age in Istria (Croatia and Slovenia), we have attempted to build a structured approach for investigating intervisibility networks. This analysis is organized in three steps. An exploratory analysis of visual connections is first carried out, particularly focusing on the relationship between the lengths of connection between sites and the network structure. Next, two new indices of integration of individual sites into the network are proposed and analyzed (connection success and visual neighborhood index). Finally, an approach for examining the statistical relevance of the observed network is proposed.

Keywords*:* Visibility Analysis, Intervisibility Network, Gis, Prehistory

Introduction

The first interpretation of prehistoric hillfort sites in Istria, a peninsula in the Northern Adriatic Sea (Figure 1) was proposed by the local historian Pietro Kandler in the mid-19th century. This promoted the idea of a vast network of mutually visible fortresses which were dated erroneously to the Roman period. Their visual connection would have enabled coordinated actions for defensive or commanding purposes (Kandler, 1869; Buršić-Matijašić, 2007: 577). This anecdote reveals a deep-rooted fascination for intervisibility phenomena, not only in Istria but in many other regions where hilltop sites abound (Crete – Soetens *et al.* , 2008; Bosnia – Benac 1985; Slovenia – Dular & Tecco Hvala, 2007; Spain – Brughmans *et al.,* 2014).

Intervisibility can be considered as a visual relationship between two socially meaningful entities: two persons, two communities, or even those of spirits or ancestors. These relationships constitute a particular type of social network, i.e. they facilitate and channel social contact (Hillier ,2005). Likewise, they participate in structuring the social space by forging visual references to particular places or constructions (Tilley, 1994, 6). In terms of the long standing debate over shortcomings of visibility analyses in landscape archaeology, that is, their limitation for understanding complex social and cognitive phenomena, the intervisibility analysis is particularly interesting because of both, its explicit addressing of intentionality or at least directionality of the gaze, and its focus on complex relationships (Frieman & Gillings, 2007; Llobera, 2012).

A crucial problem of intervisibility analysis is the relationship between the observed structure (the patterns of visual connections) and the practice which is related to the function of these connections. First, it cannot be denied that many societies, including those inhabiting prehistoric Istria, developed ‘strategies of visibility’ rendering certain aspects of social life more visible than others (e.g. funerary monuments [Criado, 1995]). However, this ‘will to visibility’ as F. Criado puts it, does not need to be conscious or rationalized (ibid). The structure may emerge as a result of complex social processes rather than direct intentionality, which is, perhaps, the essential argument of the structuralist approach (Murdoch, 2005). Hence there is a need for an in-depth analysis of the observed structure, in particular testing for the possibility of random emergence (cf. Wheatley, 1995; Brughmans *et al.*, 2014). We do, nevertheless, consider a functional hypothesis in the case of the Istrian intervisibility network: that the intervisibility network facilitated communication exchange. However, it should be noted that the examination of that hypothesis is a research incentive more than its presumed goal. The principal aim of the following analysis is to propose a methodological framework for the study of intervisibility networks as prerequisite for consideration of any functional interpretation.

The analysis is organised through a three-tiered scheme. First, we begin with an exploratory analysis of visual connections with particular attention to distances between sites. Next a connectivity analysis is performed, which is done using two new metrics, along with the standard index of betweenness centrality. Third, as a means of testing the relevance of the observed patterns (i.e. for their potentially random emergence), a background intervisibility network of all Istrian hilltops is modelled and compared against the observed network of archaeological sites.

1. Dataset and methodology

The prehistory of Istria is best known for castellieri, modestly sized Bronze and Iron Age hillforts (1 to 5 ha in surface), surrounded by one or several massive ramparts (Mihovilić, 2013). Large amounts of pottery, as well as settlement debris and traces of architecture found in excavations, indicate that the majority of sites were inhabited. Some sites display apparent indicators of higher status, either in terms of their size, imported goods or particularly rich necropoles that emerged in their vicinity (Hänsel *et al.*, 2009; Buršić-Matijašić, 2007). The density of the distribution of these sites in the region is remarkable — the distances between sites are less than 2 km on average (Table 1).

There are, however, many uncertainties regarding Istrian hillforts. The excavations tend to focus on large, high status sites, while little is known about small, less intensively inhabited sites. Many could have served as pastoral enclosures, especially in the mountainous and karstic areas (Slapšak, 1995), or even as observation outposts (Teržan & Turk, 2005). Another problem, related to the lack of research, is insufficient data on the chronology of the sites: the vast majority is dated roughly to 'protohistory' (Buršić-Matijašić, 2007). A recent study of the Northern part of Istria by M. Sakara-Sučević (2012) has shown that six out of nine sites examined could have been contemporary in the Middle or beginning of Late Bronze Age. In southern Istria, a vast majority of hillforts is supposed to be established in the Bronze Age, while fewer continued in the Iron Age (Hänsel *et al.*, 2009; Buršić-Matijašić, 2007). In any case, a conclusive survey is yet to be made, but broad contemporaneity of a large number of sites (roughly two thirds or more) may be hypothesized for the Middle to Late Bronze Age.

The dataset presented here has been compiled mainly from an exhaustive catalogue by K. Buršić-Matijašić (2007) for the Croatian part of Istria, and S. Poglajen (2007) for the Slovenian part. The 'hypothetical' group from the catalogue of K. Buršić-Matijašić has been retained in order to account for possible lack in the verified part of the dataset (Figure 1). The hypothetical sites tend to gather in the less researched interior of the peninsula, which may play a crucial role for assessing the integrity of the intervisibility network. All analyses were run parallel, with and without hypothetical sites.

Standard viewshed analysis, implemented in many GIS software packages, can be used for the purpose of analysing intervisibility relationships (Wheatley, 1995). However, this approach poses a significant computational overhead because the usual viewshed algorithm verifies all potential target points in a given radius, while in our case only a very limited number of target points is of interest. Massive viewshed calculations can take days of computing on a common personal computer (Llobera *et al.*, 2010).

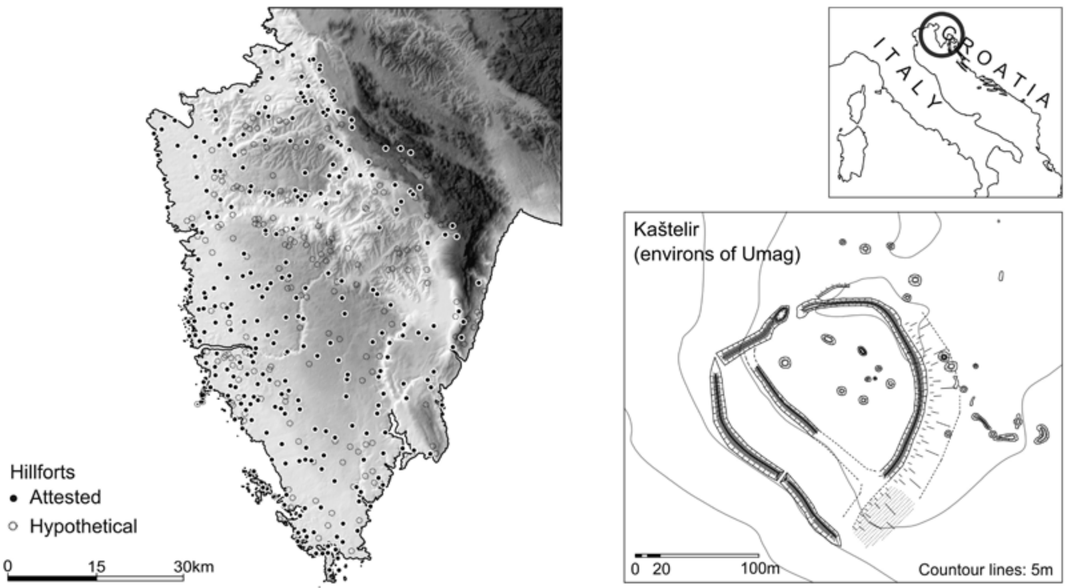


Figure 1: Hillforts of Istria with an example of a site (elevation data: SRTM).

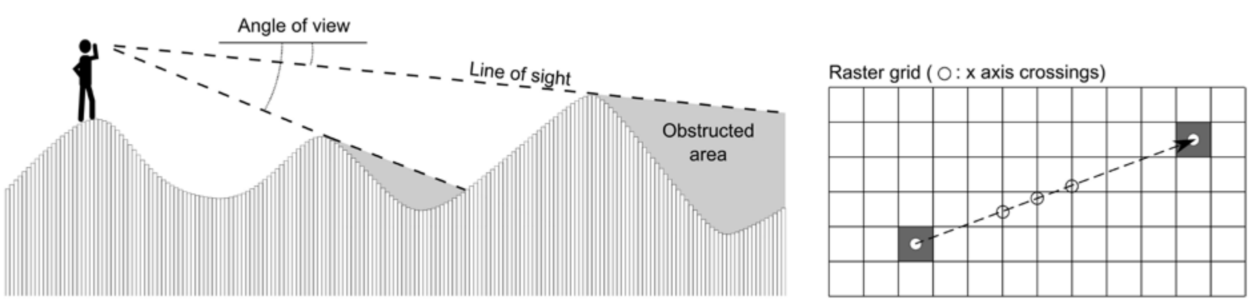


Figure 2: Intervisibility calculation.

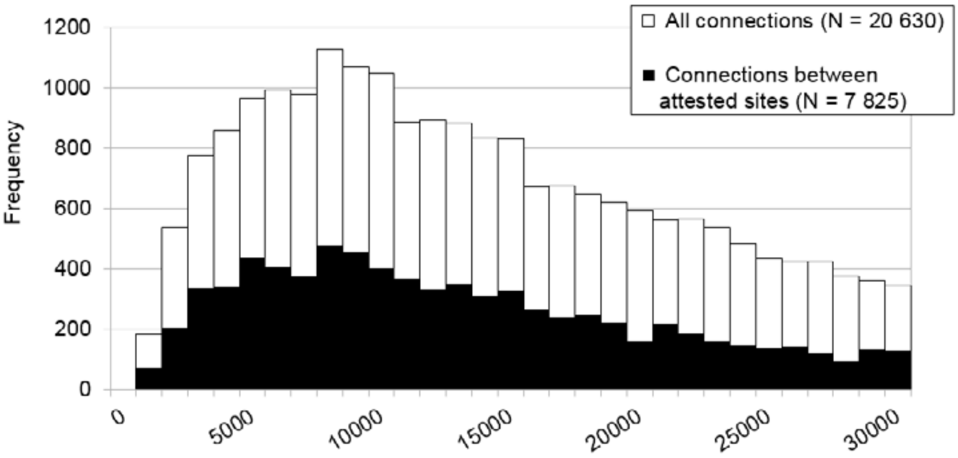


Figure 3. Histogram of in,tervisibility connection lengths.

For the purpose of efficient analysis as well as generation of an appropriate data format for network analysis (a list of pairs of connected nodes with some supplementary data), a customized intervisibility algorithm was implemented in QuantumGIS software (Čučković, 2014). The algorithm calculates the vertical angle of view towards each pixel traversed by a straight line connecting the observer and the target (Figure 2). The target is visible if situated above the maximum angle of view. Note that the straight line will rarely touch the centre of intermediary pixels (Figure 2, right): the height values are, therefore, interpolated between two neighbouring pixels on each side of the tested line, which is the common or 'exact' visibility calculation (Kaučič & Žalik, 2002).

Free SRTM data was used for the digital elevation model (DEM), with the resolution of three arc seconds, corresponding to 80 meters in the local projection (Farr *et al.,* 2007). More precise data of ASTER GDEM[1] (30 meters resolution), also available free of charge, are unfortunately very noisy in the studied area and not as detailed as might be expected for the nominal resolution. Although coarse resolution of the SRTM DEM is definitely problematic for typical viewshed analysis which deals with a handful of sites in a restricted area, it may be argued that large numbers of sites, studied over a large area (here 90 x 60 km), would bring some additional robustness to the analysis (cf. Fisher et al. 1997 for use of a 50 m DEM).

2. Assembling the network

The intervisibility relationships were calculated for observers with height (from eye level) of 4 meters, accounting for massive drystone ramparts recorded on many excavated sites (Hänsel *et al.*, 1999; Buršić-Matijašić, 2007:513). The observer point was placed in the approximate centre of each site. Only reciprocal visibility was considered, i.e. if site A sees site B, the same should be valid the other way around.[2]

The histogram displayed in Figure 3 summarizes lengths of visual connections between Istrian hillforts. The highest frequency of connections is in the range of 5 to 10 kilometers with a peak at 7 kilometers. Apparently, these are the ranges which contribute the most to the formation of the hypothesised network, which can be further verified by its visualisation (Figure 4). The network layout settles down to a shape corresponding loosely to the geographical shape of Istria at a 7 to 10 kilometer distance ('Force Atlas' layout algorithm implemented in Gephi software was used). That effect is less evident when hypothetical sites are filtered out, because those are more frequent in the interior of the peninsula, which plays a crucial role for the network integrity.

Simple as it may be, the first step in the analysis of the intervisibility connections reveals a crucial relationship between network integrity and the distance of connections. The network becomes a region-wide phenomenon when connections of 7 or more kilometres are allowed, while at 5 km most parts of Istria already get integrated. The range of 5 to 10 km is also the most frequent in the distribution of connection lengths, which may explain the ranges at which the intervisibility network would have most efficiently operated. This finding may be further used to define distance based ranges that shape contact and communication. The near range would be situated at less than 5 km from the site. Individual features of neighbouring sites can still be discerned, large ramparts in particular, and even acoustic communication can be established.[3] The middle range, on the other hand, which covers distance from 5 to 10 km, offers a larger number of visual connections so that potential messages may be passed on with fewer intermediaries over greater distances. Finally, in far ranges, at distances beyond 10 or 15 km, only large landscape features remain visible and communication

[1] ASTER GDEM is a product of METI and NASA.

[2] Non-reciprocal visibility relationships may be an important feature at close range, as in the case of surveillance systems (cf. Yekutieli 2006). However, at longer distances considered here such relations seem unlikely, especially if exchange of communication is hypothesized.

[3] Drumming, for example, can be very effective means of communication even beyond 5 km (Finnegan 2012, chap. 17)

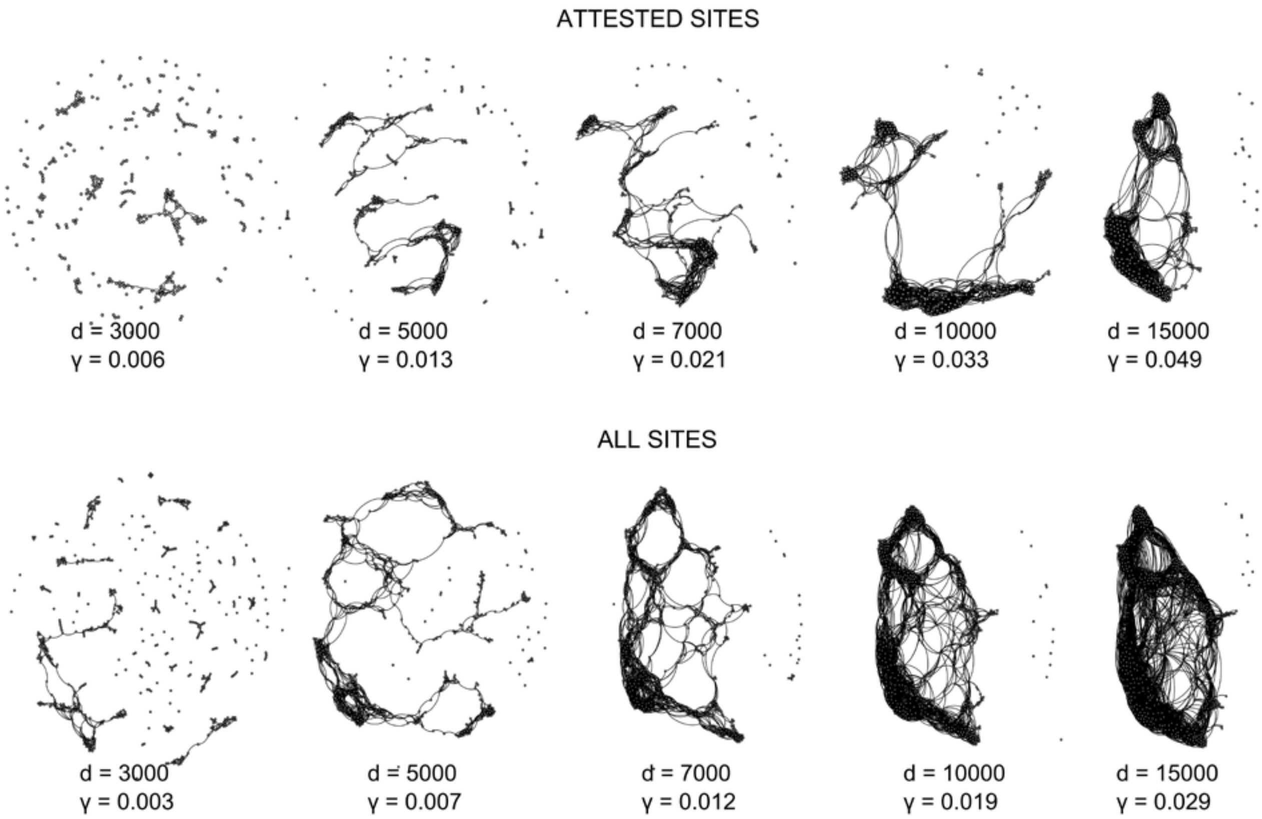


Figure 4: Progressive network integration in relation to increase of the maximum allowed connection distance (d, in meters). γ denotes graph density.
(Graph density is calculated by dividing the number of observed links by the maximum possible number of links (everybody connected to everybody - Newman 2010, chap. 6.9))

exchange becomes complicated. However, a practice of visual referencing of prominent (symbolically charged) features may still result in intervisibility networks.

3. Measuring connection

If a standard viewshed from a chosen observer point can be quantified by absolute values such as surface, average angle of view, shape etc., in the intervisibility network, which is composed of pairs of connected nodes exclusively, we have to rely on measures relative to the context of local or overall connections in the network. Additionally, the knowledge of spatial locations of analysed sites can be used to evaluate the connected sites against other available but disconnected sites in the vicinity (i.e. in the same radius of analysis; Figure 5).

Two basic intervisibility indices have been developed. Connection success of a site is defined as the percentage of the visible target sites in a set of evaluated target sites. A simple node degree (i.e. the number of connections per site), is less meaningful since it does not account for the number of available sites in the vicinity. Visual neighbourhood index is obtained by subtracting the median length of unsuccessful connections from the median length of successful connections. Negative values will, thus, indicate 'myopic' sites which keep their connections in the closer range, while sites having mostly far reaching connections will have positive index. The calculation is made on medians rather than on averages of distances because these values tend to have large variances: most commonly tested locations are spread all over the analysed zone. Note that both metrics are relative to the local context rather than to the overall configuration of the intervisibility network: their purpose is chiefly in measuring local responses to the potential for visual connection.

In order to study the integration of individual sites into a possible global flow in the network, one classic metric has been chosen as well: betweenness centrality. This metric is calculated by counting how many shortest (or optimal) paths between all pairs of nodes in a network pass through each individual node (Newman, 2010, chap. 7.7) It is very useful for detecting 'bottleneck' nodes which act as bridges connecting different clusters in the network (also known as 'hub nodes'). In the case of the Istrian intervisibility network, the betweenness centrality may help to locate sites having an important role in maintaining the integrity of the network.

Our indices were, then, obtained using the intervisibility algorithm at the search radius of 7.5 km, which is in the hypothesized optimal distance range (supra). The output produced by the algorithm, a table of both successful and unsuccessful connections, makes the calculation of proposed intervisibility indices a rather straightforward procedure. The histogram of connection success (Figure 6) displays a slight bimodal distribution, as if there may be a group of better and a group of less connected sites. The visual neighbourhood index mostly in the negative range, indicating (expectedly) that most of the sites keep their connections in the closer range. However, there are some ‘long-sighted’ sites which stand out from this ‘myopic’ majority, and which tend to connect farther away within their 7.5 km neighbourhood. In the scatter plot on Figure 6 the individual sites are distributed according to the two indices. No tendency of correlation between the success of connection and visual neighbourhood index is apparent. The spread of points indicates that even sites with long connections do not necessarily have a good success.

The relevance of these abstract metrics for understanding real-world archaeological sites can be illustrated by the two following examples (Figure 7). The site of Krug has both success of connection and visual neighbourhood index in the medium range. The hillfort itself is small to medium sized (1.8 ha) and single walled, situated on a dominant hilltop in the north of Istria (Buršić-Matijašić, 2007: 441). Its dating is not quite certain, but the surface pottery assemblage seems to indicate the Bronze Age. Krug would be an ordinary Istrian hillfort if it didn’t have one of the highest betweenness centrality values in the intervisibility network (especially when the maximum connection distance is set above 7 km) (Figure 6). The second example, located at the lower left of the scatter plot where sites with poor connection success and short connection distances gather, is the site of Kaštelir near Nova Vas (Figure 7). The site is exceptionally large for

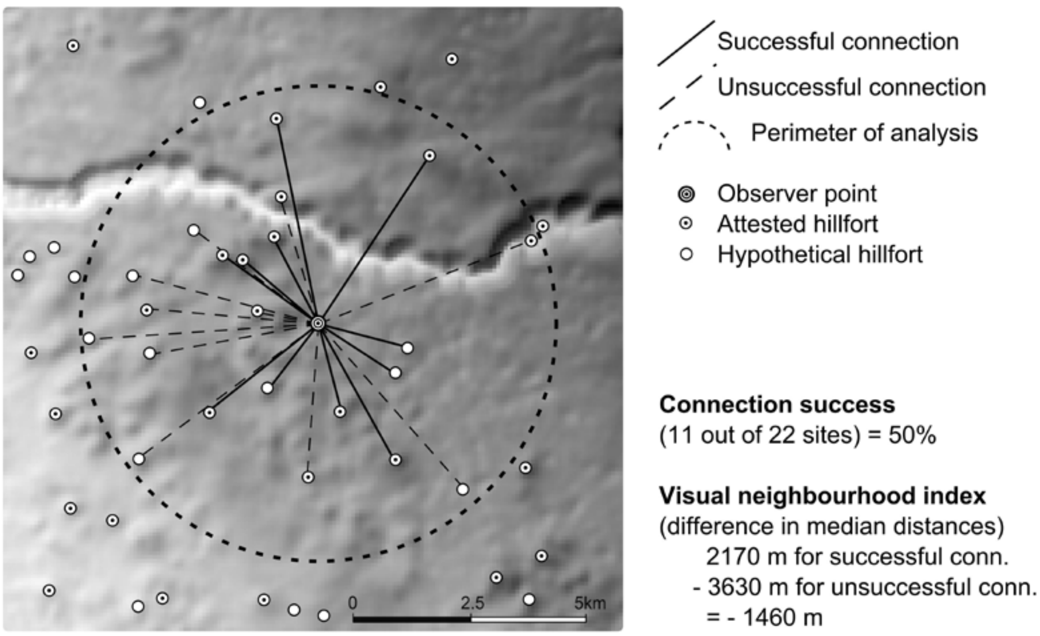


Figure 5: Calculation of intervisibility indices: the observer point corresponds to Maklavun tumulus near Rovinj (Hänsel et al. 2009).

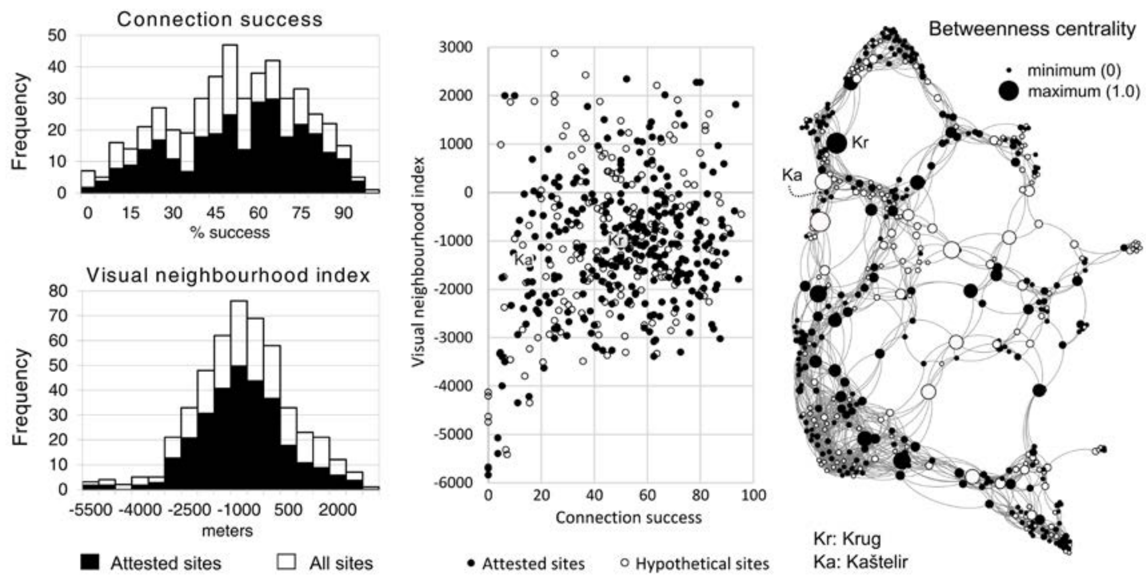


Figure 6: Intervisibility indices for analysis radius of 7.5 km.

Istrian prehistory (13 ha), and after a Bronze Age phase it became a local centre in the Iron Age (Sakara-Sučević, 2004). A glance on its topographic position – surrounded on all sides by higher ground – belies bad placement in terms of visual dominance. In fact it has only 5 visual links and measures of its connectivity are insignificant. However, the site is connected to two major hubs in terms of their betweenness centrality (Figure 6: just above and below the site), albeit both are considered as hypothetical sites.

While in the case of Krug the visual connections may have played a role in the choice of the location of the site (note the high betweenness centrality), this is definitely not true for the important settlement of Kaštelir. Apparently, intervisibility connections (or visibility in general) are not directly related to the settlement status. In fact, well connected sites tend to occupy high spots which are more often than not surrounded by rugged and poor terrain, less attractive for agricultural practices. Some of those may even have had surveillance and intervisibility connection as their primary function, as in the case of the probable Iron Age observation tower from Ostri Vrh in Slovenia.[4] In effect, the evidence of complex surveillance strategies in the Iron Age suggests that larger sites, such as Kaštelir, could still be well connected to the network through particular, intermediary sites. Such a higher level of complexity still remains to be tackled.

4. Testing for relevance

By scattering observers in the landscape according to a spatially random process, some of the observers should see each other thus creating an intervisibility network. Landscape has inherent visual connectivity. The observed prehistoric network should, then, differ significantly from such a random set if it has been shaped by particular, non-random processes.

A very popular method for statistical testing of intervisibility was proposed by David Wheatley. Essentially, his method uses viewshed surfaces generated from the analysed

[4] The tower is a massive drystone structure of oval shape, 11 meters long, with walls 1.5 to 2.5 meters in thickness, perched on top of a steep, rocky hill. Calibrated radiocarbon dates are in the range of 8th to 5th century BC. The excavators interpret the structure as an element of a surveillance system over an important land passage (Teržan & Turk 2005).

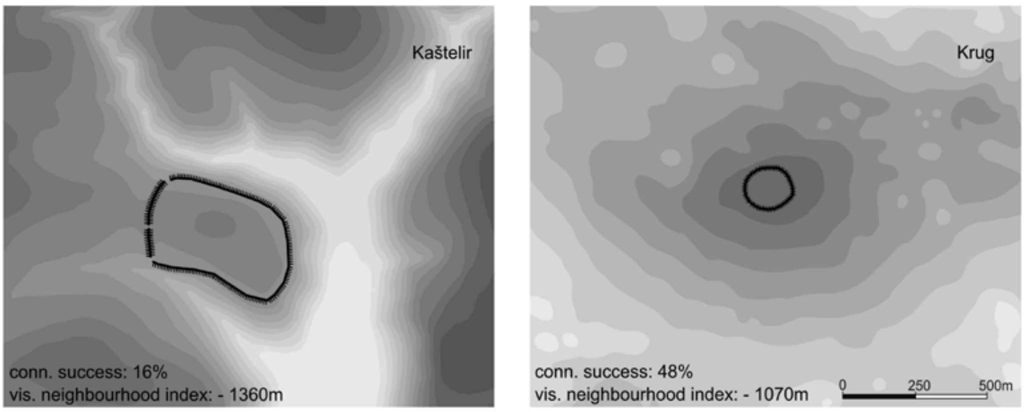


Figure 7: Sites Kaštelir near Brtonigla and Krug near Buje (same scale).

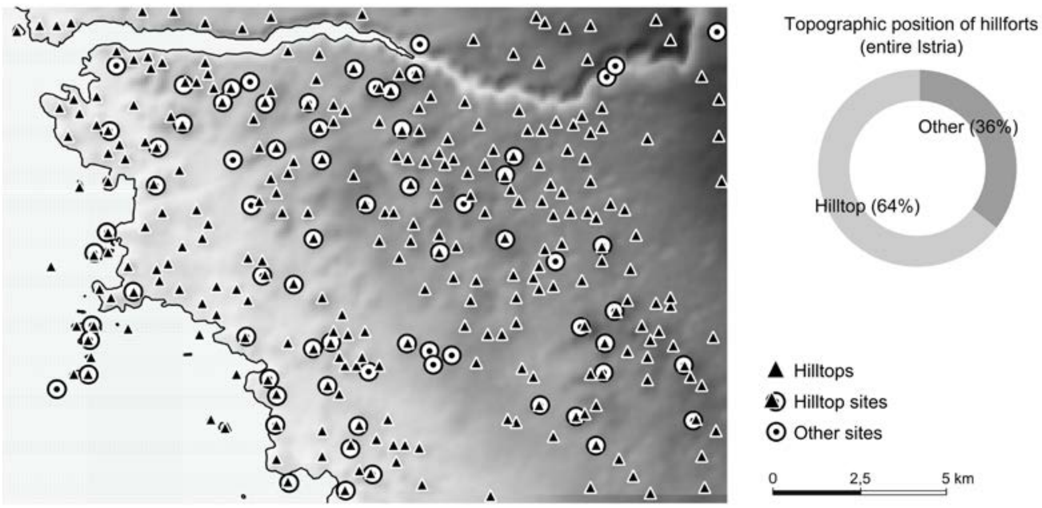


Figure 8: Correspondence of hilltops and prehistoric hillforts (map coverage: surroundings of Rovinj).

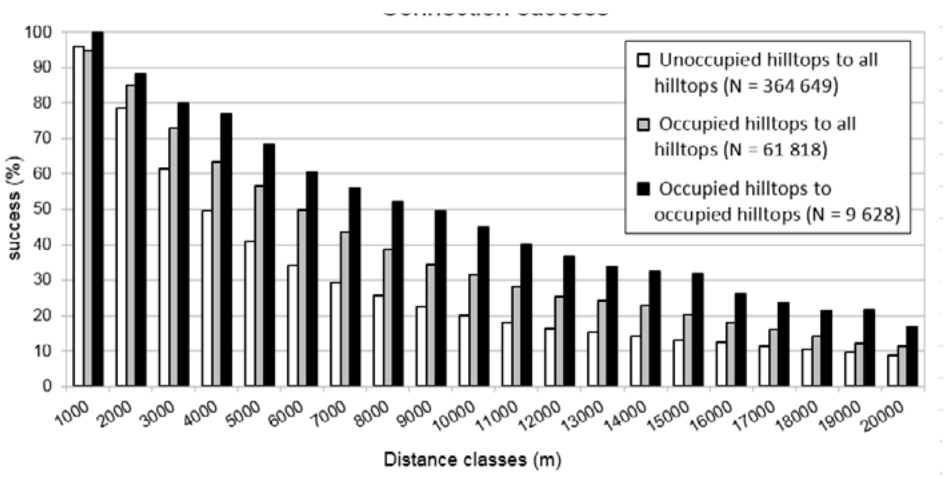


Figure 9: Connection success of analysed hilltop population.

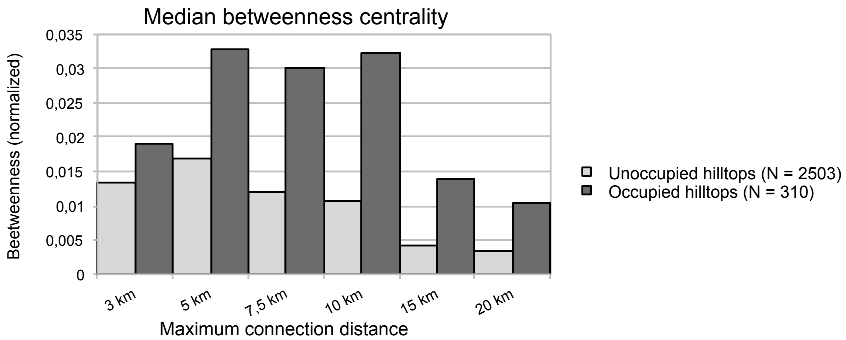


Figure 10: Median betweenness centrality for distance based models of the hilltop intervisibility network.

locations (i.e. sites) and tests whether the same sites tend to gather in zones with dense viewshed overlaps (Wheatley, 1995). However, that method operates in a uniform, indiscriminate space; the null hypothesis is that sites had equal chances of being placed anywhere in the landscape. That may work for some types of sites, such as prehistoric tumuli analysed by Wheatley, but not for hillforts which, by definition, occupy hills or elevated positions.

In order to develop a test more appropriate for our dataset, the background population of possible hillfort locations was reduced exclusively to hill-tops. These are defined as locations (here pixels in the DEM) which have the highest height value in a radius of 250 meters. Setting this radius to smaller values (equivalent to one or two pixels size) will likely result in selecting insignificant 'noise' in the DEM.[5] Close to two thirds of site locations (64%) correspond to the applied hilltop criteria (Figure 8). Apparently, many hillforts occupy other types of topographic positions, such as ridges or spurs, which have to be isolated by more sophisticated methods.[6] Nevertheless, the obtained sample of 2.813 hilltop sites, of which 310 occupied during the period under study, seems solid enough for statistical analysis, even if obviously biased by virtue of applied topographic criteria.

The intervisibility analysis was repeated on the sample of extracted hilltops, but with the observer and target heights set to two meters, considering these locations to be devoid of particular constructions (i.e. higher than half a metre). The same observer height was used regardless of the eventual prehistoric occupation of hilltops, in order to provide an unambiguous basis for statistical comparison. The complexity of the obtained network is enormous: more than 400 000 visual connections for the analysis radius of 20 km. Regarding the presence of archaeological evidence on the observer and target locations, three types of connections are considered here: 1) unoccupied hilltops to all hilltops, 2) occupied hilltops to all hilltops, and 3) occupied hilltops to each other.

The histogram on Figure 9 shows the connection success (i.e. the percentage of successful links) for one-kilometre wide distance ranges.[7] A strong tendency towards better connection between occupied hilltops is apparent. The statistics of betweenness centrality was calculated for six network models, each defined by the maximum distance allowed for establishing visual connection (Figure 10). These values have very large variances, and their

[5] The noise, i.e. insignificant topographic prominences or simply fluctuations of the DEM, is difficult to avoid in any case. For the purpose of our analysis 'hilltops' less than a meter in height and situated in completely flat areas were manually deleted.

[6] Another problem is the definition of hillfort which is inevitably ambiguous. Preferably, it is 1) relatively large site, 2) delimited by massive rampart and 3) situated on an elevated position. However, not all of these criteria are met in cases such as coastal hillforts or small, possibly seasonally occupied hillforts (cf. Buršić-Matijašić 2007, 485ff.)

[7] Previously discussed connection success index evaluates the success of individual sites while here the links are grouped according to their length, regardless of sites they connect.

distribution is extremely skewed, which is otherwise typical of this metric (Newman, 2010: chap. 7.7). Therefore, medians rather than means were used. Again, occupied hilltops differ from unoccupied ones, in particular after 5 km, which corroborates with the previous observation that the visual connections tend to concentrate in the range from 5 to 10 km. In order to verify statistically the observed differences in betweenness centrality the non-parametric Wilcoxon-Mann-Whitney test, which compares median differences of two samples, was used (Table 2). A very small (< 0.2) effect size of all differences should be noted, which indicates that these populations still resemble each other very much, even if their differences are statistically significant ($p < 0.05$).[8]

Further clues on the character of the intervisibility network can be obtained through graph theory, namely by examining the statistical distribution of node degrees (i.e. number of links per node). The simplest networks, formed by the random selection of links created with a set probability (called Bernoulli or Erdős-Rényi random graphs), will have a binomial or Poisson degree distribution (Newman 2010, chap. 12.3). Indeed, that is precisely the case with the intervisibility network of all hilltops (Figure 11). Much the same is valid for the population of all hillforts in the dataset, but the statistic becomes more interesting when hypothetical sites are eliminated. A tendency towards dispersion in degree distribution can be discerned in the group of attested hilltop sites (Figure 11). The network of hilltop sites, thus, features a higher proportion of well-connected nodes (with normalized degrees above 0.5), at least in comparison with the network of all hilltops. However, the exact relationship between these networks, including possible differences from a random pattern, should be tested by more advanced methods (see esp. Brughmans *et al.*, 2014).

It can be concluded now that hilltop sites play a key role in the connectivity of the network and that intervisibility relationships did influence choices of site location, which seems less clear for other, non-hilltop archaeological sites. Indeed, the bias introduced by selecting hilltop sites enabled at the same time to filter out sites with poor connection potential.

Conclusion and perspectives

It is hoped that the analysis of Istrian prehistoric hillforts demonstrated the existence of a web of intervisibility relations as a distinct cultural phenomenon. A consistent tendency towards settlement locations offering better intervisibility connections is discernible, particularly at the middle distance range (between 5 and 10, at maximum 15 km). Distance is apparently a key variable for understanding intervisibility networks, not only in terms of presumed function (see below), but also in terms of the potential for visual communication offered by the landscape. Nevertheless, intervisibility studies often present network models with connections ranging from 0 to beyond 30 km (e.g. Dular *et al.*, 2007; Soetens *et al.,* 2008). In the outlook, an exploration of interplay between near, medium and far distance ranges (as proposed above) in the intervisibility network is envisaged, aided by modelling networks where connections are weighted by distance.

Two new indices for the insertion of sites into the intervisibility network are proposed: the connection success and the visual neighbourhood index. Both make use of data that are absent from the intervisibility network, namely the unsuccessful connections to sites which are close by but hidden from sight. Such an approach is explicitly local inasmuch as it evaluates individual responses of sites to the intervisibility network. Consequently, the proposed

[8] The effect size was calculated as $r = z / \sqrt{N}$, where z is the z-score and N the number of all observations (Fritz et al. 2012). wilcox.test and wilcox_test functions from R software, with paired option set to False, were used for obtaining these values. The basic assumption of the Wilcoxon-Mann-Whitney test is that groups have similar distributions, while conformity to normal distribution is not required (Wilcox 2009, 273).

Radius of analysis	3 km	5 km	7.5 km	10 km	15 km	20 km
W	342391	304235	284040	276330	268640	275264
p	0,0007562	5,72E-010	< 2,2e-16	< 2,2e-16	< 2,2e-16	< 2.2e-16
Effect size (r)	0,06	0,12	0,15	0,16	0,17	0,16

Table 2: Results of Wilcoxon-Mann-Whitney test for betweenness metric.

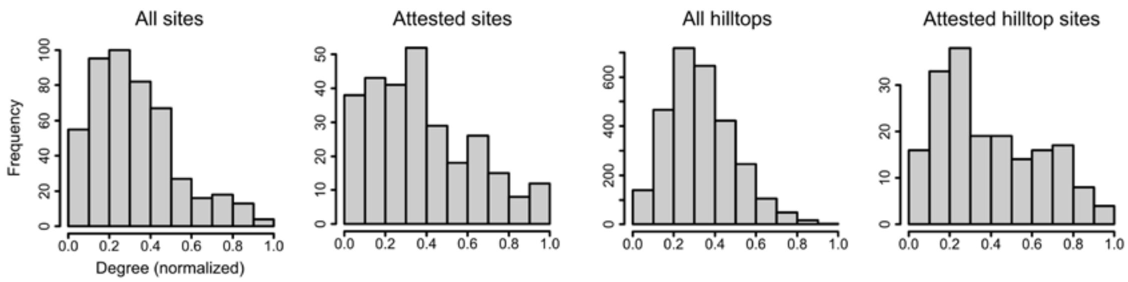


Figure 11: Degree distributions for networks with maximum connection distance of 7.5 km.

indices may be better suited for evaluating sites in their local context than the node degree metric, as demonstrated on cases of two hillforts. Many more parameters can be evaluated in the same manner, such as azimuths of connections or types of sites engaged. Such elaborate data may be used to investigate an experiential context of a site (i.e. visual references with the surrounding landscape), as well as to nuance the often globalising approach of the network analysis.

Testing for statistical relevance of the intervisibility network is not an easy endeavour. It was not possible to use the elegant and very popular solution by D. Wheatly (1995) which tests the observed population against uniform spatial distributions – our sites are normally on hilltops. All Istrian hilltops were therefore isolated as the background population of potential sites and analysed using the intervisibility algorithm. This approach is useful, but comes with an important caveat. Moving away from the expectation of uniform spatial distribution, we're stepping into predictive modelling (Verhagen & Whitley, 2012). That would certainty introduce a new scale of complexity, but would also enable evaluating intervisibility against other factors in the choice of site locations, such as topography, environment or social relationships.

The question of the use to which the intervisibility network had been put was not tackled. Formal characteristics of the network (such as its density and positive tendency for higher betweenness centrality) might be interpreted in favour of a communication exchange network, for example by smoke and fire.[9] Nevertheless, there is still no archaeological evidence in its support (beacons, particular traces of burning etc.). Moreover, some other processes may give rise to a well-connected network, such as the establishment of visual references to other occupations in the landscape, with or without intention for regular communication (cf. Marrou & Rousseaux, 2009). In any case, these hypotheses need to be tested by more sophisticated approaches. An interesting approach is recently introduced Exponential Random Graph Modelling which is based on the generation probabilistic networks (i.e. where connections are defined by varying probabilities) (Brughmans *et al.*, 2014). Its application on the intervisibility network of Iron Age sites in the Southern Spain has enabled researchers to propose the tendency of small scale, local clustering of visual connections around dominating sites as the crucial process in the emergence of the network (idem, 452).

The aim of this paper is in presenting a structured approach to intervisibility network analysis, working from the ground level up, that is, from exploratory data analysis to more complex issues. Not finding an ultimate interpretation for the functioning of the intervisibility network is not a weakness. Intervisibility studies are often accompanied with problematic concepts and quick conclusions, which have been subject to critique (cf. Briault, 2007). Diving deeper into the complexity of visual networks and examining them in their own right is crucial if these critiques are to be confronted.

[9] Smoke and fire are often mentioned in the ancient sources (Homer, Illiad, XVIII, 210-213; Aeshylus, Agammemnon, 1-34). A fire signalling network has also been recently proposed for Bronze Age Crete (Sarris et al. 2011).

Acknowledgements

I would like to thank Elise Fovet, Tom Brughmans, Phillip Verhagen and Estelle Gauthier for helpful comments and suggestions. Earlier research on intervisibility was carried out within a M.Sc. project under supervision of Estelle Gauthier at the University of Franche-Comté (Besançon). The text has benefited from corrections by Stefani Crabtree.

References

BENAC, A. (1985). *Utvrđena ilirska naselja (I.) (delmatske gradine na Duvanjskom polju, Buškom blatu, Livanjskom i Glamočkom polju). Djela, knjiga LX, Centar za balkanološka ispitivanja.* Sarajevo: Akademija nauka i umjetnosti BiH.

BRIAULT, C. (2007). Making mountains out of molehills in the Bronze Age Aegean: visibility, ritual kits, and the idea of a peak sanctuary. *World Archaeology* 39(1), pp.122-141.

BRUGHMANS, T., KEAY, S., & EARL, G. (2014). Introducing exponential random graph models for visibility networks. *Journal of Archaeological Science* 49, pp.442-454.

BURŠIĆ-MATIJAŠIĆ, K. (2007). *Gradine Istre: povijest prije povijesti.* Pula: ZN 'Žakan Juri'.

CRIADO, F. (1995). The Visibility of the Archaeological Record and the Interpretation of Social Reality. In I. Hodder *et al.* (eds). I*nterpreting Archaeology: Finding Meaning in the Past*, London; New York: Routledge. pp.194-204.

ČUČKOVIĆ, Z. 2014. Viewshed analysis 0.2 (Quantum GIS plugin). Available from: https://plugins.qgis.org/plugins/ViewshedAnalysis/. [Accessed July 16, 2014].

DULAR, J., & TECCO HVALA, S. (2007). South-Eastern Slovenia in the Early Iron Age: settlement, economy, society = *Jugovzhodna Slovenija v starejši železni dobi : poselitev, gospodarstvo, družba. Opera Instituti Archaeologici Sloveniae 12.* Ljubljana: Založba ZRC.

FARR, T.G. ET AL. (2007). The Shuttle Radar Topography Mission. *Reviews of Geophysics* 45(2). pp.1-33.

FINNEGAN, R. (2012). *Oral Literature in Africa.* Cambridge: Open Book Publishers.

FISHER, P., FARRELLY, C., MADDOCKS, A., & RUGGLES, C. (1997). Spatial Analysis of Visible Areas from the Bronze Age Cairns of Mull. *Journal of Archaeological Science* 24(7), pp.581-592.

FRIEMAN, C., & GILLINGS, M. (2007). Seeing is perceiving? *World Archaeology* 39(1), pp.4-16.

FRITZ, C.O., MORRIS, P.E., & RICHLER, J.J. (2012). Effect size estimates: current use, calculations, and interpretation. Journal of Experimental Psychology. General 141(1), pp.2-18.

HÄNSEL, B., MATOŠEVIĆ, D., MIHOVILIĆ, K., & TERŽAN, B. (2009). O socijalnoj arheologiji brončanodobnog utvrđenog naselja i grobova na Monkodonji. *Histria archaeologica* (38-39), pp.81-122.

HÄNSEL, B., MIHOVILIĆ, K., & TERŽAN, B. (1999). Monkodonja, utvrđeno protourbano naselje starijeg i srednjeg brončanog doba kod Rovinja u Istri. *Histria archaeologica* (28), pp.37-107. [Also published in German in Prähistorische Zeitschrift 1999, 74]

HILLIER, B. (2005). The art of place and the science of space. World Architecture 185, pp.96-102.

KANDLER, P. (1869). [No title]. La provincia III (5): pp.291-292.

KAUČIČ, B., & ŽALIK, B. (2002). Comparison of viewshed algorithms on regular spaced points. In A. Chalmers (ed.). *Spring Conference on Computer Graphics 2002, 24–27 April 2002, Bratislava.* Bratislava: Comenius University. p. 177–183.

LLOBERA, M. (2012). Life on a Pixel: Challenges in the Development of Digital Methods within an 'Interpretive' Landscape Archaeology Framework. *Journal of Archaeological Method and Theory* 19(4), pp.495-509.

LLOBERA, M., WHEATLEY, D., STEELE, J., COX, S., & PARCHMENT, O. (2010). Calculating the inherent visual structure of a landscape (inherent viewshed) using high-throughput computing. In: F. Niccolucci & S. Hermon (eds). *Beyond the artefact: Digital Interpretation of the Past. Proceedings of CAA2004 Prato, 13–17 April 2004*. Budapest: Archaeolingua.

MARROU, L., & ROUSSEAUX, F. (2009). Covisibilité et peuplement aux Açores. Estimation de l'importance du facteur de covisibilité dans la répartition des populations aux Açores. *Cybergeo : European Journal of Geography*. Available from: http://cybergeo.revues.org/22725. [Accessed July 16, 2014].

MIHOVILIĆ, K. (2013). Castellieri-Gradine of the Northern Adriatic. In: H. Fokkens & A. Harding (eds). *The Oxford handbook of the European Bronze Age,* Oxford: Oxford University Press. pp. 864-876.

MURDOCH, J. (2005). *Post-structuralist Geography: A Guide to Relational Space.* London; Thousand Oaks, Calif: SAGE Publications Ltd.

NEWMAN, M. (2010). *Networks: An Introduction.* Oxford: Oxford University Press.

POGLAJEN, S. (2007). *Geografski informacijski sistemi v študijah rimskega podeželja: primer severozahodne Istre.* PhD thesis. University of Ljubljana.

SAKARA SUČEVIĆ, M. (2012). *Prazgodovinska keramika med Miljskim zalivom in porečjem Mirne.* PhD thesis. Koper: University of Primorska.

SAKARA-SUČEVIĆ, M. (2004). *Kaštelir. Prazgodovinska naselbina pri Novi Vasi - Brtonigla.* Koper: Univerza na Primorskem, Znanstveno-raziskovalno središče, Inštitut za dediščino Sredozemlja.

SARRIS, A., KAKOULAKI, G., & PANAGIOTAKIS, N. 2011. Reconstructing the Minoan Communication Network in Central Crete. In: E. Jerem, F. Redő, & V. Szeverényi (eds). *On the Road to Reconstructing the Past. Proceedings of CAA2008, Budapest, 2–6 April 2008.* Budapest: Archaeolingua. pp. 503-511.

SLAPŠAK, B. (1995). Možnosti študija poselitve v arheologiji. Arheo (17), pp.1-90.

SOETENS, S., SARRIS, A., & VANSTEENHUYSE,, K. (2008. Between Peak and Palace Reinterpretation of the Minoan Cultural Landscape in Space and Time. In: *Proceedings of the 4th Symposium of the Hellenic Society for Archaeometry. National Hellenic Research Foundation, Athens, 28-31 May 2003.* BAR Int. 1746, Oxford: Archaeopress. pp. 153-161.

TERŽAN, B., & TURK, P. (2005). The Iron Age tower upon Ostri vrh. In: G. Bandelli & E. Montagnari Kokelj (eds). *Carlo Marchesetti e i castellieri, 1903-2003. Fonti e studi per la storia della Venezia Giulia, Ser. 2, Studi, vol. 9*, pp. 339-352. Trieste: Editreg

TILLEY, C. (1994). *A phenomenology of Landscape: Places, Paths, and Monuments.* Oxford: Berg.

VERHAGEN, P., & WHITLEY, T.G. 2012. Integrating Archaeological Theory and Predictive Modeling: a Live Report from the Scene. *Journal of Archaeological Method and Theory* 19(1), pp.49-100.

WHEATLEY, D. (1995). Cumulative Viewshed Analysis: a GIS-based method for investigating intervisibility, and its archaeological application. In G. R. Lock & Z. Stančić (eds). *Archaeology And Geographic Information Systems: A European Perspective,* London: Routledge, pp. 5-13.

WILCOX, R. (2009). *Basic Statistics : Understanding Conventional Methods and Modern Insights: Understanding Conventional Methods and Modern Insights.* Oxford: Oxford University Press.

YEKUTIELI, Y. (2006). Is Somebody Watching You? Ancient Surveillance Systems in the Southern Judean Desert. *Journal of Mediterranean Archaeology* 19(1), pp.65-89.

Chapter 7. Mathematics and Statistics in Archaeology

Intentional Process Modeling of Statistical Analysis Methods

Charlotte HUG, Rebecca DENECKÈRE, Ammar AYMEN

Université Paris 1 Panthéon-Sorbonne, Centre de Recherche en Informatique, Paris, France
charlotte.hug@univ-paris1.fr, rebecca.deneckere@univ-paris1.fr, ammar.aymen@gmail.com

Abstract
Each Humanities researcher has its own way to deal with data (collection, coding, analysis and interpretation). All these specific ways of working are not shared - each researcher is reinventing his/her own method while analyzing data without any previous experience. Nevertheless, developing and sharing these methods should be useful to the research community and students in Humanities. Moreover, a lot of data analysis is done with statistical analysis methods, to find correlations between events, to make predictions or assumptions on facts or artifacts; and the use of one method or another requires good statistical knowledge. We conducted interviews among archaeologists and historians to understand their ways of working and collected information on the methods they used. We then built a method to guide researchers in using statistical analysis methods.

Keywords: statistical analysis methods, method, intentional process

1. Introduction

Humanities include the sciences having for objective to study the human beings: their actions, their relationships as well as their traces (human cultures, lifestyles, social behaviors, societies). Studying the human beings lifestyles and their social behaviors means looking for and analyzing data that can be described under various formats or written in foreign languages, whether in the form of texts or drawings. This data can be large-sized and stored under various forms. The development of the computing tools allowed the preservation and the exploitation of big masses of data and numerous statistical methods were designed since the beginning of the 20th century to analyze them.

Statistical analysis methods are widely used in Humanities in order to find correlations between events, to make predictions or assumptions on facts or artifacts (Canning 2014). However, the use of statistical analysis methods requires previous knowledge in statistics, selecting one method or another depends on the type of data, the project's criteria, and the objectives of the researcher. The selection of a method has then a great impact on the obtained results, therefore on the interpretation. Using an unsuitable method can then lead to misinterpretations which have to be avoided at all cost for research sake. Statistical tools as R (R 2014) or SAS (SAS 2014) allow to manipulate data and to run statistical methods but researchers need to be guided during this process and during the interpretation of the obtained results, especially if they are beginners. Moreover, humanities researchers have their own ways to deal with data: the collection, coding, analysis and interpretation are often specific to each researcher. These methods are not shared - each researcher is reinventing the wheel when he/she analyzes data without previous experience. Nevertheless, modeling and sharing these methods could be useful to the research community and students in Humanities, but it is difficult to formalize them in an understandable way.

Our objective is to provide guidance to any humanities actor, from students to researchers while conducting data analysis.The provided guidance must fit the situation at hand: each project has its own constraints and specificities. It is necessary to provide a method that is flexible to better guide and support the actors.

In this paper, we present a method to guide researchers to use statistical analysis methods. This method is described as a flexible process based on the intentions of the researcher. This paper is organized as follows, section 2 presents the followed research methodology, section 3 briefly introduces the modelling language we used to represent the method and the process models of the statistical analysis methods, section 4 presents the proposed method and section 5 concludes the paper.

2. Research methodology

We first conducted a literature review on the main statistical methods used in Humanities. We found that the Principal Component Analysis (Pearson, 1901; Hotelling, 1933), the Correspondence analysis (Benzecri, 1982), the Multiple Correspondence Analysis (Benzécri, 1973), the Hierarchical Clustering (Sokal & Sneath, 1963) and the Logistic Regression (Berkson, 1944) were the most used statistical analysis methods.While we conducted the literature review and studied the different methods, we defined in parallel the corresponding process models using Map (Rolland, Prakash &Benjamen 1999), an intentional process modeling language (see section 3).

We then conducted interviews with researchers of the university Paris 1 Panthéon-Sorbonne from different fields: anthropology (Subject 1), medieval history (Subject 2), social history(Subject 3) and statistics. The objectives of these interviews were: to understand and analyze the statistical methods Humanities researchers use, to understand what were the methods they useto collect data, to understand their ways of working and the projects they were working on and to establish a generic method to guide the use of statistical analysis methods. We built a questionnaire comprising 12 questions; the interviews lasted around 2 hours. Table 1 presents the questions asked to the researchers.

We present below a synthesis of the answers to questions 2 to 10 of the three subjects.

- Q1 and Q2 were introductory questions to put the subjects at ease.
- Q3: what is the followed process to carry out a project? The subjects start by defining the problematic and its interest by analyzing the sources (feasibility, reliability). Then they design a database and encode the data to do statistical analysis. They select the variables to analyze and analyze the results from the point of view of the domain (as archaeologist or historian).
- Q4: how do you formulate a research hypothesis? For Subject 2, the hypothesis is formulated from the sources, from reading on a particular topic, or by attending a seminar. For Subject 3, it consists in determining the research space, searching for personal data and formulating the hypothesis when the data is encoded.
- Q5: what are the methods you use to collect data? Where and how do you collect data? How do you know which data to collect? Subject 1 collects two types of data: the intrinsic data which is extracted from the remains or the objects themselves and the contextual data that concerns the environment in which the object was found. Subject 2 and 3 use archives, documentary research (paper, image or multimedia), online data (collaborative tools), archaeological traces.... Subject 2 structures the information by transcription (Excel, csv…) and he selects the variables until finding the good ones. Subject 3 select the data on what interest him in relation to the research project and determine whether this data can be used or not and how.
- Q6: what are the statistical analysis methods you frequently use? The usual statistical analysis methods used by the subjects are ACP, CAH, ACM AFM, and more and more often sequential analysis (subject 2).
- Q7: upon which criteria do you select one statistical method to analyze your data? Subject 1 and 2 state the selection of the statistical method is based on the nature of the variable. The approach of Subject 3 is empirical as he tests the methods and tries to exploit the results.
- Q8: upon which criteria do you base yourself to select the variables to analyze? Subject 1 uses variables related to his research topic. Subject 2 tends to use all the variables to determine two different groups of population to apply statistical methods on the group where the data is complete and the other one where it will be possible to determine the hypothesis. Subject 3 first validates the variables through khi-2 test and through his knowledge of the topic.
- Q9: how do you interpret the results obtained by statistical analysis? Subject 1 uses the chronology given by the results to interpret them. Subject 2 and 3 state that they do not use any particular methods to interpret the results of statistical analyses.
- Q10: do you use other statistical analysis to confirm the results of a first analysis? Other statistical analysis methods can be used to confirm and balance the obtained results.
- Q11: how do you take into account the incomplete data during the analysis? Subject 1, as an archeologist, tends to use samples that are carefully selected and classified to be the best as possible. Subject 2 analyses the incomplete data separately and Subject 3 explains that some colleagues use

Q1: could you present yourself (function and research domain)?
Q2: what are the projects you are working on?
Q3: what is the followed process to carry out a project?
Q4: how do you formulate a research hypothesis?
Q5: what are the methods you use to collect data ? Where and how do you collect data? How do you know which data to collect?
Q6: what are the statistical analysis methods you frequently use?
Q7: upon which criteria do you select one statistical method to analyze your data?
Q8: upon which criteria do you base yourself to select the variables to analyze?
Q9: how do you interpret the results obtained by statistical analysis?
Q10: do you use other statistical analysis to confirm the results of a first analysis?
Q11: how do you take into account the incomplete data during the analysis?
Q12: could you look at this model and tell us if there are things you do or not? Do you do something else?

TABLE 1: THE QUESTIONNAIRE PRESENTED TO THE RESEARCHERS.

smoothing techniques to deal with incomplete or incoherent data, but he prefers to isolate them.

- Q12: could you look at this model and tell us if there are things you do or not? Do you do something else? All the subjects recognized their ways of working in the presented models.

The main intentions were represented and the strategies to achieve them were adequate. However, the defined terminology was not always clear.

These interviews then helped us to better understand the researcher problems while analyzing data. The interviews also helped us to better formalize the proposed method by using the adequate terminology and by providing more detailed guidance. They also confirmed that the selected statistical methods were relevant. We gave the researchers feedback after the interviews and they were satisfied with the final version of the method.

We then built the final method as a method family that consists in describing ‘a set of several organized method components for a specific domain’ (Kornyshova, Deneckère & Rolland, 2011).

3. State of the art

3.1. Map: Intentional process modelling language

The modeling language we used to formalize the method is called Map (Rolland, Prakash &Benjamen 1999). This language allows representing processes focusing on the intentions to achieve and on the strategies to adopt to reach these intentions. Figure 1 presents a simple example of a map process model where one intention is defined as a node Visit Paris. A map process model always begins with a start intention and ends with a stop intention. The strategies are defined as edges between nodes. In this example there are 3 strategies to achieve the intention Visit Paris: By visiting websites, By traveling for leisure or By attending CAA 2014 (which took place in Paris). When the intention Visit Paris is achieved, the process can stop By satisfaction of the intention if the visit was good, or by discontent if the person was not satisfied of the visit.

Map modeling language allows representing flexible processes as several strategies can be applied to achieve a given intention. For instance, if the strategies By visiting websites is not sufficient to achieve the intention Visit Paris, it is possible to apply other strategies, as long as necessary. On the contrary to activity oriented process models, Map models do not have to be enacted sequentially: one can enact the process as long as the intentions are not achieved without following a particular order.

Map modelling language has proved to be efficient and useful in a lot of different domains: to study strategic alignment (Thévenet, 2007; Rolland, 2004), to specify the outcome of business process models (Salinesi, 2003), to support guidance (Rolland, 1993; Deneckère, 2010), to describe intentional services (Rolland 2010),to express pervasive information systems (Najar, 2011), to define systems requirements (Ralyté ,1999), to study users' behavior to identify and name use cases, to tailor methods (Ralyté, 2003), and also in Humanities to model scientific processes (Hug, Salinesi, Deneckère & Lamassé 2012).

We believe the Map process modelling language is adapted to represent the process of using statistical analysis methods as it is flexible and close to the human ways of thinking and working.

3.2. Statistical analysis methods

In this section we will briefly describe the selected statistical analysis methods and their corresponding process model represented with Map. Figure 2 presents the five different methods with their specific intentions and strategies.

The Principal Component Analysis (PCA) (Pearson 1901) (Hotelling 1933) is a technique of factor analysis. The analysis takes as input a table with n rows and p columns and aims to reduce the size of the table by determining new reduced variables containing more information. This method provides a scatter plot (with adjustment of the cloud computing individuals and inertia calculation to measure the dispersion of the cloud using a correlation matrix).

The Correspondence Analysis (Benzecri, 1982) is similar to PCA but while the PCA applies to continuous numeric variables, the CA applies to two categorical variables. The method is used to study the link between two qualitative variables. The data normalization is performed using a Chi-square test (particularly by correlation analysis). A contingency table is established as a homogeneous comprehensive table. The average profiles (centroids)

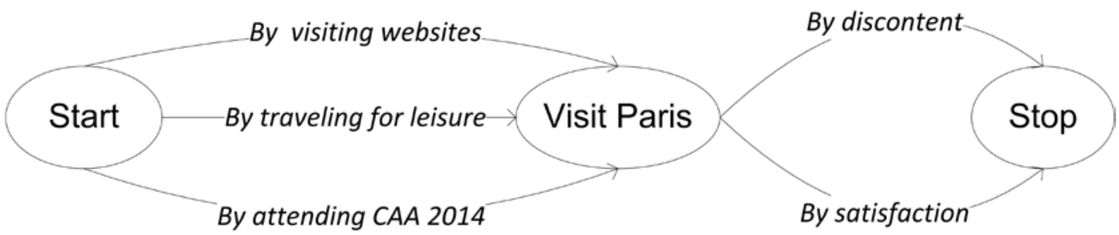


FIGURE 1: EXAMPLE OF A MAP PROCESS MODEL.

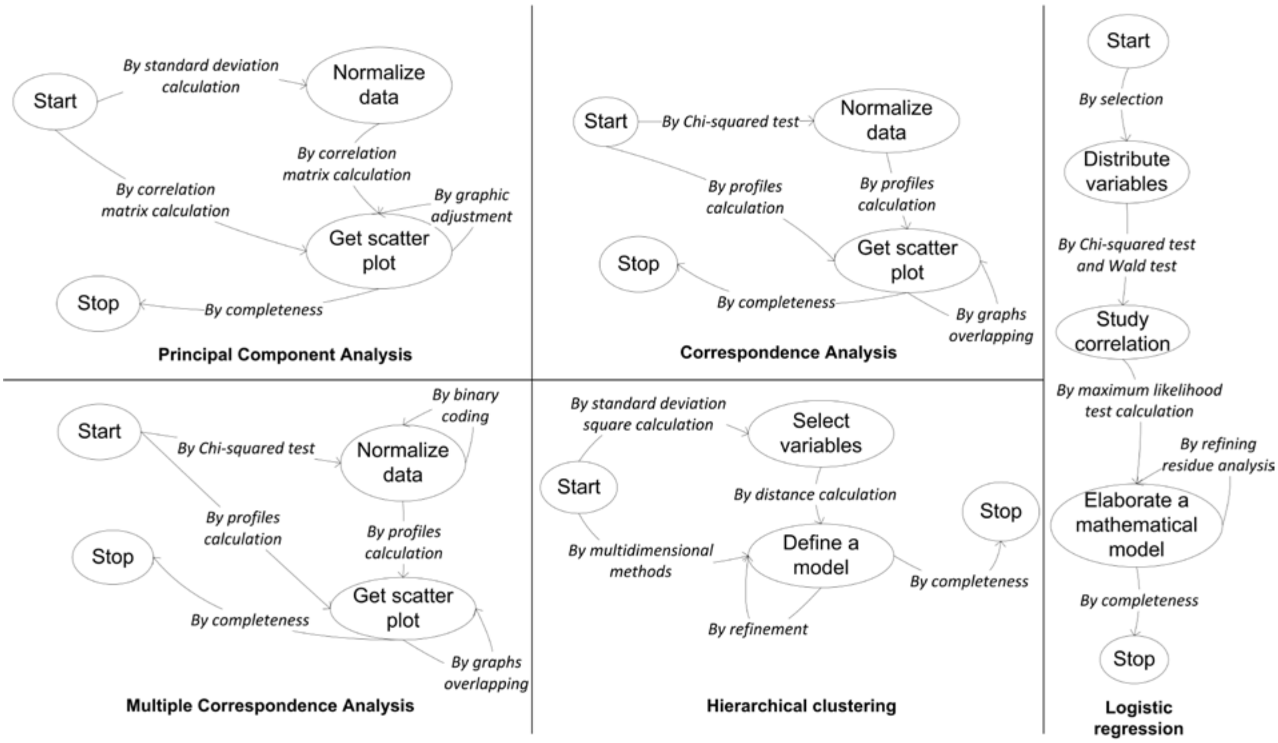


FIGURE 2: THE MAPS OF THE DIFFERENT STATISTICAL ANALYSIS METHODS.

allow determining whether and how a class of individuals differs from the general population.

The Multiple Correspondence Analysis (MCA) (Benzécri, 1973) is similar to a CA applied to more than two variables. The data normalization table is performed so as to obtain binary numbers. The rows represent the individuals, whereas the columns represent the variables. This table is then transformed into a complete disjunctive table. The rows still represent individuals, but the columns represent the terms (each term is connected to a variable). Once the number of variables increases, the data is represented as a hyper-contingency table (called Burt table). The similarity between individuals is determined by the number of terms in common. Thus, two terms are similar if they are present or absent in many individuals. The profile calculation is also used to get the total of the rows and columns, as in CA. Each profile forming a cloud of points is then projected into a different space. The projection onto a single plane allows highlighting a series of orthogonal directions, to study the projections of two clouds, which allows to choose the number of projections axes and to study the values representing the inertia of each axis.

The objective of the Hierarchical Clustering (Sokal & Sneath, 1963) is to classify individuals sharing similarities from a set of variables. The selection of the variables representing the individuals is done using the standard deviation. The distance between the individuals is then calculated (e.g. with the calculation of the Euclidean distance or the Chebyshev distance), then the distance between groups (one can calculate the minimum distance, maximum distance or the distance defined by the method of Ward (1963)) to determine a dissimilarity and aggregation index. The result of a HC is a hierarchy of classes such that any classis not empty, every individual belongs to a class an deach classis the union of the classes that are included. To interpret the dendrogram, a partition is selected from the class hierarchy. The cleavage is carried out at a level where the inertia between classes increases suddenly and significantly. The data of the remaining subsets are considered relevant.

The Logistic Regression (Berkson, 1944) first consists in studying the distribution of different variables by calculating the conditional probabilities (Rakotomalala 2011). Two approaches areproposed to estimate this probability: based on frequencies or based on the calculation of the likelihood ratio. It is alsopossible to use models such as logit (Berkson, 1944) or probit (Bliss, 1934) to 'score' each individual. Then the correlations between the independent variables and the relationship between each explanatory variable and the dependent variable (Wald test) are calculated. The logistic regression model is finally built by applying the Chi-square test of the model (the variables are they relatedto the dependent variable?) and the Chi-square maximum likelihood (Hosmer & Lemeshow, 2000) to determine if the data contradict the established relationships. The adjustment of the model is necessary as it allows to judge the qualityof the final model according to the data.This adjustment can be made with the tests of residual analysis as residual deviance and Pearson. The presentation and interpretation of the models can be made by Odds ratio test and residue analysis to check the quality of the regression.

4. The Method

4.1 The global map

We then used Map to represent the method to guide researchers to use statistical analysis methods. We first defined a global map (see Figure 3), presenting the whole analysis process, from the collect of the data to the interpretation of the results. This map was built based on the literature review and on the interviews.

We first defined four different strategies to collect data: by interviews, by documentary research, by field research (in the case of archaeologists for example) or by interpreting existing works. The researcher can enact these strategies as long as the intention collect data is not achieved. The researcher can also collect data by critical analysis of his own collect or by transcription when the data is entered into a computer file as a spreadsheet, a data base, an xml file, etc. If the researcher is experienced, he/she can directly analyze data.

During and after the process of the data collection it is important to ensure that the sources are complete and valid. We then defined the intention Ensure sources completeness and validity and the strategies by source checking and by expert validation. An expert of the domain will be able to detect problems (corrupt or incomplete sources), and to advise the researcher to study new potential sources.

The researcher now has the data checked. He/she can then establish a hypothesisthat is the question he/she wants to explore according to the collected data. This can be done using 3 different strategies: by continuity of previous works, by counter-hypothesis (when a hypothesis has already been established) or by intuition.

Before starting the statistical analysis of the data, the researcher should verify if the hypothesis he/she wants to establish and study is original, that no researcher has already answered it in a similar way (Guarantee hypothesis originality intention). We then propose two strategies: by systematic review and by expert validation. If the hypothesis is weak, not original, not specific enough or on the contrary too specific, it can be changed by applying the different strategies: by modification, by specialization or by generalization to achieve the intention Establish hypothesis.

The next step is then to analyze the data by statistical analysis. This particular section will be detailed in section 4.2 as it is complex and requires more guidance.When the data is analyzed, the researcher can decide to change the hypothesis if the results are not satisfying: by intuition if he/she feels the hypothesis has to be changed, by specialization if it is too general, or by generalization if it is too specific.

When analyzing the data, the researcher can also realize the sources are not complete or should be modified; the strategy by return to sourcesis then followed. The results of the analysis can be refined to improve the visualization of the results by refinement.

Finally, the analyzed data can be interpreted. This is the last step of the method (by interpretation).

4.2 Analyze data by statistical analysis

Analyzing data by statistical analysis can be complex as there are many statistical methods and researchers need guidance while executing them.We then refine the section <Establish hypothesis, Analyze data, by statistical

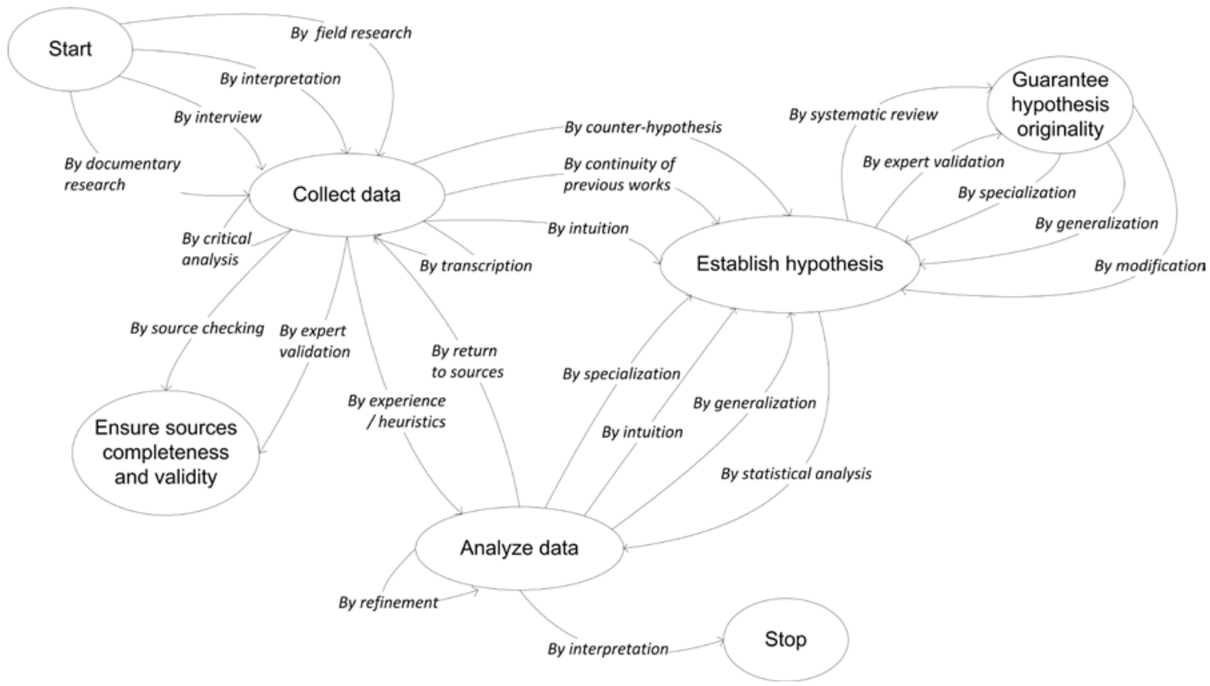


Figure 3: Global view of the proposed method.

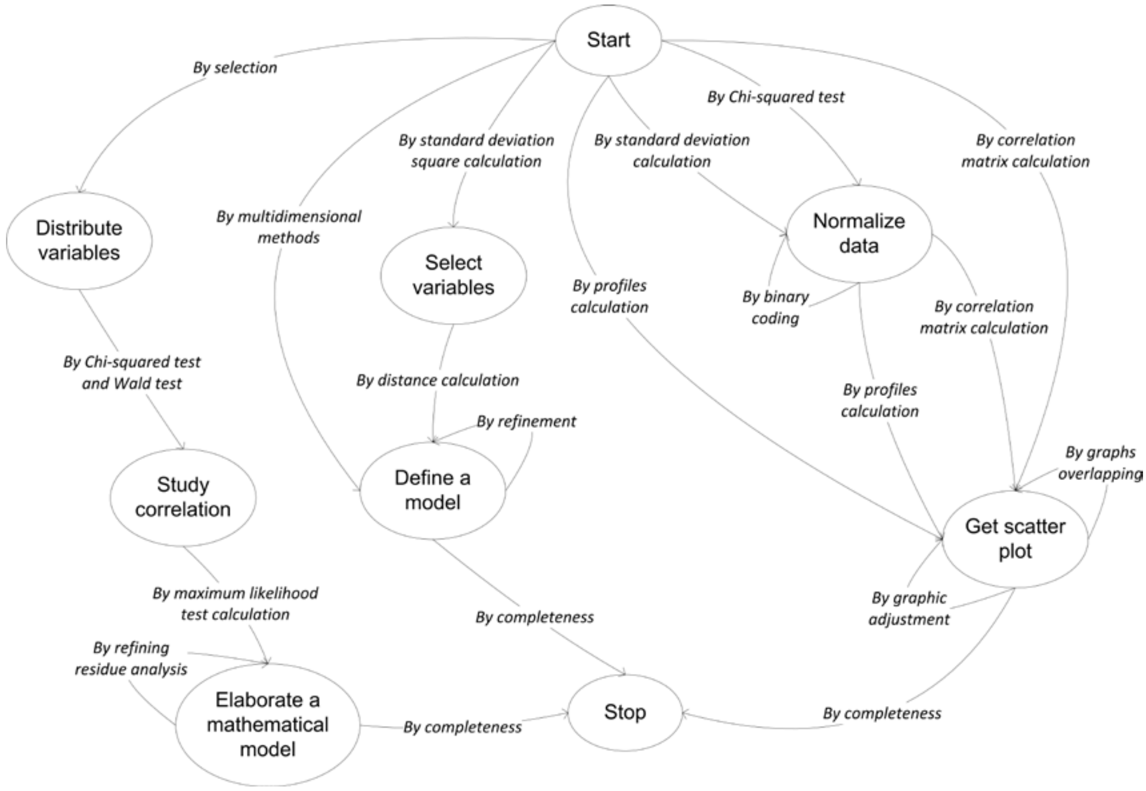


FIGURE 4: REFINEMENT OF THE SECTION <ESTABLISH HYPOTHESIS, ANALYZE DATA, BY STATISTICAL ANALYSIS>.

analysis> into another map process model. This map also comprises Start and Stop intentions. It was built from the literature review and the interviews conducted with the Humanities researchers. Figure 4 presents the map process model refining the section <Establish hypothesis, Analyze data, By statistical analysis>.

The statistical methods used to construct this process model are the Principal Component Analysis, the Correspondence Analysis, the Multiple Correspondence Analysis, the Hierarchical Clustering and Logistic Regression methods. We built this map according to the models we presented in section 3.2.

Each path of this model corresponds to a particular method. For instance, the path comprising the intentions Distribute variables, study correlation, elaborate mathematical model were extracted from the Logistic regression method. The path comprising Select variables and Define a model were extracted from the Hierarchical clustering method.Some paths are common to different statistical analysis method as the intentions Normalize data and Get scatter plot that are defined both in Principal Component Analysis, Correspondence Analysis and Multiple Correspondence Analysis methods. However the strategies allow us to differentiate the methods.

5. Conclusion and future work

This work is a first step towards a tool to guide humanities researchers in using statistical analysis method. We provide a set of process models that describe the collect of data, the establishment of hypothesis and the analysis of he data. We conducted a first evaluation of the proposed models with four researchers.

We plan to validate the proposed method with master and PhD students from university Paris 1 Panthéon-Sorbonne to measure its understanding and ease of use. The experiment platform is ready and the evaluations will start in October 2014 using Google Forms and R.

The next step of this research is then to implement the method into the Online Method Engine (Vlaanderen, Spruit, Dalpiaz & Brinkkemper 2014) to provide an online tool to Humanities researchers. This tool will allow gatheringall the process models for statistical analysis methods to make them available to the community, including new ones added by researchers themselves.

We also need to define the context of the data analysis projects (as the volume or type of data) to better guide the use of one statistical method or another, to help the researchers in selecting the best fitted method.

Currently, the proposed models are suitable for multi-dimensional methods, but we also plan to introduce other methods to analyze other type of data such as natural language or image. This will also be done through the Online Method Engine.

References

BENZÉCRI, J.-P. (1973). *L'analyse des données: L'analyse des correspondances*, Dunod, Paris.

BENZÉCRI, J.-P. (1982). *Histoire et préhistoire de l'analyse des données*, Dunod, Paris.

BERKSON, J. (1944). Application of the Logistic Function to Bio-Assay, *Journal of the American Statistical Association* 39, pp. 357-365.

BLISS, C. I. (1934). The method of probits, *Science*, vol. 79, no. 2037, pp. 38–39.

CANNING, J. (2014). Statistics for Humanities, Available from http://statisticsforhumanities.net [July 2014]

DENECKÈRE, R., KORNYSHOVA, E. (2010) Process line configuration: An indicator-based guidance of the intentional model MAP, *Enterprise, Business-Process and Information Systems Modeling* 50, pp. 327-339, Springer, Berlin Heidelberg.

HOSMER, DW., LEMESHOW, S. (2000). *Applied Logistic Regression*, Wiley, New-York.

HOTELLING, H. (1933). Analysis of a complex of statistical variables into principal components, *Journal of Educational Psychology* 24, pp. 417-441 & 498-520.

HUG, C., SALINESI, C., DENECKÈRE, R., LAMASSÉ, S. (2012). Process modeling for Humanities: tracing and analyzing scientific processes, *Proceedings of Annual Conference of Computer Applications and Quantitative Methods in Archaeology*, Available from: Amsterdam University Press, pp. 245-255.

KORNYSHOVA, E., DENECKÈRE, R., ROLLAND, C. (2001). Method Families Concept: Application to Decision-Making Methods, Enterprise, Business-Process and Information Systems Modeling, *Lecture Notes in Business Information Processing*, vol. 81, pp 413-427.

MIRBEL, I., RALYTÉ, J. (2006). Situational method engineering: combining assembly-based and road map driven approaches, R*equirements Engineering* 11(1), pp. 58-78.

NAJAR, S., KIRSCH-PINHEIRO, M., SOUVEYET, C. (2011). Towards semantic modeling of intentional pervasive information systems, In *Proceedings of the 6th International Workshop on Enhanced Web Service Technologies*, ACM, NY, pp. 30-34.

PEARSON, K. (1901), On Lines and Planes of Closest Fit to Systems of Points is Space, *Philosophical Magazine Series* 2(6), pp. 559-572.

R, computer software 2014.Available from: http://www.r-project.org/

RAKOTOMALALA, R. (2011). *Pratique de la Régression Logistique*, Université Lyon 2.

RALYTÉ, J. (1999). Reusing scenario based approaches in requirement engineering methods: CREWS method base, In *Proceedings of Tenth International Workshop on Database and Expert Systems Applications*, IEEE , pp. 305-309,.

RALYTÉ, J., DENECKÈRE, R., ROLLAND, C. (2003), Towards a generic model for situational method engineering, *Lecture Notes in Computer Science* 2681, pp. 95-110.

ROLLAND, C. (1993). Modeling the requirements engineering process, in: H. Kangassalo, H. Jaakkola (Eds.) *Proceedings of the Third European-Japanese Seminar*, Budapest, Hungary, May, pp. 85-96.

ROLLAND, C., KIRSCH-PINHEIRO, M., SOUVEYET, C (2010). An intentional approach to service engineering, *IEEE Transactions on Services Computing* 3(4), pp. 292-305.

ROLLAND, C., PRAKASH, N., BENJAMEN, B. (1999). A Multi-Model View of Process Modelling. *Requirements Engineering* 4(4), pp. 169-187.

ROLLAND, C., SALINESI, C., ETIEN, A. (2004) Eliciting gaps in requirements change, *Requirements Engineering* 2(1), pp. 1-15.

SALINESI, C., ROLLAND, C. (2003). Fitting Business Models to System Functionality Exploring the Fitness Relationship, *Lecture Notes in Computer Science* 2681, pp. 647-664.

SAS, computer software (2014). Available from: http://www.sas.com/en_us/software/analytics.html

SOKAL, R.R., SNEATH, P.H.A. (1993). *Principles of Numerical Taxonomy*, San Francisco: W. H. Freeman and Company, .

THÉVENET, L.H., SALINESI, C. (2007). Aligning IS to organization's strategy: the INSTAL method, *Lecture Notes in Computer Science* 4495, pp. 203-217.

VLAANDEREN, K., SPRUIT, S., DALPIAZ, F., BRINKKEMPER, S. (2014). Demonstration of the Online Method Engine, In joint Proceedings of the CAiSE 2014 Forum and CAiSE 2014, Doctoral Consortium colocated with the 26th International Conference on Advanced Information Systems Engineering, *CEUR workshops* 1164, pp. 169-176.

WARD, J.H. (1963). Hierarchical grouping to optimize an objective function, *Journal of the American Statistical Association* 58, pp. 236-244.

Ancient Mesopotamian Glyptic Products, Statistics and Data Mining: A Research Proposal

Alessandro DI LUDOVICO,[1] Sergio CAMIZ[2]

[1] Dipartimento di Scienze dell'Antichità, Sapienza Università di Roma, Italy
[2] Dipartimento di Matematica Guido Castelnuovo, Sapienza Università di Roma, Italy
alediludo@gmail.com, sergio.camiz@uniroma1.it

Abstract
A stratified and complex investigation of the figurative language of a corpus of Mesopotamian glyptic artefacts will be described here. The methodologies adopted and the formal description of the products under investigation are the result of a series of experimental works carried out in the past. Two main problems have been faced in the former experiences: the proper translation into an adequately coded form of what could be observed on the artefacts and the choice and adapting of a specific investigation methodology. The development of these researches has lead to the use of different models, and also to their parallel adoption, and to the comparison of the relevant results and logistic or logical weaknesses and advantages. Here a further research course will be pursued, as a continuation of the former ones. The corpus of Ur III presentation scenes will be explored through an integration of models based on different logics.

Keywords: Mesopotamian Glyptic, Correspondence Analysis, Ur III figurative Languages, presentation scenes.

1. Introduction

Glyptic iconography has historically been the preferred subject for the use of quantitative methods in the art history of Western Asia.[1] The compositions of motifs carved on the surfaces of cylinder seals are thus very tempting for quantitative analyses, perhaps also because of the long and widespread tradition in the use of this class of artefacts in the ancient Near East. Nevertheless, their study through quantitative methods can become a difficult challenge, more or less complex and stimulating – depending on the chronological and geographic references of the chosen corpus, its inner largeness and the ultimate aims of the investigation. A research project dealing with the presentation scenes of Mesopotamian glyptic of the second half of the third millennium began over ten years ago with the development of a coding method based on linguistic principles (Di Ludovico, 2005).

That coding was used for some initial investigations in which mathematical algorithms permitted the outlining of some diachronic comparisons and classifications (Di Ludovico and Ramazzotti 2008). Later on the data set changed, in order to perform more specific investigations: thus its reference period was narrowed and a number of new records were introduced (Di Ludovico, 2011). The chronological reference period was reduced to Ur III Dynasty one, an era in which the theme of the presentation can be considered central, besides becoming relatively more homogeneously structured.[2] The new data set was thus more homogeneous, but was also made up of new codings, prepared in the view of processing based on statistical algorithms (Di Ludovico, Camiz and Pieri, 2013; Di Ludovico and Camiz, in press).

2. The Data Set for the Textual Correspondence Analysis

The new codings of the scenes produced a data set that collects 354 records, 85 of which are known from ancient impressions and only 269 from cylinders. A place of origin or find-spot is documented or can be inferred for 212 scenes, while for the dating we mostly ascribed the relevant seals to three main phases (early, core, and late Ur III period), based on stylistic and iconographic criteria, since useful elements giving more precise dates are only available for a minority. Toponyms and chronological phases have been here understood and used as the external features. The presentation scenes were coded as formalised and adapted verbal descriptions of their iconographic content and inner compositional relations.

The codings were then processed through the Textual Correspondence Analysis (TCA) algorithm, which is included in the SPAD software pack:[3] after a preliminary exploration of the corpus, a succession of different applications of TCA led to outline iconographic classes of the scenes. Then the relevant provenance sites and relative chronological subdivisions were consequently characterised (Di Ludovico and Camiz, in press). The algorithm allowed the relative proximity of forms used

[1] See the other contribution on the use of quantitative methods in the art history of the pre-Classical Near East by the same authors in this publication, pp. 29-34.
[2] See, e.g., Collon 1982: 129. Under the rule of Ur III Dynasty, the concrete expressions of this theme progressively become almost exclusive in glyptic, besides showing an increasing homogeneity in the compositional logics.
[3] Lebart and Salem 1994. Past applications of these algorithms in glyptic studies: Camiz and Rova 2001; 2003; Di Ludovico, Camiz and Pieri 2013; Di Ludovico and Camiz in press

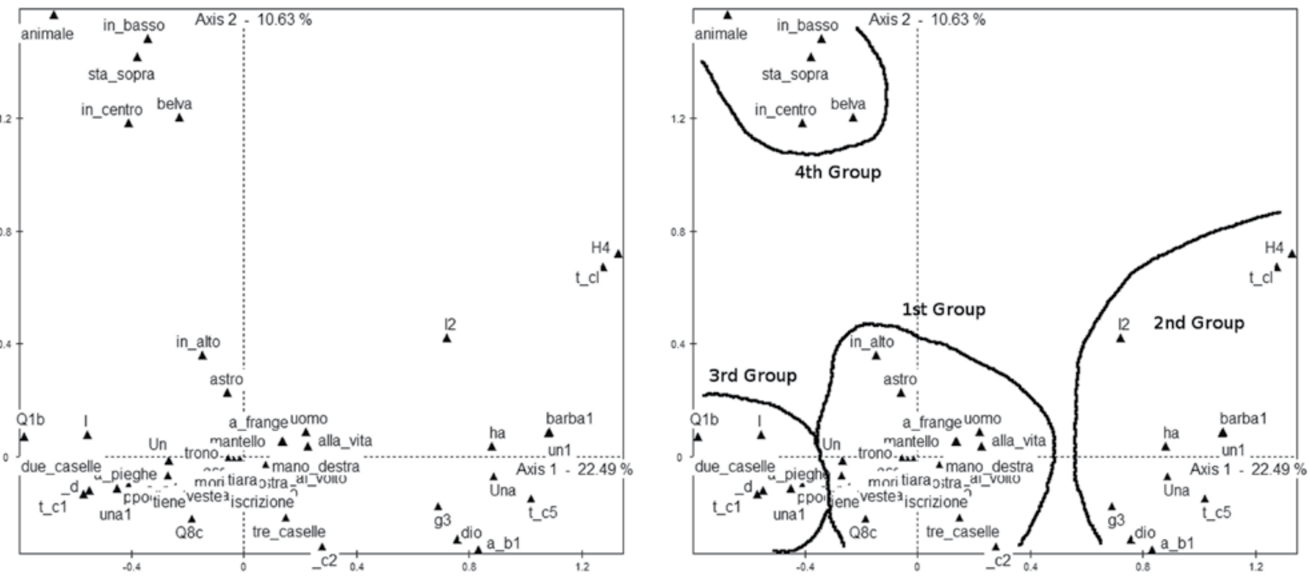


FIGURE 1: RELATIONS DETECTED THROUGH THE FORMS IN THE FIRST STAGE OF THE INVESTIGATION (FORMS CROSSED WITH EXTERNAL FEATURES).

to describe the scenes to be measured on the basis of the similarity among the contexts in which the forms appeared. The classes of scenes were outlined through a Hierarchical Ascendant Classification (HAC) based on the Euclidean distance and using the Ward's (1963) aggregation criterion.

3. Classification of scenes through forms and investigations through external features

During the first stages of this project, the relations between the forms used in the descriptions of the scenes and the external features were investigated, and the scenes were classified using these forms. From this, an initial frame in which some toponyms and sub-periods were opposed to each other along a number of major axes was obtained.

In the outcomes of the first analyses (Fig. 1) the forms distributed along the axes according to logics which seem to distinguish those more connected to the basic concept of Ur III presentation (axis 1, negative side) from those that are symptoms of more specific features, like the ones related to "royal" presentations (axis 1, positive side), or re-working activities (axis 2, positive side).

A similar general picture of the forms emerged in the analysis in which they were crossed with the external

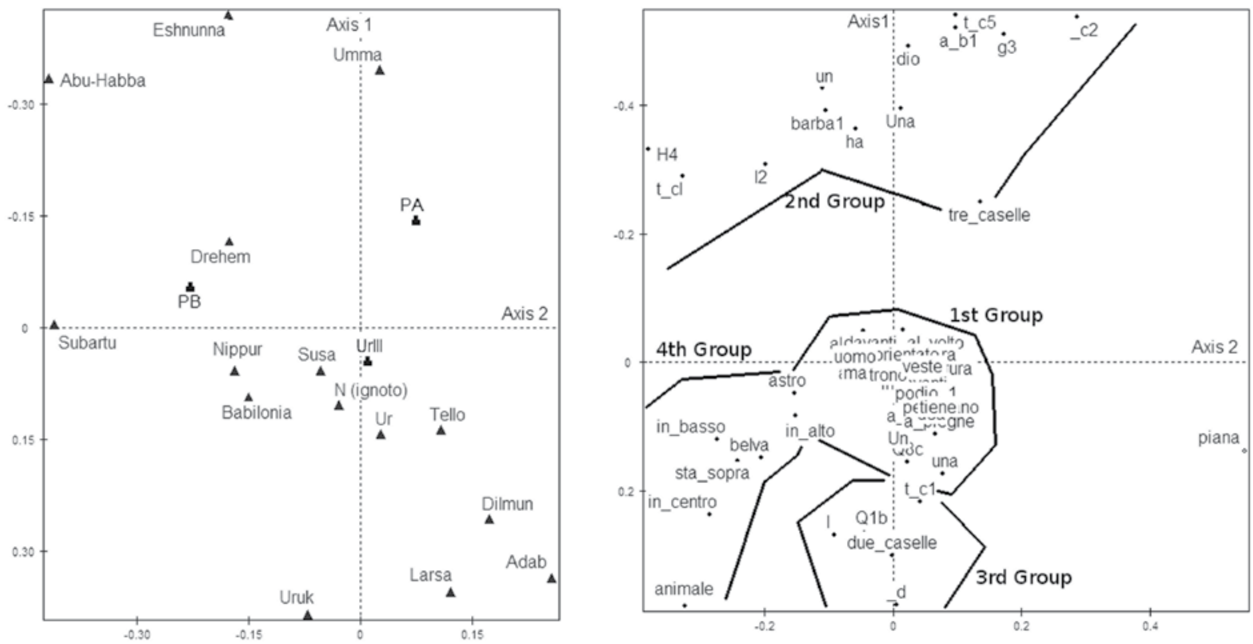


FIGURE 2: RELATIONS DETECTED THROUGH THE EXTERNAL FEATURES IN THE FIRST STAGE OF THE INVESTIGATION (LEFT), AND RELATION THROUGH THE FORMS CROSSED WITH EXTERNAL FEATURES (RIGHT).

features (Fig. 2). The latter were arranged in the graph so that they gave a quite clear regional distinction, but parallel to this, a qualitative difference in the materials under analysis also emerged, especially in the case of the toponym of Umma, to which a very large number (almost all) of original impressions included in the data set are related. On the other hand, axis 2 distinguishes the regional characters of the most southern regions (Dilmun, Larsa, Adab...) from those of the northern ones (Subartu, Nippur), with Ur and Susa representing the logical centre and the Diyala Region (here represented by Eshnunna) closer to the northern sites, but in a special relation with specific administrative features (axis 1, negative side).

The classification of the scenes through forms led to the outlining of 9 groups which are encapsulated in the 4 resulting from a different cut of the relevant tree (Fig. 3). The classes show an opposition which corresponds quite clearly to that observed in the analysis of the forms. On the one hand, the scenes which represent what we interpreted as the basic idea of presentation of this period (class I), with the outlining of some earlier specimens (class 2), formed a quite large and compact group, with a relatively stable structure.

On the other hand, a smaller group (class II) collecting later or reused scenes (class 3), or heavily reworked ones (class 4), and others with a frequent presence of a male god receiver (class 5) was much more heterogeneous and dispersed through the graph. A clear opposition to the latter showed, along both axes, a further group which contained a large number of original impressions (class 6/III), thus representative of the scenes which correspond more than others to the idea of compositions associated with administrative activities.

The last major class was also made up of fairly scattered specimens, with presentations before a male figure and with few integrating motifs (class 7) and two types of royal presentations differing from one another in the compositional pattern (class 8; class 9).

4. New investigations on segments

Single processes of binary correspondence analysis were then carried out in order to compare the behaviour within the data set of couples of different types of features pertaining to the scenes or the cylinders on which they are depicted. The first two couples of features that have been compared in this way are, on the one hand, integrating motifs (that is all objects, animals, symbols, and similar figures placed in the field without playing an evident active role in the scene) and legends, and, on the other hand, integrating motifs and external features.

4.1. Cross-comparison of integrating motifs and legends

The distribution obtained in the comparison of legends and integrating motifs (Fig. 4) shows that the main contribution to the positive part of the first axis is the one provided by quite short legends mentioning name, profession and patronymic of the seal owner (according to the formula: «PN, professional qualification, DUMU [son of] PN2»). Less meaningful, but still quite important, is the contribution of larger legends with the royal dedication, mentioning the royal titles and ending in the formula «his servant» (ÌR.ZU). Here, the contribution of integrating motifs comes mainly from the standard topped by a bull, that topped by an element resembling a prong, the lion, the small figure of the hero with a sprinkling vase, the table with offerings, and the pot for libations. Less meaningful is the contribution given by the anzu-like eagle and the lizard (similarly weak is the contribution of the sun disc with the moon sickle).

The negative part of the same axis develops through the main contribution of legends with the formulae: name,

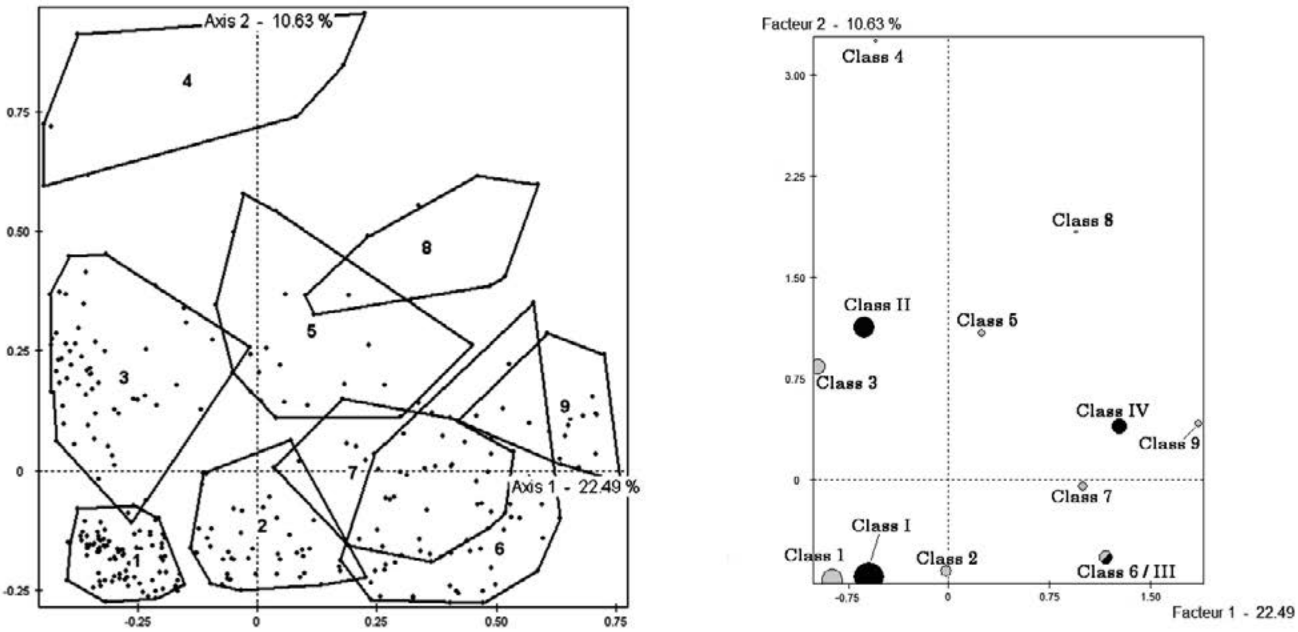


Figure 3: Results of the classification of the scenes through the forms.

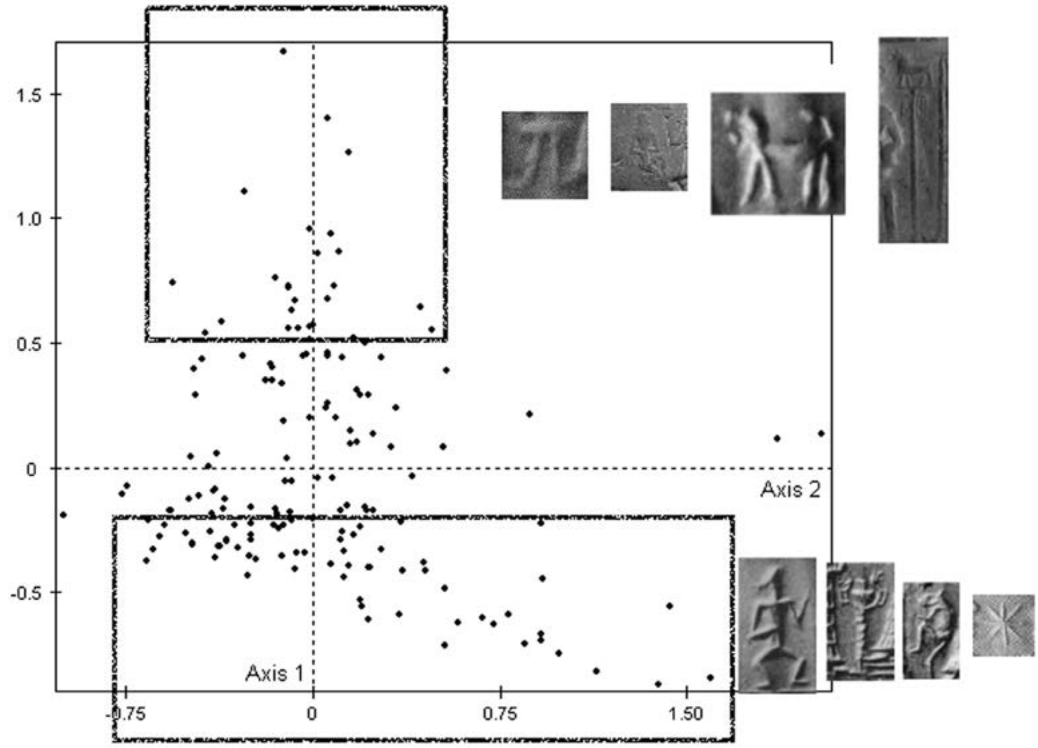


Figure 4: Results of the second stage of the investigation (legends through integrating motifs).

patronymic and profession; divine name; not readable; name, patronymic, and the «servant of [divine name]» formula; simple name. They are usually in a narrow arrangement, that is one to three lines.

Contributions from integrating motifs here come mainly from the figure of the dwarf, the standard with lion heads, the monkey, the star, the ball-and-staff, the mace, the scorpion, the standard with crescent, the goose, the long snake, and a few other types of wild animal figures.

The basic opposition features along the first axis are thus the largeness of the seal legend and the relevant content, but also, and especially, the order of formulae expressing name, patronymic, and profession. This is probably connected to diachronic changes, with an increasing official standardisation of the seal legends that lead to the arrangement prevailing in the positive part of axis 1. However, it is also likely that this diachronic explanation is valid only for the mentioned set of formulae, while the position of the content of other short legends is rather to be interpreted in the light of its relation with the official administrative destination of the seal. Such an interpretation seems to be confirmed by the behaviour of the integrating motifs, those along the positive part of the axis are less various than the opposite ones, and are also closer to the elements ordinarily included in scenes known from impressions on administrative documents. Furthermore, along the negative side of the axis, one also finds integrating motifs typical of later (sub-)periods, like the ball-and-staff, or in typical arrangements of recut seals, as expressed by the forms which refer to a position of the element in the upper or central part of the field.

On the positive side of axis 2 (Fig. 5) the main contribution is provided by the royal gift legends, those bearing just a name, and the legends with name, patronymic, and profession or with name, patronymic, and «servant of [divine name]» formula. They are usually short legends, but important contributions also come here from very large inscriptions, like those known as royal and royal gift legends.

The integrating motifs which give a positive contribution are the dwarf, the standard with the prong-like element, the monkey, the vase for libations, the asterisk, the ball-and-staff, and others, mainly wild animals, the weight of which is lighter for the formation of the axis.

The negative contribution is also primarily provided by infrequent features, like the legends with name and the formula «wife of PN», two divine names, and the dedication for someone's life, besides those mentioning a name and a profession. Furthermore, a relatively less significant contribution comes here from legends that mention in various order name, patronymic, and profession. The dimensions of the legend can vary a great deal along this side of the second axis. Integrating motifs that contribute most here are the table with offerings, the griffins holding maces, the bull, the altar with the date palm, the anzu-like eagle, and various others.

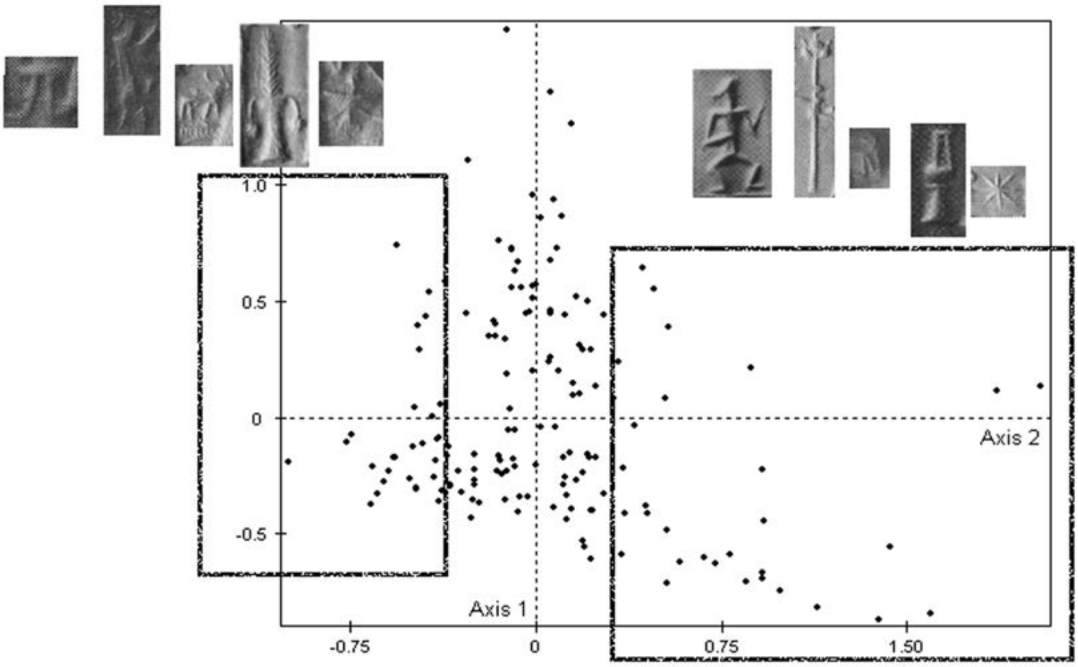


Figure 5: Results of the second stage of the investigation (legends through integrating motifs).

The hidden logics of the oppositions along the second axis are not as clear as for axis 1, but a number of clues suggest that here distinctions of a geographic nature prevail. Along each side of the axis the elements are actually of various kinds, and can be related to different historical periods. The location of sites and periods on the map provide a certain confirmation of this, since southern sites are mainly related to the negative side of the axis, or very close to its zero point (Fig. 6).

4.2. Cross-comparison of integrating motifs and external features

In the analysis of external features and integrating motifs, axis 1 receives a meaningful positive contribution especially from Umma, a little less from north-eastern and southern sites, as for the places of origin, and from Early UrIII as for the (sub-)periods (Fig. 6). On the negative side, the contributions are not very significant, and are those of Tello and Unknown place of origin and Late Ur III sub-period.

The integrating motifs which give the most important contribution to the positive side of the first axis are almost all placed in the lower part of the field, and are few: the standard with bull, the lion, the lizard, the hero with sprinkling vase, the table with offerings, the altar with date palm, the libation vase.

A negative contribution comes mainly from the dwarf, the standard with griffin heads, the asterisk, the monkey, the

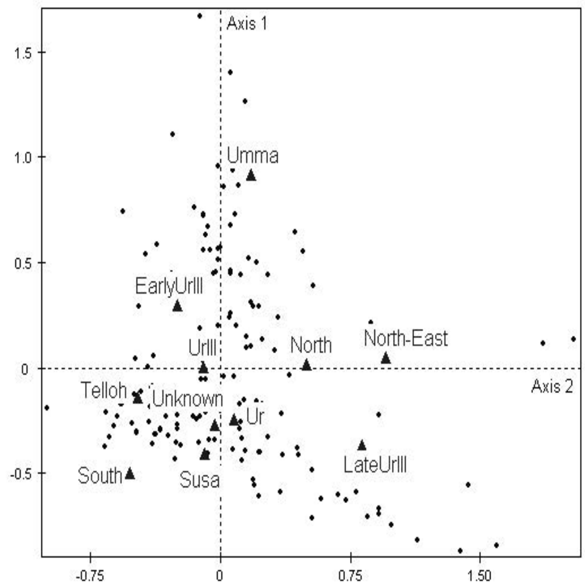


Figure 6: Relations among the external features from the second stage of the investigation.

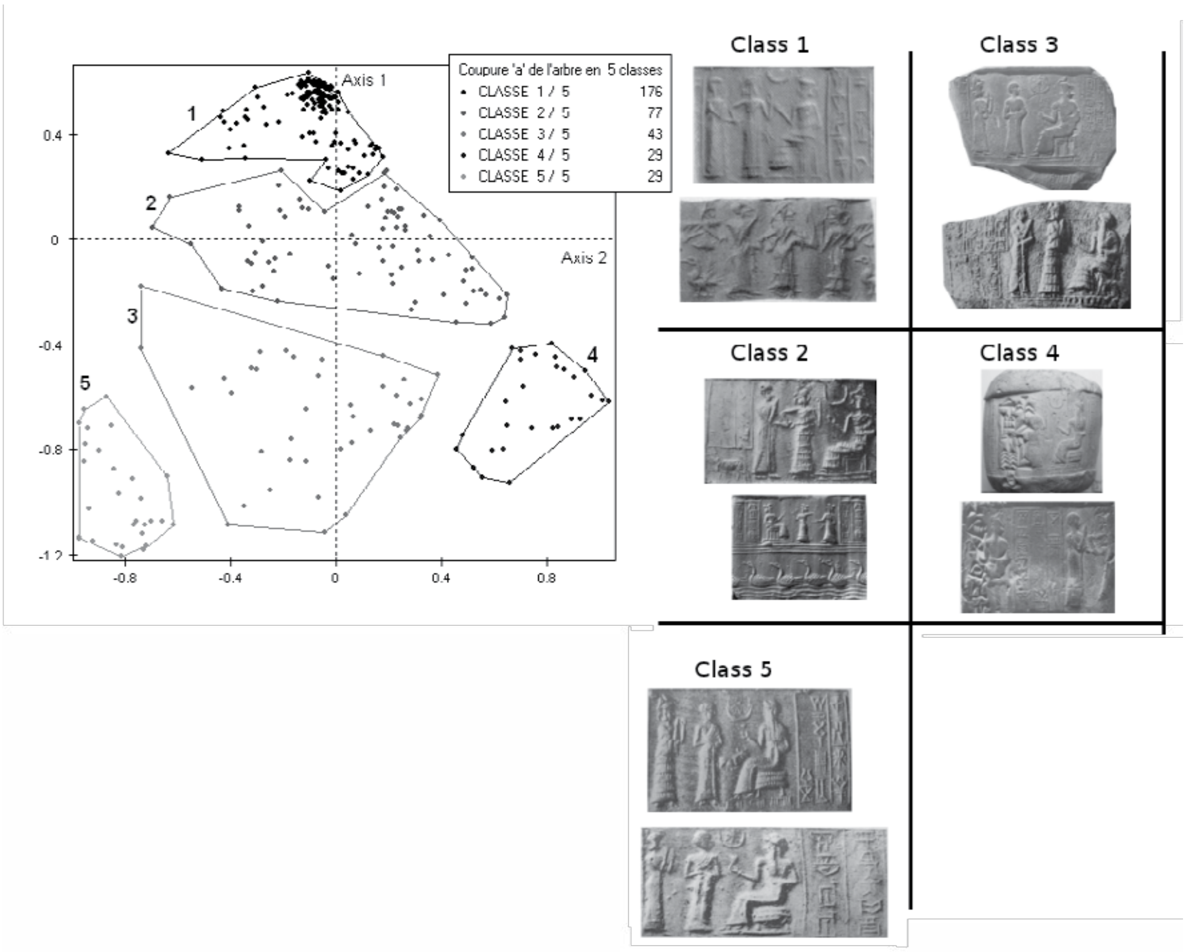


Figure 7: Classes of scenes obtained in the second stage of the investigation (partition in 5 classes).

mace, the ball-and-staff, and much less from a number of other animal figures. Generally speaking, the opposition also here seems mainly based on logics related to the needs of bureaucracy, that is to a nature of the seal which is more or less bound to its official administrative use. Scenes from Umma are in fact recorded on original impressions made on administrative documents, while scenes belonging to the Late Ur III sub-period can often show signs of recutting. On the other hand, many of the integrating motifs located along the negative side also often recur in reworked scenes (that is scenes which were altered through secondary carving, an operation that in most cases probably changed the official meaning and role of the relevant seal), while those on the positive one are not infrequently on original impressions.

The situation recorded along the second axis seems not to be very clear, partly because of what has already been observed by the analysis on legends and integrating motifs. Positive contributions to it come mainly from Unknown sites of origin and peripheral sub-periods (Early and Late Ur III), while various geographical origins (southern sites, northern ones, Susa, Ur, etc.) and the core sub-period find their place on the negative side. Integrating motifs distributed along the positive side of this axis are of various kinds, and are in part those that are often used in re-cut workings of the seals, or in relatively late productions, like lion-griffins or the dwarf. Many elements that are more typical of the core of the Ur III age and of its administrative iconography find their place on the opposite side. Examples of these are the anzu-like eagle or the hero with a sprinkling vase, but also some other, more generally used motifs, like the scorpion or some types of birds. Especially the position in the top part of the field is quite unusual for scenes that are typical of official administrative seals. It is possible that the seemingly contradictory nature of this axis is due to the badly defined and quite heterogeneous group of scenes with unknown origin, which acquired too much weight in positively defining it.

4.3. Classification of scenes

The classification of the scenes through the segments gives a general picture that looks very similar to that obtained from the analysis through the forms, but suggests that some compositional arrangements which could seem not to be very evident at first sight bring a remarkable contribution to the semantic of the scene.

Five classes are obtained here (Fig. 7).

In the first class, which is also the largest, one finds scenes of presentation before a goddess with three figures, but many of them bear the signs of re-workings, have a relatively large number of integrating motifs or show traits that can be related either to a late or to a very early date.

The second class is much less compact than class one, and is split into two minor parts along axis 2; on its positive side, scenes are of the kind that show three figures with the female deity as a receiver and with few and consistently used integrating motifs. The legends are here never very short and there are few signs of re-workings; few scenes with a male divine receiver are included by the positive side.

The remaining three classes are much less numerous. In the third class one can locate different kinds of royal presentations, mostly with figures which are not physically connected (that is, without pairs of figures depicted hand in hand); in the fourth group scenes with the male deity in the role of the receiver and the hand in hand couple do appear, and in the fifth one there are various kinds of royal presentations, but all with very few or no integrating motifs.

This general picture is a stimulus for a deeper and systematic investigation on the smaller iconographic features and their arrangements.

5. Future Proposals

The results obtained through the use of TCA until now provide encouragement for the continuation of this research project with further investigations of the same kind. The next stages of the project will deal with details and specific features of the iconography of the scenes, to provide us with clues about the logics of their use. Over a longer time-span, these same investigative procedures will be used on a larger data set, in which other types of contemporary scenes would also be included: basically the contest scenes. Furthermore, the relationships between sites of origin and original impressions should be better balanced and the content of legends could be investigated more in depth, introducing the succession of the names of the signs in the coding. Before accomplishing these tasks, however, a number of technical problems need to be solved. Firstly, the compositional structure of contests is very different from that of presentations, which would enhance the complexity of the data set and thus require either a more complex coding system or a strongly simplified version of it. Moreover, original impressions are only easily accessible in few cases and are almost always published in a form that is not sufficient for iconographic analyses or descriptions. As for the content of legends, the nature of the Sumerian language and the state of research into it, is such that developing a proper coding system for them requires a relatively long period of time and some specific tests, perhaps in cooperation with philologists.

References

CAMIZ, S. and ROVA, E. (2001). Exploratory Analyses of Structured Images: A Test on Different Coding Procedures and Analysis Methods. *Archeologia e Calcolatori* 12. pp. 7-45.

CAMIZ, S. and ROVA, E. (2003). Quantitative Study of Images in Archaeology: I. Textual Coding. In: Schader, M., Gaul, W. and Vichi, M. (eds.). *Between Data Science And Applied Data Analysis*. Berlin: Springer. pp. 624-632.

COLLON, D. (1982). *Akkadian – Post Akkadian – Ur III Periods. Catalogue of the Western Asiatic Seals of the British Museum 2.* London: British Museum.

DI LUDOVICO, A. (2005) Scene-in-frammenti: una proposta di analisi delle “scene di presentazione” dei sigilli a cilindro mesopotamici orientata all’elaborazione statistica ed informatica dei dati. In: Di Ludovico, A. and Nadali, D. (eds.). *Studi in onore di Paolo Matthiae presentati in occasione del suo sessantacinquesimo compleanno, Contributi e Materiali di Archeologia Orientale. X – Numero Speciale*. Roma: Sapienza. pp. 57-95.

DI LUDOVICO, A. (2011). Experimental Approaches to Glyptic Art Using Artificial Neural Networks. An Investigation into the Ur III Iconological Context. In: Jerem, E., Redő, F. and Szeverényi, V. (eds.). *On the Road to Reconstructing the Past, Proceedings of the 36th Annual Conference on Computer Applications and Quantitative Methods in Archaeology, Budapest, April 2-6 2008*, Budapest: Archaeolingua. pp. 135-146.

DI LUDOVICO, A. and CAMIZ, S. (in press). A Quantitative Approach to Ur III Mesopotamian Figurative Languages: Reflections, Results, and New Proposals. *Archeologia e Calcolatori* 25.

DI LUDOVICO, A., CAMIZ, S. and PIERI, G. (2013) .Comparative Use of Mathematical Models in an Investigation on Mesopotamian Cylinder Seals. In Contreras, F., Farjas, M. and Melero, F.J. (eds.). *Fusion of Cultures. Proceedings of the 38th Annual Conference on Computer Applications and Quantitative Methods in Archaeology, Granada, April 6-9 2010.* British Archaeological Reports – International Series 2494. Oxford: Archaeopress. p. 495-498.

DI LUDOVICO, A. and RAMAZZOTTI, M. (2008). Reconstructing Lexicography in Glyptic Art: Structural Relations between the Akkadian age and the Ur III Period. In: Biggs, R.D., Myers, J. and Roth, M. (eds.). *Proceedings of the 51st Rencontre Assyriologique Internationale, Held at the Oriental Institute of the University of Chicago, July 18-22, 2005. Studies in Ancient Oriental Civilization* 62. Chicago: Oriental Institute. pp. 263-280.

LEBART, L. and SALEM, A. (1994). *Statistique textuelle.* Paris: Dunod.

WARD, J.H. (1963). Hierarchical Grouping to Optimise an Objective Function. *Journal of the American Statistical Association* 58. pp. 236-244.

Intrasite Spatial Analysis Applied to the Neolithic Sites of the Paris Basin: From the Archaeological Feature to Global Analysis

François GILIGNY

Université Paris 1 Panthéon-Sorbonne – UMR Trajectoires, Paris, France
francois.giligny@univ-paris1.fr

Abstract

So far only a few intra-site spatial analyses have been carried out for Neolithic settlements in the Paris Basin. Case studies were performed mostly by means of distribution mapping and simple frequency calculations by material category, but rarely with statistical multivariate analysis.

Different cases can be recognised among the published studies: models were constructed for one type of feature, for sites with juxtaposition of elementary features, for the reconstruction of activity areas within an extensively excavated single phase occupational layer and recorded by surface unit and finally for the comparison between assemblages of domestic units.

One case study is presented here. The Neauphle-le-Vieux settlement (Yvelines) near Paris is an Early Neolithic settlement located in the Seine valley. It was assigned to the Post-Linear Pottery culture of the Paris Basin, the Villeneuve-Saint-Germain, dated between 5000 and 4700 BC.

In this small settlement that yielded a single occupation level and two pits, distribution maps for each material category made it possible to identify three refuse areas at the surface. A factor analysis highlighted differences within refuse areas and pits.

Activity areas were then reconstructed and explained based on the functional interpretation of the artefacts.

Keywords: Intrasite spatial analysis, multivariate statistics, early Neolithic.

1. From distribution maps to quantitative analyses: examples of spatial analyses in Neolithic sites of the Paris basin

Intra-site spatial analyses occupy a particularly important place within the study of pre- and protohistoric sites.

Since the palethnological approach developed by Leroi-Gourhan we perceive space ─ settlement or funerary space ─ as a social space and all space formed by individuals or groups as a space that acquired a social meaning. The analysis of the *hypogée* of Les Mournouards can be cited as being at the root of this approach (Leroi-Gourhan *et al.,* 1962). From this perspective a distinct number of case studies referring to Neolithic sites made it possible to better understand the funerary space as well as domestic space or activity areas.

The introduction of quantitative methods linked with intra-site spatial analysis in the 1970s provided archaeologists with more or less effective tools for identifying spatial patterning (Hodder and Orton, 1976; Hietala and Larson, 1984; Blankholm, 1991; Kroll and Price, 1991; Djindjian, 1991, 1997, 1999). Once these patterns had been identified, the phenomena which had triggered these spatial distributions still had to be reconstructed and, first and foremost, the cultural phenomena still had to be distinguished from natural phenomena related to differential preservation and post-depositional processes. The procedure of such an analysis is summarised as follows by F. Djindjian (1997: p. 13):

- identification of the significance of post-depositional processes;
- selection of categories of vestigial remains likely to reveal spatial patterns mirroring economic and social activities;
- spatial analysis carried out with appropriate methods;
- interpretation in palethnological terms of the structures obtained.

Only a small number of intra-site spatial analyses of Neolithic settlements in the Paris basin have used quantitative methods. A brief overview of the methods used leads us to state that in most cases we are dealing merely with map-based quantitative distributions and with frequency calculations of vestigial remains by category dominated by the archaeological feature and its internal structure.

1.1. Modelling at the scale of the feature

Spatial analysis carried out at the scale of the archaeological feature leads us to construct a model of space organisation depending on the archaeological feature. Two emblematic examples will be cited from the Neolithic in the Paris basin: an Early Neolithic house and a Late/Final Neolithic collective burial.

The model of the Linear Pottery house is doubtlessly one of the most famous models of spatial organisation of the European Neolithic (Soudsky, 1969; Coudart, 1998).

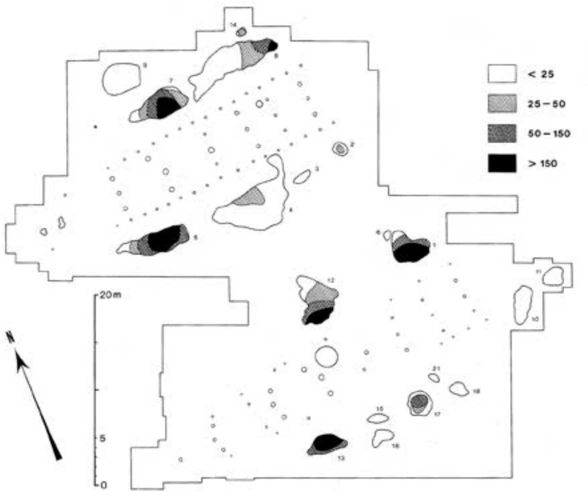


Figure 1: Projection onto plan sections of the cumulated quantities recorded from the pits of the Linear Pottery houses. House 1 and 2 of the Echilleuses site, total number of objects in numbers by 0.1 cubic metre (except for daub remains) (Simonin 1996: fig. 39).

The internal space of the Linear Pottery house is indeed subdivided into different parts (front, central, rear, separated by corridors) along a longitudinal axis. It is also structured with regard to its external spaces, more particularly with regard to the lateral so-called 'construction pits' dug into the soil to extract clay materials, most probably used for the making of daub and later re-used as refuse pits. The repeatedly observed regularity of the distribution of vestigial remains in these lateral pits made it possible to advance interpretative hypotheses with regard to lateral entrances (Chataigner and Plateaux, 1986;, Ilett *et al.*, 1982; Ilett *et al.,* 1986; Simonin, 1996; Bostyn *et al.,* 2003). The analyses of these distribution maps were made by projection on plane sections of the cumulative quantities recorded in the pits, the quantities being represented by proportional circles per square metre (fig. 1). A different method of representation consists of a projection along a longitudinal axis following the two walls (Bostyn *et al.*, 2003: fig. 62 and 215).

A second example refers to the burial features analysed by M. Sohn (Sohn, 2002, 2006). This author reconsidered the ordination proposed by A. Leroi-Gourhan with regard to the collective or individual character associated with the funerary goods of the collective burials in the Paris basin. She highlighted a spatial organisation of the collective grave goods according to the different areas within the gallery graves (*allées sépulchrales*), from the exterior to the interior: vestibule, sill, antechamber, burial chamber. She concluded that the situation is unclear and that it does not mirror segregation between the antechamber and the burial chamber. Individual grave goods were found primarily within the sepulchral layer and tended to be personal equipment (stone or bone tools, personal ornaments, metal objects). The collective grave goods were distributed within a fairly restricted space: the vestibule, the entrance sill or next to the inhumations. Late Neolithic pottery, stone axes and transversally perforated sleeves are mainly collective grave goods. Axes were closely associated with the sills and they were found in the circulation areas between the inhumation units or between the antechamber and the burial chamber. Further objects are in addition part of collective deposits: transverse arrowheads, blades, awls, sleeves and flakes. The spatial analysis proposed in the present study aims to highlight recurrent characteristics observed from drawings and plans. The quantitative analysis applied to the corpus was mainly based on the calculation of frequencies and it was visualised in a schematic form (Sohn, 2002: fig. 3).

The use of geographic information systems is still exceptional for the analysis of collective burials (Guillot and Guy, 1996). In most cases, simple distribution maps or plans of associations and refittings are presented, which are not connected with a data base (Chambon, 2003).

1.2. From the elementary feature to the site structure

A distinct number of sites are composed of juxtaposed or repeated features. This is more particularly the case for the enclosures with interrupted ditch and for the flint mines.

The analysis carried out for the Bazoches-sur-Vesles enclosure is based on the principle of evidencing structural geographic elements, more particularly crossing points (Dubouloz *et al.,* 1997). This enclosure is dated to the Middle Neolithic II and – for a circumference of 12,000 metres enclosing an area of 9.2 ha – it is composed of four interrupted ditch systems and one or two palisades. The investigations carried out in 1997 concerned only half of the previously excavated enclosure. The vestigial remains were projected onto a plan by units of two linear metres in order to identify concentrations and to place them in relationship to the circulation axes. An overall plan of the pottery concentrations plotted on six linear metres on either side of the apertures highlights their distribution along the circulation axes. This plan is correlated with the mapping of the deposits of more or less complete vases (Dubouloz *et al.,* 1997: fig. 5 and 7). The vestigial remains were mainly concentrated in two pits located close to the palisades and in the linear metres of the pits close to the circulation axes (65% of the pottery remains). A second significant element is the dissymmetry observed with regard to the concentrations on either side of the pits. Three types of possible circulation areas are evidenced by distinct characteristics : first, the complete 'passages' leading from the inner part of the enclosure to the outside and crossing both ditches and palisades with more or less significant concentrations with regard to the two instances recorded; second, the 'pseudo-passages' providing access from the exterior towards the palisade but without crossing it (7), and which present significant concentrations; third, the 'gateways', which make it possible to cross the palisade from the interior to the exterior but not to go through the ditches, are lacking in concentrations. Refuse areas and deposits distributed from the exterior to the interior of the enclosure therefore appear to be a dominant feature in Bazoches. This dominant feature was observed

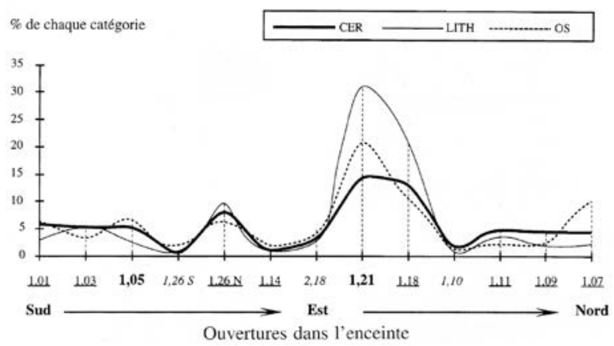


Figure 2: Bazoches-sur-Vesles. Frequency curve of the three categories of objects along the main entrances and deviations from the mean of the cumulated quantities of vestigial remains (Dubouloz et al. 1997: fig. 10 and 11).

in other enclosures of Central Europe, for example in the Sarup enclosure (Andersen 1997). The frequency curve of these three categories of features based on percentages and deviations from the mean rather highlights a fair correlation and evidences the most abundant concentrations (fig. 2).

Flint mines are a perfect example of a site with juxtaposed and regularly spaced features. This spacing echoes the exploitation strategy of the chalk substratum, which involved digging a mine shaft down from the surface level followed by the extraction of flint blocks according to several modalities: simple exploitation at the foot of the shaft, digging of bell-shaped pits, niches or galleries, and even horizontal galleries supported by chalk pillars. The spacing of the shafts at the surface depends on the volume exploited underground: the greater this volume, the greater the distance between the shaft heads at the surface. This model of distribution makes it possible to define a supposed quarrying area underground, based on the equidistance between the shafts according to the geographic model of Thiessen polygons. Initially this model was defined in order to model the relationship between the size of a settlement and the surrounding area according to a scheme of uniform distribution of sites located at the centre of a hexagonal cell. The boundaries of the hexagons or polygons are obtained from the perpendicular bisections of the segments connecting each point to its closest neighbours (Haggett, 1968). This model was used to draw the plan of the Jablines mine, one of the most important Neolithic flint mines of the Paris basin (Bostyn, Lanchon *et al.,* 1992). According to the aerial photographs, the total area occupied by this mine was estimated at nearly 40 ha. Excavations were carried out in an area encompassing 3.5 ha. Almost 800 shafts were identified, 56 of which were excavated. The pattern of the Thiessen polygons indeed matches the underground extension of the exploitation if this is superimposed on the area in which are located the 43 features effectively excavated in depth: space management at the surface consequently took into account the underground network (Bostyn and Lanchon, 1992: fig. 78). The size of the polygons tends to decrease downwards given that at the point where the flint seam is the deepest the galleries are the largest and consequently the distances between the shafts are the more significant. Based on this principle and on a mine with the same type of functioning it should be possible to estimate the extension of the underground exploitation from the distance between the shafts. This is, for example, the case for the Flins-sur-Seine mine, for which the depth of the exploitation can be estimated at 2 or 3 m.

1.3. Reconstructing work areas

The reconstruction of work areas requires spatial data stemming from systematic excavations carried out over an area which is not too small and which was not subjected to very intensive post-depositional processes.

The state of preservation of the Neolithic occupations in the Paris basin is not very favourable with regard to detailed spatial analysis given that in most cases the occupational layers were destroyed and only features dug into the soil were preserved. In some cases occupational layers were partially and unevenly preserved due to particular taphonomic conditions (materials trapped in shallow pits, overlaid by alluvial deposits or by accumulations of colluvium). In the absence of a more appropriate term these deposits are called 'layers'.

Among the Neolithic sites of the Paris basin suitable for intra-site spatial analysis very few Early Neolithic sites yielded occupational levels preserved in relationship with features dug into the soil. The best-known is Jablines 'La Pente de Croupeton', assigned to the Villeneuve-Saint-Germain group, and excavated over 2,500 m^2 during several campaigns between 1986 and 1991. Thanks to a colluvium deposit an occupational layer was preserved. The material was sampled by quarters of square metres and almost two-thirds (72.7%) was found within the occupational layer (Lanchon *et al.* 1997). The analysed area corresponds to two dwelling units and their comparison evidenced quantitative differences. Refitting maps, density maps and simple counting were used to process these data (Lanchon *et al.*, 1997: fig. 1 and 11). Several additional Villeneuve-Saint-Germain sites revealed a more or less well preserved occupational layer and density maps of the vestigial remains were drawn. At Pontpoint 'Le Fond de Rambourg' a living surface was also preserved around three dwelling units established on sloping sandy ground (Arbogast *et al.,* 1998). The accurate quantities of the vestigial remains contained in the occupational layer and in the pits respectively are not mentioned, but the distribution maps indicate higher densities for the pits. At Ocquerre 'La Rocluche', in the lower Marne valley, a production site of blades made from Bartonian flint yielded an occupational layer with two or three dwelling units in an area encompassing 4,800 m^2 (Praud *et al.*, 2009). In this case density maps also evidence concentrations of vestigial remains at the surface and in the features dug into the soil (ibid: fig. 65).

With regard to occupational layers in alluvial sedimentary context, some rare attempts were made to reconstruct the dwelling units and their task-specific activity areas from a palethnological perspective. The interpretation of these levels, as is the case for the palaeosols of Palaeolithic sites, supposes that the remains were not unduly displaced by post-depositional processes after their abandonment. Such an analysis was carried out for the site of Choisy-au-Bac 'Le Confluent', where wall effects and negative imprints were analysed (Prodéo, 1997). The wall effects were correlated with the posts in order to reconstruct the building the extent of which unfortunately reaches beyond the excavated area (Prodéo, 1997: fig. 8). The originality of this approach, greatly inspired by similar approaches in pile-dwelling sites the depositional context of which is quite different, is the reasoning based on the functional attribution of the artefacts. This approach was, however, criticised. Although it may be based on good intentions, the condition underlying a plausible reconstruction of work areas is indeed the need to obtain detailed functional data for the material the distribution of which is analysed. In this example two functional interpretations considered the flint tools – picks, chisels and tranchet type axes – to be related to wood crafting and the burins to testify to bone processing. However, since the functional analyses carried out on similar toolkits in the Paris basin, tranchet type axes were first and foremost associated with the processing of stone materials, although wood and other materials are not excluded, and burins were used for the processing of plant materials, most probably during activities related to the transformation of textile fibres (Allard *et al.,* 2004; Caspar *et al.,* 2005, 2007). Given that to date no functional analysis has been carried out at Choisy-au-Bac and that only one site attributed to this period – Louviers 'La Villette' (Giligny, 2005) – has been subject to functional analysis, these functions should not be considered a priori.

Data analyses of chemical elements are exceptional and only a single case study dealt with phosphorus levels (Fechner *et al.,* 2011). The elevated values with regard to concentrations provide support to the hypothesis that livestock was kept in the Early Neolithic buildings. By contrast, lower values were measured for Final Neolithic buildings and thus generate several hypotheses: shorter duration of occupation or functional difference with low-polluting domestic activities and the absence of livestock.

By contrast to occupational layers in pile-dwelling contexts, not a single site in the Paris basin has been subject to an analysis according to the principle of spatial organisation (Djindjian, 1988) or *unconstrained clustering* de R. Whallon (1984). An example of this approach was proposed by C. Tardieu at the sites of Charavines-les-Baigneurs and Egolzwil 3 (Tardieu, 2005). Sites with preserved occupational levels are rare in the Paris basin and most of the sites are represented by features dug into the soil which are partially preserved and disjoined.

1.4. Towards a global intra-site spatial analysis of the settlements

Currently, the most conclusive intra-site spatial analyses at the scale of the Neolithic sites in the Paris basin were carried out in two Early Neolithic sites: Cuiry-lès-Chaudardes and Poses.

The approach developed by L. Hachem with regard to the spatial analysis of the faunal remains stemming from Cuiry-lès-Chaudardes (Hachem, 1995, 1997) renewed spatial analyses of Danubian settlements through its methodology and its results. This analysis dealt with faunal remains, a rare record in that they are generally badly preserved in Danubian settlements and that they stem from an exhaustive excavation, another particularity of this site. The use of factorial correspondence analysis by domestic unit, applied to the quantities of faunal remains, enabled the author to propose a model of distribution highlighting significant differences with regard to the consumed species.

The spatial organisation revealed by this analysis is based on a partition of the village space into four quarters along two axes: the one east-west and the other north-south. The houses attributed to five occupational phases are unevenly distributed across the area with the first house erected in the eastern part. Nonetheless, a tendency of overrepresentation of some species appears to be recurrent during all the phases. The houses which group together the most significant proportion of domestic animals are located in the eastern part and those in which wild animals are dominant are located in the north-western quarter (Hachem 2011: fig. 141). Three species are particularly significant: cattle, wild boar and sheep. If the two eastern quarters are grouped together, three areas can be distinguished with regard to the faunal composition, and during all phases at least two houses were located in each of the three areas. In the eastern part cattle were the most abundant, in the north-western part wild boar and in the south-western part sheep (fig. 3).

This model triggers a social interpretation, more particularly with regard to a potential segmentation of the society, a model that may have more general value and may reveal the fundamental organisation of the society it is related to; it is therefore of great interest.

The settlement of Villeneuve-Saint-Germain de Poses 'Sur la Mare' is another example of global spatial analysis looking at differences in the composition of the vestigial remains associated with the domestic units (Bostyn *et al.,* 2003). Although the site yielded a rather large number of artefacts, i.e. pottery as well as flint industry (860 vases, 40,000 pieces), other remains such as animal bones prohibit the approach adopted at Cuiry-lès-Chaudardes. The number of faunal remains is indeed sufficient to form a global picture of the meat diet but not on a household scale. As for any analysis of Linear Pottery settlements, the chronological issue is a prerequisite for reasoning in terms of complementarities within a same phase. Pottery enabled the identification of two chronological phases corroborated by those identified from the lithic industry. Although the time interval represented by each phase cannot be determined, the authors propose, based on the interpretation of the plan, a continuous development between the two phases,

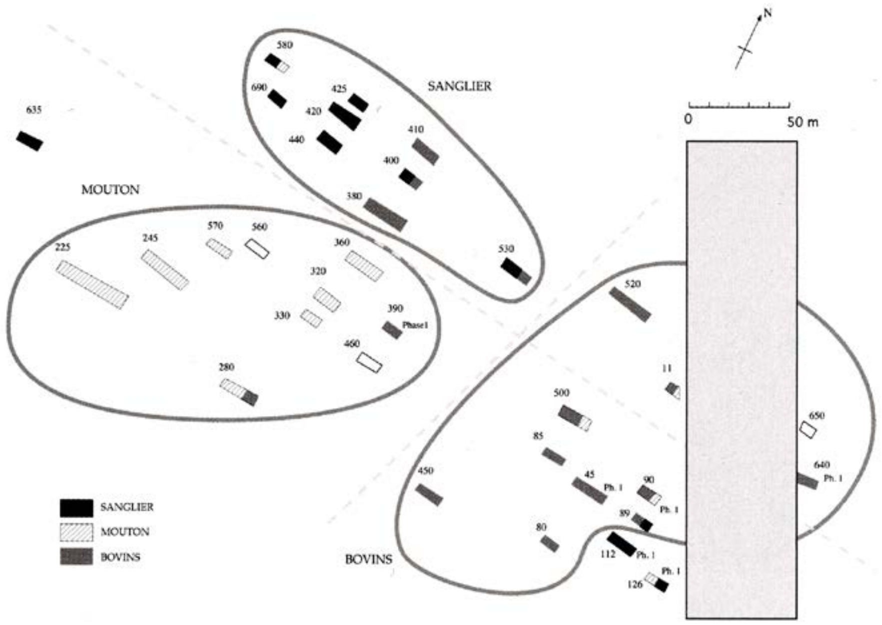


Figure 3: Spatial distribution of three species strongly represented in the refuse from the mean for domestic and wild animals (Hachem 2011: fig. 141).

reflecting an arrangement of the houses in two rows along a north-south axis that structures the village space. Of a total of eleven domestic units, the first phase comprises three units (40, 80 and 120), the second four units (70, 90, 6, 60), one unit cannot be clearly attributed (50) and three other incomplete units are not incorporated into the scheme (23, 60bis, 130). The trapeze-shaped ground plan of the houses is more strongly marked during the second than during the first phase, a chronological indication that fits well with the proposed pattern. The authors remind us that all the analyses led to the same conclusion, i.e. the systematic presence of a common set and the invariably unchanged composition of the pottery, lithic or sandstone assemblages associated with the house units which never present fundamentally varying characteristics. The differential representation of each type of object and consequently of each type of activity to which it is specific, highlights variations and gives them a meaning. From this perspective, differences were recognised within the phases and transposed schematically (Bostyn *et al.*, 2003: fig. 291). Several observations are instanced:

- houses 40 and 120, dated to the early phase, yielded a high number of burins made on blades and denticulates, whereas house 120 concentrates a large part of the grinding material and house 80 half of the polishing tools;
- houses 50 and 90, dated to the late phase, exhibit a high rate of blade products, by contrast to house 6; houses 6 and 60 group together most of the polishing tools made from sandstone; house 60 hosts the most abundant collection of large 'storage' containers.

Finally, with regard to importations, during the late phase, a contrast becomes apparent between arm rings made on slate and blades made on Bartonian flint, which leads to the assumption of a generalised redistribution in the village except for one house. The differences are also reinforced during the late phase.

2. Case study: Neauphle-le-Vieux 'Le Moulin-de-Lettrée': a partially preserved occupational level

2.1. Presentation

The site was excavated during investigations carried out by AFAN in 1996 on the occasion of rescue operations related to the planned road alignment of the detour road west of Paris (Giligny *et al.*, 1996, 1997). Several pre- and protohistoric occupations were identified and partially excavated on either side of the Guyonne River, a tributary of the Mauldre River, including one Early Neolithic site.

As a particularity the Neauphle-le-Vieux site yielded a partially preserved occupational layer, which is preserved at the bottom of a brown silty clay layer, 30 to 60 cm thick, followed over nearly 70 metres. It is characterised by the presence of vestigial remains distributed within a 10 cm thick stripe revealing at its base a large number of elements of a large size and often in a horizontal position. The formation of brown silty clay continued after the abandonment of the site. Significant bioturbation explains the vertical distribution of remains. We are not dealing here with an archaeological layer strictly speaking but rather with a degraded condition. The richest archaeological features as well as the most important concentrations of vestigial remains were located in a shallow pan along the southern limit of the excavated area.

A detailed excavation was carried out on 1500 m^2 of this archaeological layer, with the sampling of the vestigial remains made by square metre and the piece-by-piece dismantling of some outstanding concentrations. The objective that guided the excavation was the spatial

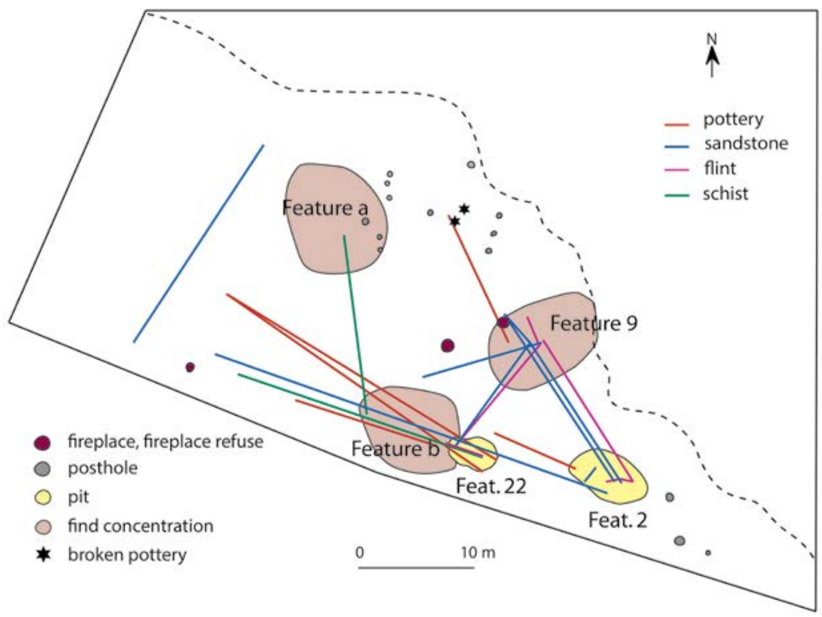


FIGURE 4: NEAUPHLE-LE-VIEUX. PLAN OF THE REFITTINGS OF DIFFERENT CATEGORIES OF OBJECTS IN LAYERS AND IN FEATURES AT A DISTANCE OF MORE THAN ONE METRE.

interpretation of the distribution of the vestigial remains within the occupational layer. Sites in which such a level is preserved, are indeed rare and this type of site provides elements that make it possible to estimate the volume of vestigial remains, which is not possible when only the features dug into the soil are preserved.

The method we intend to present here consists of characterising – once the areas of concentrations of vestigial remains have been identified – the objects with regard to the task-specific activities they attest to. This method is based on the functional data of the artefacts, if possible stemming from the same site and on contextual data linked with the global spatial organisation of the site and with the function of the features.

2.2. Refittings and distribution maps

Refittings of pieces stemming from the filling of the occupational layer, from open-air deposits and from pit deposits are proof of the temporal consistency of the assemblage. Long- distance refittings (from 10 to 35 metres) could be made on all types of materials (fig. 4). The dominant south-east/north-west and north-south oriented distribution of the refittings follows the slope and these displacements are interpreted as being contemporaneous with the occupation of the site. The strong temporal and spatial consistency of the excavated assemblage therefore justified an intra-site spatial analysis in order to reconstruct activity areas from a palethnological perspective.

Based on the inventories of the objects by category and by square metre, computer processing made it possible to draw distribution maps using quantitative mapping software: global distributions by categories (in numbers and weight), and distributions of distinct types of artefacts. These distribution maps made it possible to define areas of concentration (fig. 5) that correspond to features dug into the soil (features 2, 22) or to open-air concentrations of artefacts within the layer (features a, b, 9).

The distribution of products stemming from flint debitage (fig. 6) reveals significant concentrations in the pits (feature 2 and feature 22) and two less significant concentrations in their close vicinity (feature 9 and feature b). In the same manner, the tools were also clustered in and around the pits. In addition, a third area of concentration (feature a) appears in the north-western part.

Pottery shows significant clusters corresponding to the locations of the two pits, and a third further north comprised of several crushed pots.

The sandstone pieces are clustered in the two pits and in the features 9 and a. Fragments with unmodified surface are primarily distributed in feature 22; those with worked surface are well represented in features 2, 9 and a.

Differences emerged from the mapping of the artefact types: blades, waste left by retouching, microburins and *piquants trièdres* were grouped together in features 2, 22 and 9. The burins were most abundant in features 2 and 9, the denticulates in feature 22, the hammer stones in features 22, 9, a, and b. By contrast, other tools, for example the end-scrapers, are distributed more evenly.

It can therefore be assumed that these distributions mirror specific activities which took place next to refuse areas.

23. Description of the activity areas based on data tests and data analysis

The extension of open-air concentrations was delimited according to the density of archaeological remains. All the vestigial remains were counted by zone in order to carry out a chi-squared test. The chi-squared values were calculated from two data sets grouping together the debitage products and the tools (tab. 1). The deviations from independence are also expressed in percentages.

The result highlights a strong relationship between the types of flint artefact and the features ($p < 0.8879$ for the

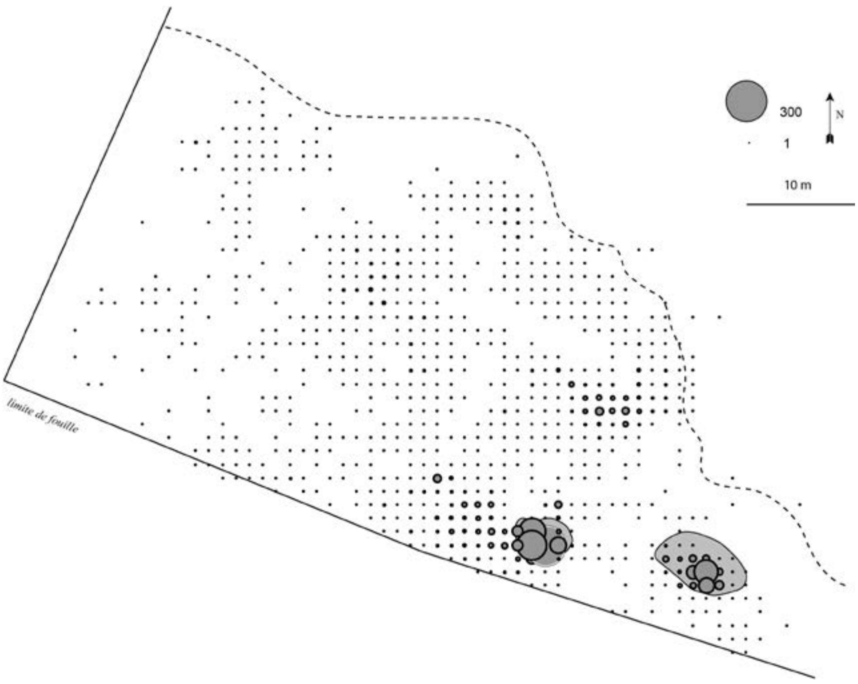


Figure 5: Neauphle-le-Vieux. Distribution of the flint debitage products by square metre.

tools (n=65; X = 80.9) and p= 0.9999 for the debitage products (nu= 25; X=80.6). A second significant result obtained from the test is the low chi-squared value of the 'remaining' assemblage, i.e. all the items located outside the pits and the concentrations, with regard to both tools and debitage products. We can therefore describe each area based on the surplus or the lack of distinct pieces for the significant chi-squared values (generally representing about 5% of the total count).

A second chi-squared test carried out on the data set grouping together the sandstone pieces is also significant (p= 0.992, with n = 60, c = 89).

A factorial correspondence analysis was applied to the data matrix of the flint, sandstone and *meulière* artefacts. The characteristic 'other' was introduced as an additional element of the analysis. The first factorial axis (48.1% of the characteristic values) contrasts the flint flakes with the sandstone debris with unworked or worked surface, the hammer stones made of *meulière* (a siliceous sedimentary rock), the hammer stones made of sandstone and the flint hammer stones associated with the retouched pieces. This axis evidences the particularity of the area of feature a as opposed to feature 22 concentrating most of the flint flakes. Axis 2 (22.8% of the characteristic values) contrasts the burins, micro-burins, side-scrapers and grinding stones with the flakes and sandstone debris, axis 3 (15.8% of the characteristic values) contrasts the blades with the cores and microburins in order to differentiate feature 2 and to contrast it to the features 9, b and a. Axis 4 contrasts the axe tools with the *meulière* hammer stones and again individualises feature a.

On the factorial level axis 1 - axis 2 (fig. 19), the seriation of the objects and variables displays two superimposed elongated horseshoe curves on axis 1 with a distinct number of variables that are not incorporated (sandstone flakes, flake with worked surface, microburin, side-scrapers, grinding stones, sandstone hammer stone, grooved polishing stone). Within this seriation are distributed, from the left to the right, the features 22 and b, quite close together, followed by feature 2, feature 9, the whole of the layer ('other') and then by feature a. Ascendant hierarchical clustering makes it possible to group together the following assemblages by pairs: feature 22 with feature b, feature 2 with feature 9, then 'others' with feature a (fig. 6).

absolute frequency	flake	blade	microburin	nucleus	rework waste	blade preparation	TOTAL DEBITAGE	retouched tool	scraper	denticulate	burin	notch	splintered piece	hammer	truncation	arrow	borer	sickle	axe	side-scraper	other	TOTAL TOOLS
feat a	316	4	0	26	0	2	348	39	15	6	9	5	3	13	1	3	1	0	0	0	3	98
feat b	733	5	7	29	6	2	782	47	17	13	10	2	7	5	3	0	0	1	2	1	5	113
feat 9	683	16	11	47	5	2	764	81	27	20	20	8	8	12	3	1	3	2	0	3	2	190
feat 2	698	29	3	19	7	6	762	52	18	13	17	11	5	6	1	0	3	2	0	3	2	133
feat 22	1040	25	5	31	15	1	1117	59	13	20	7	5	0	8	4	4	1	1	0	0	1	123
area	2313	45	9	109	23	11	2540	255	69	62	37	34	23	39	12	12	8	6	9	4	5	588
Total	5783	124	35	261	56	24	6283	533	159	134	100	65	46	83	24	20	16	12	11	11	18	1245
features	3470	79	26	152	33	13	3805	278	90	72	63	31	23	44	12	8	8	6	2	7	13	657

Table 1: Data table for the Chi-suare test and factor analysis.

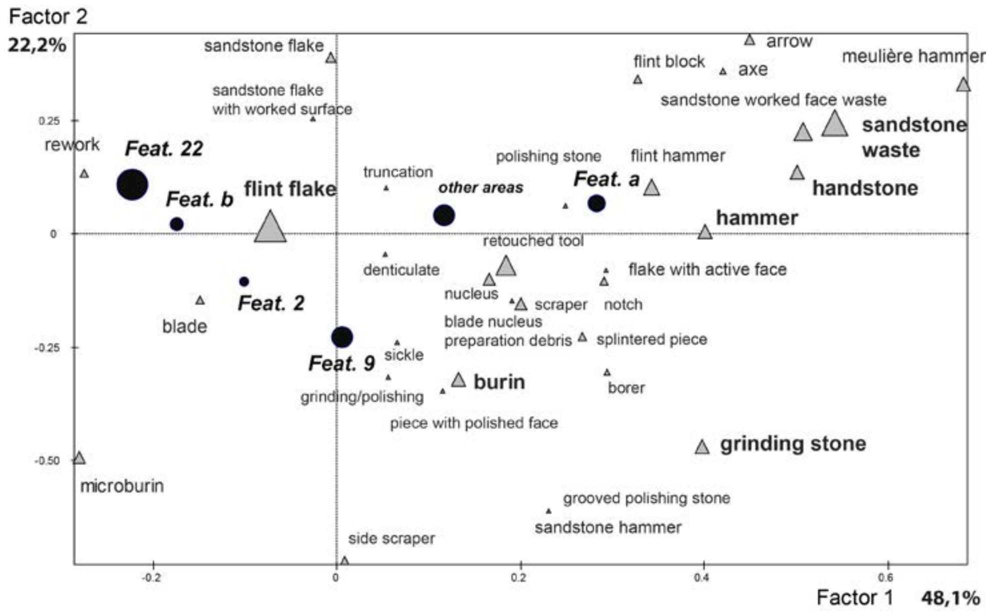


Figure 6: Neauphle-le-Vieux. Factorial map axis 1 - axis 2 stemming from the correspondence analysis of the counts of the flint and sandstone tools.

The summary figure portrays the main characteristics of each area. First of all a north-south opposition can be recognised (fig. 20): features a and 9 include mostly tools; features 2, 22 and b contain a larger number of debitage products and clearly fewer cores. A more detailed analysis shows that features 2 and 22 exhibit a surplus of blades and waste stemming from blade debitage. By contrast feature b reveals a lack of products stemming from blade debitage (the flake production is evenly distributed). Moreover, the pit feature 22 and the open-air concentration b are complementary with regard to blade production; in feature b the blade products are in a minority, whereas microburins and *piquants trièdres*, by-products of intentional blade knapping, are overrepresented. In features 2, 22 and b the best represented tools differ: burins are most abundant in feature 2, denticulates in feature 22 and splintered pieces in feature b.

Features a and 9 also differ. They contrast with each other more particularly with regard to the representation of blade blanks and microburins as well as to the representation of retouched blades and burins. An association of cores and hammer stones (made of flint and of siliceous rock of the *meulière* type) can be identified in feature a. It can be assumed that the cores which were not discarded in the pits were stored in this place to be re-used later as hammer stones. Most of the found hammer stones were indeed re-used cores. These tools were probably used during the shaping process (knapping, pecking) of sandstone tools as is suggested by the large number of sandstone debris and hand stones, frequently reshaped from debris. In feature 9, an association of cores and grinding stones exceeding the usual counts was identified.

24. Functional interpretation

These results lead us to conceive a functional interpretation of the space concerned (fig. 7), organised according to the activities that took place there.

The two east-west-oriented pits are most probably construction pits located next to the northern wall of a house located outside the excavated sector. These pits were used as domestic refuse pits, in which a variety of materials were rejected: pottery, animal bones, notably blade debitage waste and distinct, abundantly represented flint tools such as burins in feature 2, denticulates and retouched flakes in feature 22.

The high density of vestigial remains scattered over the surface adjacent to the pits (feature b) may identify this area as a work and refuse area in a complementary relationship with the adjacent pit feature 22: in this locus the Neolithic inhabitants knapped blades according to the microburin technique and abandoned the characteristic useless fragments, and they also used and abandoned splintered pieces.

Feature 9 in the northern part was a working and refuse area, which specialised more particularly in blade knapping and the manufacturing (as illustrated by the presence of burin spalls) and use of burins. The recurrence of distinct objects and complementarity of the refuse linked with blade debitage and the manufacturing of tools made on blades in feature 9, in the pit feature 2 as well as in feature b, suggests that a relationship existed between these three areas, linked to the same activity or the same activities.

In the north-western part feature a may be considered as a grinding area requiring the storage of cores related to the use (and the abandonment) of hammer stones and hand stones. In addition, a set of in situ crushed vases was discovered close to this area.

All these observations enable us to study the organisation of the excavated space as well as distinct human behaviours. We are dealing with the northern periphery of a house unit wherein a distinct number of domestic activities carried out by the inhabitants of the site were evidenced.

3. Conclusion

The examples of intra-site spatial analyses applied to Neolithic sites in the Paris basin are still not very abundant. The establishment of a protocol of applied intra-site spatial analysis requires the development of a model adapted to the features and the task-specific activity areas they were related to. In this case, the map-based quantitative analysis is an initial approach and it is precisely the most frequently implemented technique. Moreover, the analysis of spatial patterns of the vestigial remains is more rarely attempted. Qualitative models are sometimes proposed based on the presence or absence of distinct types of objects given the lack of statistical analyses.

In the case of occupational layers, once the areas of concentrations have been identified, the activities mirrored by the associations of objects can be analysed thanks to a variety of statistical methods, such as chi-squared tests or data analysis. Our objective is the reconstruction of work areas within a palethnographic approach.

In the case of settlement analysis the comparison of the assemblages stemming from domestic units is conditioned by a clear chronology. Variability factors inherent to one and the same temporal stage may occur between the houses and stylistic variability is not constant over time. The taking into account of different artefact categories during the factorial correspondence analysis efficiently emphasises the specificity of each domestic unit. Recently excavated sites, whether these be settlements, enclosures or collective burials, provide documentation that is just waiting to be processed with suitable methods.

Translation : Karoline Mazurié de Keroualin

References

ALLARD P., AUGEREAU A., BEUGNIER V., BOSTYN F., BURNEZ-LANOTTE L., CASPAR J.-P., GILIGNY F., HAMARD (D.), MARTIAL E., PHILIBERT S. (2004). Fonction des outillages lithiques dans le Bassin parisien au Néolithique, In: P. Bodu et C. Constantin (eds). *Approches fonctionnelles en Préhistoire,ctes du XXVe Congrès Préhistorique de France, Nanterre-nov. 2000,* Paris: Société Préhistorique française, pp. 181-192.

ANDERSEN N. H. (1997). *The Sarup Enclosures. The Funnel Beaker Culture of the Sarup site including two causewayed camps compared to the contemporary settlements in the area and other European enclosures*. °Arhus. Jutland Archaeological Society Publications, 33, vol. 1.

ARBOGAST R.M., BOSTYN F., LORIN Y., PRODEO F. (1998). Un nouveau site d'habitat du groupe de Villeneuve-Saint-Germain à Pontpoint 'Le Fond de Rambourg' (Oise). *In* : Actes du XXIIIe colloque interrégional sur le Néolithique, Bruxelles 1997. *Anthropologie et préhistoire*, 109, pp. 41-62.

BLANKHOLM H. P. (1991) *Intrasite Spatial Analysis in Theory and Practice*. Aarhus: Aarhus University Press.

BEURION, C., BILLARD, C., BOSTYN, F., GUILLON, M., HACHEM, L., HAMON, C., LANCHON, Y., PRAUD, I., RECKINGER, F., ROPARS, A., MUNAUT, A.-V., (2003) *Le village Villeneuve-Saint-Germain de Poses 'Sur la Mare' et les sites de la boucle du Vaudreuil*. Paris: Société Préhistorique Française (travaux 4).

BOSTYN, F., LANCHON, Y. (dir.); BOGUSZEWSKI, A., FRUGIER, C., JEREMIE, S. *et alii*. (1992). *Jablines,*

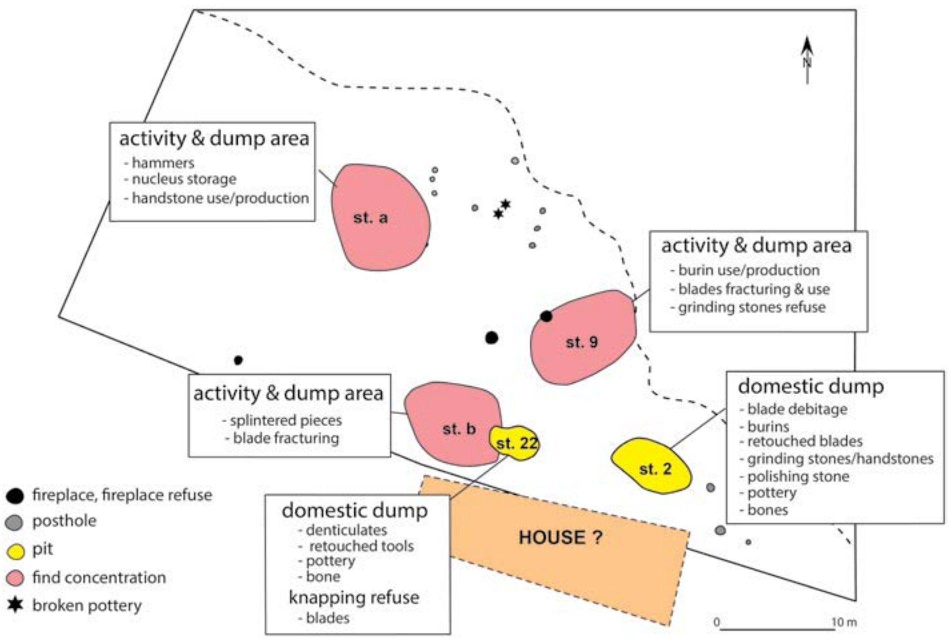


Figure 7: Neauphle-le-Vieux. Functional interpretation of the activity areas and the refuse areas.

Le Haut Château (Seine-et-Marne) : une minière de silex au Néolithique. Paris : Maison des Sciences de l'Homme, Documents d'Archéologie Française 35.

CASPAR, J.-P., MARTIAL, E. and FERAY, P. (2005) Identification et reconstitution des traces de teillage des fibres végétales au Néolithique. *Bulletin de la Société préhistorique française*, 102, pp. 867-880.

CASPAR, J.-P., MARTIAL, E. and FERAY P. (2007) Le teillage des plantes fibreuses au Néolithique. *In* : *Relations interrégionales au Néolithique entre Bassin parisien et Bassin rhénan*, Actes du 26ème colloque interrégional sur le Néolithique, Luxembourg, 89 novembre 2003, Archaeologia Mosellana, 7, p. 613-623.

CHAMBON P. (2003) *Les Morts dans les sépultures collectives néolithiques en France. Du cadavre aux restes ultimes,* Paris, CNRS editions (XXXVe Supplément à Gallia Préhistoire).

CHATAIGNER, C., PLATEAUX, M. (1986) Analyse spatiale des habitats rubanés et informatique, *Bulletin de la Société préhistorique française,* 83 (10), 1986, p. 319-324.

DJINDJIAN, F. (1988) Improvement in intrasite spatial analysis techniques, Computer and quantitative methods in archaeology, In: Ratz S.P.Q. (ed.). Computer and Quantitative Methods in Archaeology 1988, Oxford Bar International series 446, pp. 95-106.

DJINDJIAN, F. (1991) *Méthodes pour l'archéologie*. Paris: Armand Colin.

DJINDJIAN, F. (1997) L'analyse spatiale de l'habitat pré- et protohistorique. Perspectives et limites des méthodes actuelles. *In* : Auxiette (G.), Hachem (L.), Robert (B.) (dir.) *Espaces physiques, espaces sociaux dans l'analyse interne des sites du Néolithique à l'Age du Fer*. Actes du 119e congrès National des Sociétés Historiques et Scientifiques, Amiens, 1994. Paris : CTHS, pp. 13-21.

DJINDJIAN, F. (1999) L'analyse spatiale de l'habitat : un état de l'art. *Archaeologia e Calcolatori*, *10, pp.*17-32.

DUBOULOZ, J., HAMARD, D., LE BOLLOCH, M. (1997) Composantes fonctionnelles et symboliques d'un site exceptionnel : Bazoches-sur-Vesle (Aisne), 4000 ans av. J.-C. In: Auxiette G., Hachem L., Robert B. (eds.). *Espaces physiques, espaces sociaux dans l'analyse interne des sites du Néolithique à l'Age du Fer*. Actes du 119e congrès National des Sociétés Historiques et Scientifiques, Amiens, 1994. Paris, CTHS, 1997p. 127-144.

FECHNER, K., DE LIL, A., CLAVEL, V., HUS, J., TEHEUX, E. (2011) Cartographie du phosphore dans des bâtiments allongés du Néolithique dans le nord de la France, en Belgique et au Luxembourg, *In* : F. Bostyn, E. Martial, I. Praud (eds.), *Le Néolithique du Nord de la France dans son contexte européen : habitat et économie aux 4e et 3e millénaires avant notre ère*, Actes du 29e colloque interrégional sur le Néolithique, Villeneuve-d'Ascq, 2-3 octobre 2009, Revue Archéologique de Picardie, n°28, 2011, pp. 275-298.

GILIGNY, F., MARTIAL, E., PRAUD, I. (1996) Le Villeneuve -Saint-Germain de Neauphle-le-Vieux 'Le Moulin de Lettrée' (Yvelines). In *Internéo 1*, Paris : Association pour les Etudes interrégionales sur le Néolithique/Société Préhistorique Française, pp.15-31.

GILIGNY, F. (ed.) (2005) *Louviers 'La Villette'. Un site Néolithique moyen en zone humide.* Document Archéologique de l'Ouest, Rennes : Presses Universitaires de Rennes.

GILIGNY, F. (2008) Une méthode d'analyse spatiale des rejets domestiques par les remontages céramiques. Etude d'un site néolithique jurassien: Chalain station 3, *Archeologia e Calcolatori*, 19, pp.173-186.

GOMART, L. (2010) Méthodes d'étude technologique de corpus céramiques issus de contextes détritiques : le cas des vases retrouvés en contexte d'habitat rubané dans la vallée de l'Aisne, *Les Nouvelles de l'archéologie*, 119, pp. 25-29.

GUILLOT, H., GUY, H. (1996) L'utilisation d'un système d'information géographique (SIG) pour l'étude d'une sépulture collective, *Internéo* 1, p. 103-112.

HACHEM, L. (1997) Structuration spatiale d'un village du Rubané récent, Cuiry-lès-Chaudardes (Aisne). Analyse d'une catégorie de rejets domestiques : la faune. In : *Espaces physiques espaces sociaux dans l'analyse interne des sites du Néolithique à l'Age du Fer* (Actes du 119e congrès national des sociétés historiques et scientifiques, Amiens 1994), Editions du CTHS, Paris, pp. 245-261.

HACHEM, L. (2011) *Le site néolithique de Cuiry-lès-Chaudardes - I. De l'analyse de la faune à la structuration sociale.* Internationale Archäologie 120, Rahden/Westf. : Verlag Marie Leidorf GmbH.

HAGGETT, P. (1968) *L'analyse spatiale en géographie humaine*. Paris: P.U.F.

HIETALA, H. J. and LARSON, P. A. (1984) *Intrasite spatial analysis in archaeology* Cambridge/New York: Cambridge University Press.

HODDER, I., ORTON, C. (1976). *Spatial Analysis in Archaeology*. Cambridge, Cambridge University Press.

ILETT, M. J., CONSTANTIN, C., COUDART, A., DEMOULE J.-P. (1982). The late Bandkeramik of the Aisne valley. Environment and spatial organisation. *In* : *Prehistoric Settlements patterns around the southern North Sea*, Leyde, Leiden University Press, 1982, (Analecta Praehistorica Leidensia, 15), pp. 45-61.

ILETT, M. J., PLATEAUX, M., COUDART, A. (1986). Analyse spatiale des habitats du Rubané récent. Problèmes actuels. In: Demoule J.-P., Guilaine J.

(eds.). *Le Néolithique de la France*, Paris, Picard, pp. 131-140.

KROLL, E.M., PRICE, T.D. (eds.) (1992). *The Interpretation of Archaeological Spatial Patterning*. New York: Plenum.

LANCHON, Y., BOSTYN, F., HACHEM, F. (1994). L'étude d'un niveau archéologique Néolithique et ses rapports à la compréhension d'un site d'habitat : L'exemple de Jablines, 'La Pente de Croupeton' (Seine-et-Marne). In: G. Auxiette, L. Hachem, B. Robert (eds.). *Espaces physiques, espaces sociaux dans l'analyse interne des sites du Néolithique à l'Age du Fer*. Actes du 119e congrès National des Sociétés Historiques et Scientifiques, Amiens, 1994. Paris: CTHS, pp. 327-344.

LEROI-GOURHAN, A., BAILLOUD, G., BREZILLON, M. (1962). L'hypogée des Mournouards (Mesnil-sur-Oger, Marne), *Gallia-Préhistoire* 5, pp. 23-133.

PRAUD, I., BOSTYN, F., HAMON, C., LANCHON, Y., CASPAR, J.-P., DIETSCH-SELLAMI, M.-F., MARTIAL, E., MICHEL, L. and VACHARD, D. (2009). *Le Néolithique ancien dans la basse vallée de la Marne : un site Villeneuve-Saint-Germain producteur de lames en silex tertiaire à Ocquerre 'La Rocluche' (Seine-et-Marne)*, Paris: Société préhistorique française (Travaux 9).

PRODEO, F. (1997). Une méthode d'analyse des sites fossilisés en contexte alluvial. L'exemple de Choisy-au-Bac 'Le Confluent' (Oise) au début du néolithique moyen II. In: G. Auxiette, L. Hachem, B. Robert (eds.). *Espaces physiques espaces sociaux dans l'analyse interne des sites du Néolithique à l'Age du Fer*, 119è Congrès CTHS, 1997, pp. 207-232.

SIMONIN, D. (1996). *Les habitats Néolithiques d'Echilleuses (Loiret), Analyse spatiale des documents archéologiques*. Edition de la Fédération Archéologique du Loiret, Revue Archéologique du Loiret, 21-22.

SOUDSKY, B. (1969). Etude de la maison néolithique. *Slovenska Archeologia* 17, pp. 5-96.

SOHN, M. (2002). Place et rôle du mobilier dans les sépultures collectives du Bassin parisien à la fin du Néolithique. *Bulletin de la Société Préhistorique Française*, 99 (3), pp. 501-520.

SOHN, M. (2006). *Du collectif à l'individuel : évolution des dépôts mobiliers dans les sépultures d'Europe occidentale de la fin du IVe à la fin du IIIe millénaire av. J.-C.* Thèse de doctorat, Université de Paris 1.

SOHN, M. (2008). Entre signe et symbole. Les fonctions du mobilier dans les sépultures collectives d'Europe occidentale à la fin du Néolithique. *Préhistoire méditerranéennes* 14, 2008, pp. 53-71. Available from: http://pm.revues.org/index317.htm [consulted February 6, 2012]

TARDIEU, C. (2005) L'outil multidimensionnel dans l'analyse spatiale en milieu lacustre : l'exemple des habitats d'Egolzwil 3 (Ch) et de Charavines-les-Baigneurs B3 (F), *ArchéoSciences* 29, pp. 133-150. Available from: http://archeosciences.revues.org/628 , [consulted October, 27, 2011]

WHALLON, R. (1984). Unconstrained clustering for the analysis of spatial distributions in archaeology. In Hietala H. (ed.). *Intrasite spatial analysis in archaeology*, New York: Cambridge university Press, pp. 242-277.

Statistical and Mathematical Models for Archaeological Data Mining: A Comparison

Nevio DUBBINI,[1] Adam LODOEN[2]

[1] University of Pisa, Pisa, Italy
[2] Bournemouth University, Fern Barrow, Poole, Dorset, UK
nevio.dubbini@gmail.com , alodoen@bournemouth.ac.uk

Abstract
Within the field of archaeological predictive modelling, a range of analysis techniques and methods are possible. However, given the amount of archaeological data now available, it may be necessary to develop new automatic, computer-based methods to process and analyse the archaeological data, and to rethink methods currently in use. Specifically, methods that make use of patterns and spatial correlations within the datasets should be explored. To this aim, this paper compares the potential of mathematical and statistical methods to use existing data for archaeological prediction and analyses.

Keywords: Predictive modelling, Statistical modelling, Kriging, Mathematical modelling

1. Introduction

The last two decades has seen a huge increase in the amount of data held by archaeological data bases. For many datasets, a point has been reached where manual analyses of the data by individual researchers or even groups of researchers are impractical. Ultimately, modelling using manual means becomes impossible if the task is to model the whole dataset rather than individual parts. This is due not just to the huge amounts of data but also to its heterogeneous and complex nature. As a consequence, it is necessary to develop automatic, computer-based methods to process and analyse the archaeological data. One possibility would be to use archaeological predictive modelling techniques to mine existing data sets.

The aim of this paper is to compare the performance of a statistical prediction method (a cokriging model) with a mathematical method (a PageRank based model). Both models make use of environmental variables as well as cultural variables sensu stricto to try to predict medieval agricultural land use. Cultural variables *sensu stricto* were defined by Verhagen *et al.* (2007, 205) as 'measurable attributes of the archaeological sample that are not related to an environmental factor' (e.g. site location, size, functional type, or period of occupation). When discussing cultural variables sensu stricto, it is important to note that these cultural variables are actually part of the archaeological dataset. In other words, any models using cultural variables sensu stricto must use the character and / or location of recorded archaeological remains, in addition to environmental factors, to make their predictions. The two methods that we compare have been chosen since they are able to use cultural variables *sensu stricto* as inputs.

For automatic computerised analysis of data many possibilities exist in the general framework of statistical or mathematical models. Within the field of predictive modelling, statistical methods have been used since the 1980s and were pioneered by Kenneth Kvamme and others (e.g. Kvamme, 1980, 1983; Scholtz, 1981). As a consequence, these methods are relatively well understood by archaeological researchers. However, for the analyses of large amounts of heterogeneous archaeological data these methods may need to be adjusted and rethought.

The use of mathematical methods for archaeological predictive modelling is a more recent development (see e.g. Wheatley and Gillings (2002) and references therein), and much less well understood by the archaeological community. A mathematical model is distinguished from a statistical one mainly in making assumptions about the model: in the case of mathematical models these assumptions are put in the form of rules governing the system under study (e.g. equations), while in the case of statistical models the assumptions deal only with probability distributions. A graphical illustration of such difference is given in figure 1. With this distinction in mind, we observe that mathematical models are still uncommon in archaeological predictive modelling. Though mathematics increasingly find applications in many fields of science, including fields commonly thought of as far removed from mathematics, such as economics and finance (Roman, 2004), medicine and biology (Epstein, 2008), social networks and search engines (Langville *et al.*, 2006).

2. The data set

The study area is the former government districts of Kerrier and Penwith in Cornwall, southwest England (figure 2). The size of the study area is approximately 805 square kilometers.

The data chosen for the modelling was medieval data from the former local government districts of Kerrier and

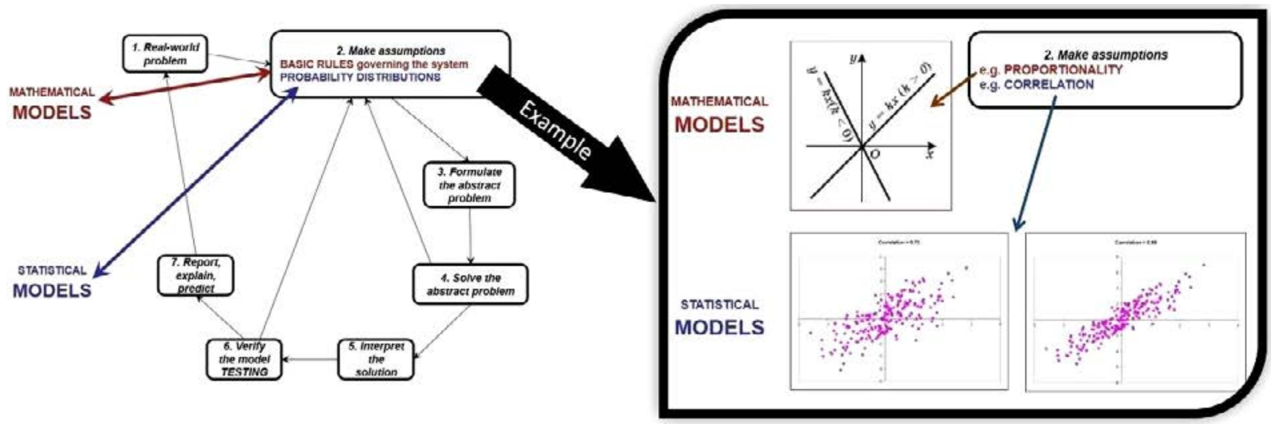


FIGURE 1. THE PROCESS IN DEVELOPING MATHEMATICAL AND STATISTICAL MODELS, WITH AN EXAMPLE OF THEIR MAIN DIFFERENCE, I.E. IN MAKING ASSUMPTIONS ABOUT THE SYSTEM UNDER STUDY.

Penwith in Cornwall, England. The data was taken from the Cornwall Historic Environment Record (Cornwall County Council 2013). 3264 data points in total, divided in the following 12 categories before modelling: Maritime, Agricultural, Settlement, Rabbit warrens (pillow mounds), Tin mining related, Animal husbandry, Manor houses, Wells, Deer Parks, Stone quarries, Chapels and Water mills.

Explanation of the data categories:

- Maritime: (Remains directly related to maritime activities): Quays, harbours etc.
- "Agricultural (Features directly related to the use of the land for agriculture): Ridge and furrow, headlands, clearance cairns (created through the clearing of the land of large stones to make ploughing easier), field system evidence etc.
- Settlement: This is almost exclusively an historical category, since very little actual archaeological evidence exists for medieval settlement in this area. However, the location of many are known from historical sources. A typical settlement would consist of 2-6 messuages (dwelling houses /farm houses). The settlement pattern in the area is highly dispersed.
- Rabbit warrens (pillow mounds)
- Tin mining related: A huge range of different kind of features, including shafts, streamworks, prospecting pits etc., all directly associated with tin mining.
- Animal husbandry: Features associated with animal husbandry – mainly different kinds of enclosures and dew ponds.
- Manor houses

FIGURE 2: THE STUDY AREA. SOURCES: ESRI, HERE, DELORME, TOMTOM, INTERMAP, INCREMENT P CORP., GEBCO, USGS, FAO, NPS, NRCAN, GEOBASE, IGN, KADASTER NL, ORDNANCE SURVEY, ESRI JAPAN, METI, ESRI CHINA (HONG KONG), SWISSTOPO, MAPMYINDIA, © OPENSTREETMAP CONTRIBUTORS, AND THE GIS USER COMMUNITY.

- Wells: Most of these are substantial stone structures, and most are 'holy' or sacred in some way.
- Deer Parks
- Stone quarries
- Chapels
- Water mills: This category could further be divided into fulling mills (wool processing), corn mills, and Tin processing mills.

50% of the Agricultural data were removed from the data before the modelling to make it possible to assess the performance of the models.

Other variables used

Only two purely environmental variables were used for the modelling:

- Slope: The slope of the ground surface measured in degrees, where 0° denotes a perfectly horizontal surface and 90° denotes a perfectly vertical surface.
- Distance to coastline: The distance to the nearest coastline measured in meters.

3. The statistical model: cokriging based

There may be a number of statistically significant correlations within archaeological datasets that could potentially be used for predictive purposes. How the data are correlated spatially is perhaps the most important of those. A possible way of making use of those correlations would be to base the predictions fully or partially on what is sometimes called Tobler's first law of geography. This principle states that 'Everything is related to everything else, but near things are more related than distant things' (Tobler 1970), and is also known as *spatial autocorrelation*.

Kriging is a method (or more precisely a family of methods) of interpolation which uses statistical methods to choose a model of spatial autocorrelation which minimises the error of the predictions (Lloyd & Atkinson 2004). Kriging is a complicated and vast subject, and it is not possible to here give an exhaustive description of the method. However, in very general terms, the way to estimate an unknown value Z in location x_0 using some kriging method is to:

a) Construct a model of spatial dependence;
b) With the help of this model, calculate a series of weights w_i for each location x_0, that determine the influence that known values have on that location;
c) Calculate the value Z through building a weighted average of a sample of known values.

Kriging has found uses in areas as diverse as soil science (Marchetti *et al.,* 2012; Ballabio *et al.*, 2012), prediction of air pollution (Robinson *et al.,* 2013), climatology (Aalto *et al.*, 2013), mining (van der Meer, 2006), and even real estate appraisal (Liu, 2011). Kriging has been used in archaeology for very different purposes, e.g. deposit modelling (Vermeer *et al.*, 2014), population estimation (Jones & DeWitte, 2012) and landscape reconstruction (Benito-calvo *et al.*, 2008).

A number of variants of the kriging method exists, among which:

- *Simple kriging* assumes known stationary means;
- *Ordinary kriging* assumes unknown stationary means;
- *Universal kriging* assumes a general polynomial trend model;
- *Indicator kriging* estimates the probability that a value at would exceeds (or conversely, is lower than) a chosen threshold value;
- *Cokriging* makes use of additional measured values that correlate with the predicted quantity.

Ordinary cokriging was chosen as a method, since this makes it possible to use a number of additional variables in addition to the main predictive category (the 'Agricultural' category), and so make better use of correlations within the data set. The general formula for ordinary cokriging to estimate an unknown value Z in location x_0 is

$$Zi_0(x_0) = \sum_{i=1}^{N} \sum_{\alpha=1}^{n_i} w_\alpha^i Zi(x\alpha)$$

where:

- i_0 refers to a particular variable of the set of N variables used for the prediction.
- n_i is the number of samples taken of the additional measured values used for the predictions (Wackernagel 2003).

The model

The data were coded as dummy variables in a spread sheet before modelling. The main kriging category and the cokriging categories were coded '1', and the rest of the data were coded '0'. In other words, the data with the exception of the kriging and cokriging categories were used as negative evidence for agricultural land use.

To find the best model and the best predictors, three different models were compared. The main kriging category, the 'agricultural' category, was the same for all three models but the models used different cokriging categories:

- Model 1 used the categories 'settlement', 'animal husbandry' and 'slope' as predictors.
- Model 2 used the categories 'maritime', 'wells' and 'mills' as predictors.
- Model 3 used the categories 'tin mining related', 'manor houses' and 'chapels' as predictors.

4. The mathematical model: PageRank based

PageRank was developed for attributing a value of importance to web pages regardless of the value of their content and solely on the basis of the interconnections between the pages. The criterion used to compute the ranking of the web pages may be summarised as follows: *a page i that points to pages ji,...jk, distributes its importance* *in equal parts to pages ji,...jk, and therefore gives 1/k of its importance to the pages it points to.*

A general introduction to the PageRank model can be found in (Langville *et al.,* 2006), while the application of PageRank based techniques to the estimation of the archaeological potential can be found in (Dubbini and Gattiglia, 2013).

In order to apply a PageRank model to the determination of agricultural land use, we first create a rectangular space fully containing the study area. The size of this 'box' is 40250 m north to south and 52250 m east to west. This is then divided into 161 x 209 cells, each of which 250 m x 250 m in size. The model works based on a matrix of weights S, representing the weights (value) of the links between cells. The matrix S is computed in the following way:

- each cell which has recorded archaeological remains distributes its importance to a square mask of 30 x 30 cells, centred on the cell itself;
- the weights relative to the mask above are added to the matrix of weights S if the cell contains data belonging to the following set of categories: 'Settlements', 'Agricultural', 'Animal husbandry', 'Manor houses', 'Wells', 'Mills'.
- the weights relative to the mask above are subtracted from the matrix of weights S if the cell contains data fulfilling the following criteria: cells containing 'Maritime', 'Tin mining', or 'Deer parks' categories data, or where slope > 10°, or distance to coastline < 100 m.

In other words:

- the more the nearby presence of the first set of categories, the more probable agricultural land use is assumed to be;
- the more the nearby presence of the second set of categories, the less probable agricultural land use is assumed to be;
- the two conditions are in competition;

The PageRank model is thereafter applied. The algorithm consists of these steps:

a) The matrix of weights S is generated as described above.

b) The following iterations are performed

for i = 1,…,1000

$$A = S \cdot x + \begin{bmatrix} 1/n & \cdots & 1/n \\ \vdots & \ddots & \vdots \\ 1/n & \cdots & 1/n \end{bmatrix} \cdot x;$$

A= A/2 + x/2;

u = [1 1 … 1];

y = A; $y = y/(\sum^{n}_{(i=1)} y_i)$

x = y;

end

D = x;

In these formulas x is a stochastic (i.e. the sum of its component equals 1) random column vector of dimension n, used as an initial condition for the application of the iteration described in the 'for' cycle. The result of these iterations are independent of the initial condition.

5. Comparison

The performance was measured using *Kvamme's gain*.

Kvamme's gain statistic estimates performance of archaeological predictive models by comparing the accuracy and the precision of the model to what can be expected by chance. The accuracy is defined as the ability of the model to accurately predict probable locations of archaeological sites, whereas the precision is the ability of the model to pinpoint precise areas of high probability for finding sites. Technically, accuracy and precision are defined in the following way:

Precision = P_a = proportion of total area designated as high probability = A_{high}/A_{total}

Accuracy = P_s = proportion of sites in high probability areas = S_{high}/S_{total}

Kvamme's gain can then be calculated as: $G = 1-P_a/P_s$ (Kvamme 1988).

Cokriging model

The second cokriging model is clearly is the best performing one. Not only is Kvamme's gain the highest at 67% for the second model, but it also performs much better than the other two models at high predicted values.

Predicted values	Model 1	Model 2	Model 3
>0.1	0,2016	0,1959	0,1902
>0.2	0,3426	0,3403	0,3267
>0.3	0,4516	0,4761	0,4513
>0.4	0,5705	0,5734	0,5709
>0.5	0,6317	0,6173	0,602
>0.6	0,5953	0,6047	0,2536
>0.7	0,5449	0,6356	0,5465
>0.8	0,3558	0,6706	0,3705

TABLE 1: COKRIGING MODEL PERFORMANCE – KVAMME'S GAIN

PageRank model

As can be seen in tables 1 and 2, the PageRank performed much better, at 75%, than the best cokriging model did at 67%.

Predicted values	Gain
>0.1	0,1212
>0.2	0,1834
>0.3	0,2067
>0.4	0,2079
>0.5	0,3939

Predicted values	Gain
>0.6	0,5524
>0.7	0,6447
>0.8	0,6176
>0.9	0,7506

Table 2. PageRank model performance - Kvamme's gain

Performance comparison

Since the models predict a range of values, it would be preferable to make a number of comparisons across these values. All models all predict values between 0 and 1, and it may therefore seem logical to compare the models at points where the predicted values are the same. However, the predicted values are not directly comparable. The predicted values have no direct interpretation but are only meaningful in relation to how well they divide the study area into areas of low- and high-probabilities. To compare like-for-like, the best way may be to compare the models may therefore be the following:

If the models are compared at points where the *precision* of the models are identical, Kvamme's gain will be higher for the better model. In other words, Kvamme's gain statistic here measures which of the models are the most *accurate* at points where the precision is identical (Kvammes gain

$$= 1 - \frac{Pa}{Ps} = 1 - \frac{\text{precision}}{\text{accuracy}}$$

Compared in this fashion, it is clear that the PageRank model performs better than the Cokriging model at both low and high values of precision, but the cokriging model is much better in the mid-range values.

High-probability area / Total area	PageRank	Cokriging model 2
90%	0,0983	N/A
80%	0,1763	0,133
70%	0,2139	0,200
60%	0,2161	0,273
50%	0,2529	0,307
40%	0,3291	0,358
30%	0,3833	0,435
20%	0,4666	0,532
10%	0,6092	0,615
5%	0,6537	0,634
1%	0,7403	0,600

Table 3. Performance comparison

Graphic comparison

The cokriging model predicts more localised concentrations of agricultural fields (figure 3), whereas the PageRank model predictions are more spread out (figure 2). It is difficult to imagine that large areas of the study area would have been practically devoid of agricultural fields, as the cokriging model seems to imply. However, the exact pattern of field agriculture in the medieval period is not known for the study area, and it is in general therefore not easy to be certain of which one presents the most realistic picture.

A small part of the study area – the Lizard (the southeast part of the study area) does however present an intriguing picture. The cokriging model clearly predicts that the areas of moorland visible on figures 5 and 7 (darker areas) have a high probability for agricultural use (i.e. were likely under the plough, or more accurately under the spade). Surprisingly, the model predicts the physical extent of the

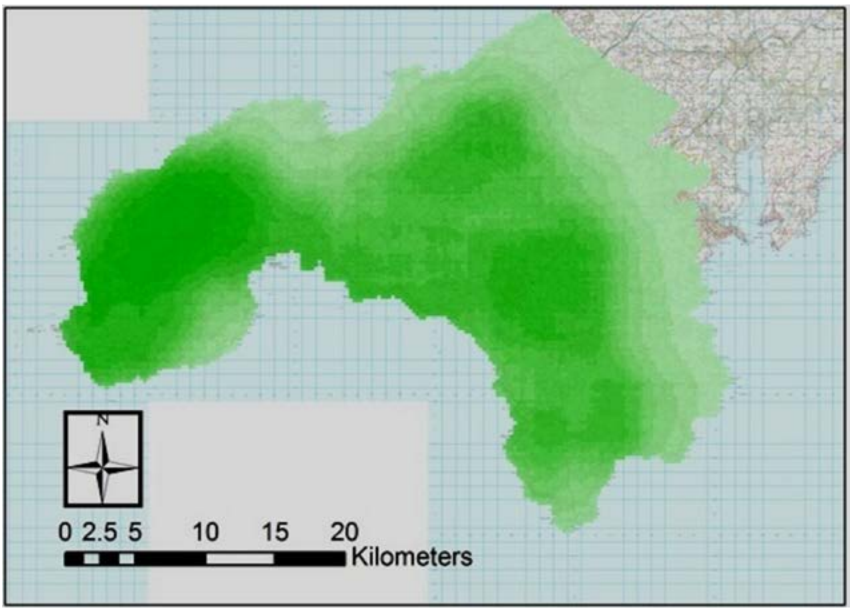


Figure 3: PageRank model. Source: Author. © Crown Copyright/database right 2014. An Ordnance Survey/EDINA supplied service.

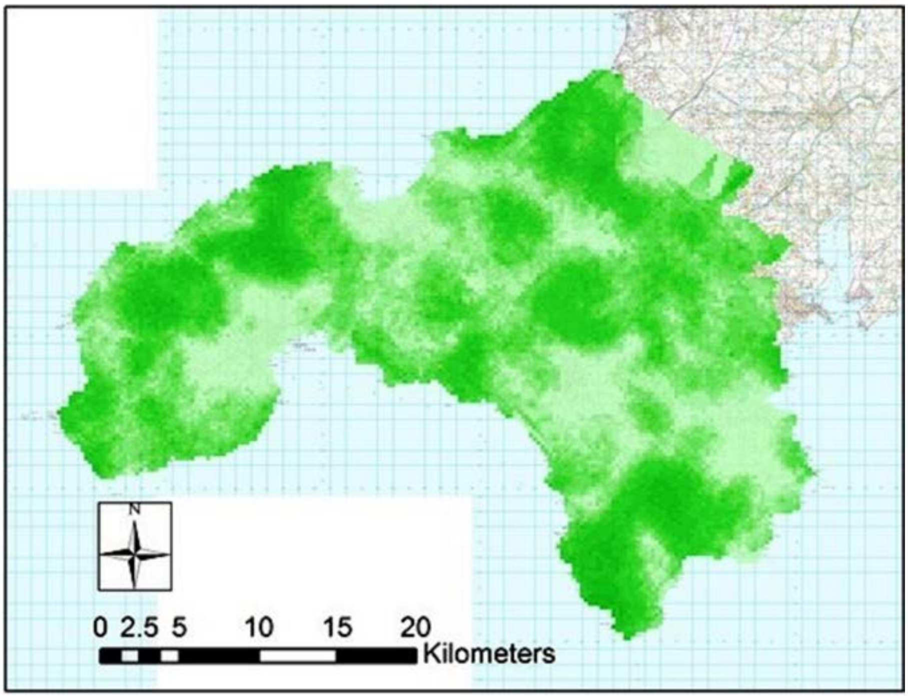


Figure 4: Cokriging model 2. Source: Author. © Crown copyright/database right 2014. An Ordnance Survey/ EDINA supplied service.

Figure 5: Cokriging model - The Lizard. Source: Google Earth. ©2014 Google. Data SIO, US Navy, NGA, GEBCO. Image Landsat. ©Infoterra Ltd and Landsat.

Figure 6: Aerial photo - The Lizard. Source: Google Earth. ©2014 Google. Data SIO, US Navy, NGA, GEBCO. Image Landsat. ©Infoterra Ltd and Landsat.

Figure 7: PageRank model - the Lizard. Source: Google Earth. ©2014 Google. Data SIO, US Navy, NGA, GEBCO. Image Landsat. ©Infoterra Ltd and Landsat.

Figure 8: Aerial photo - The Lizard. Source: Google Earth. ©2014 Google. Data SIO, US Navy, NGA, GEBCO. Image Landsat. ©Infoterra Ltd and Landsat.

moors quite precisely. The moors quite obviously are poor agricultural land. They are not currently used for agriculture. It is likely that the model predicts areas where agricultural remains have survived, not necessarily medieval land use in general. A possible interpretation of the model could be that the moors were used for agriculture up to the mid-14th century (the bubonic plague), but abandoned shortly afterwards for better land. This would have preserved the evidence of agricultural fields (earthworks etc.) better than in the surrounding areas.

By contrast, the PageRank model does not clearly predict that moorland visible on the aerial photographs on the right was used for agriculture in the medieval period. It also does not define the extent of the moorland very well (figures 7 and 8).

The differences between the results can be explained through the differences in how the models reach their predictions.

Kriging makes use of the spatial autocorrelation inherent to the dataset. For the kriging model, this implies that the closer one is to recorded agricultural archaeological remains, the higher the probability is of agricultural land use will be. Naturally, in areas where extant archaeological remains of agricultural land use are plentiful (e.g. in the Lizard moorlands), the model predict high probability for agricultural land use. The only real surprise is how precise the physical extent of the moorland is predicted.

By contrast, the PageRank model does not rely on spatial autocorrelation. The results, and the quality of the predictions, are instead influenced by the assumptions that the researcher makes on how each cell influences the neighbourhood. This means that the predictions are less sensitive of the precise spatial arrangement of the data points.

6. Summary

The PageRank model clearly performed best of the models tested (gain of 0.75). The PageRank model performs better than the best Cokriging model at both low and high values of precision, but the cokriging model is much better in the mid-range values.

It is not quite clear which of the models present the more realistic graphical picture of landuse. The cokriging model in particular predicts large contiguous areas apparently devoid of agricultural fields (or with very few fields present) unlikely to be correct.

The kriging models are potentially more sensitive to the actual pattern of recorded data.

References

AALTO, J., PIRINEN, P., HEIKKINEN, J., VENÄLÄINEN, A. (2013). Spatial interpolation of monthly climate data for Finland: comparing the performance of kriging and generalized additive models. *Theoretical & Applied Climatology*, 112(1/2), pp.99–111. Available at: http://search.ebscohost.com/login.aspx?direct=true&db=a9h&AN=86196957&site=eds-live&scope=site. [Accessed July 22, 2014].

BALLABIO, C., FAVA, F. & ROSENMUND, A. (2012). A plant ecology approach to digital soil mapping, improving the prediction of soil organic carbon content in alpine grasslands. *Geoderma*, 187/188, pp.102–116. Available at: http://search.ebscohost.com/login.aspx?direct=true&db=lah&AN=20123210391&site=eds-live&scope=site. [Accessed July 22, 2014].

BENITO-CALVO, A., PÉREZ-GONZÁLEZ, A. & PARÉS, J.M. (2008). Quantitative reconstruction of Late Cenozoic landscapes : a case study in the Sierra de Atapuerca (Burgos, Spain). *Earth Surface Processes and Landforms*, 33(2), pp.196–208.

CORNWALL COUNTY COUNCIL [producer] 2013. Cornwall HER database.

DUBBINI N., GATTIGLIA G. (2013). A PageRank based predictive model for the estimation of the archaeological potential of an urban area, *Digital Heritage International Congress (Digital Heritage) Proceedings*, 2013, 1, pp. 571-578,

EPSTEIN C.L. (2008). *Introduction to the mathematics of medical imaging*. Philadelphia (USA): SIAM.

KVAMME, K.L. (1980). Predictive model of site location in the Glenwood Springs resource area. In: *Class II Cultural Resource Inventory of the Bureau of Land Management's Glenwood Springs Area, Grand Junction District, Colorado (Part II).* Montrose, (CO): Bureau of Land Management.

KVAMME, K. L. (1983). Computer processing techniques for regional modelling of archaeological site locations. *Advances in Computer Archaeology,* 1, pp. 26-52.

JONES, E.E. & DEWITTE, S.N., (2012). Using spatial analysis to estimate depopulation for Native American populations in northeastern North America, AD 1616–1645. *Journal of Anthropological Archaeology,* 31(1), pp.83–92. Available at: http://linkinghub.elsevier.com/retrieve/pii/S0278416511000547 [Accessed July 22, 2014].

KVAMME, K. (1988). Development and testing of quantitative models. In W. J. Judge & L. Sebastian, (eds.) *Quantifying the Present and Predicting the Past: Theory, Method and Application of Archaeological Predictive Modeling.* Denver, Colorado: U.S. Department of Interior, Bureau of Land Management, pp. 325–418.

LANGVILLE A. N., and MEYER, C. D. (2006). *Google's PageRank and Beyond: The Science of Search Engine Rankings*. Princeton: Princeton University Press.

LIU, X. (2011). Towards a Microsimulation Residential Housing Market Model: Real Estate Appraisal and New Housing Development. Available at: http://search.

ebscohost.com/login.aspx?direct=true&db=edsoai&AN=edsoai.755635579&site=eds-live&scope=site. [Accessed July 22, 2014].

LLOYD, C.D. & ATKINSON, P.M. (2004). Archaeology and geostatistics. *Journal of Archaeological Science*, (31), pp.151–165.

MARCHETTI, A, PICCINI, C., FRANCAVIGLIA, R. MABIT, L. (2012). Spatial distribution of soil organic matter using geostatistics: a key indicator to assess soil degradation status in central Italy. *Pedosphere*, 22(2), pp.230–242. Available at: http://search.ebscohost.com/login.aspx?direct=true&db=lah&AN=20123132757&site=eds-live&scope=site. [Accessed July 22, 2014].

VAN DER MEER, F. (2006). Indicator kriging applied to absorption band analysis in hyperspectral imagery: A case study from the Rodalquilar epithermal gold mining area, SE Spain. I*nternational Journal of Applied Earth Observations and Geoinformation*, 8, pp.61–72. Available at: http://search.ebscohost.com/login.aspx?direct=true&db=edselp&AN=S0303243405000590&site=eds-live&scope=site. [Accessed July 22, 2014].

ROBINSON, D.P., LLOYD, C.D. & MCKINLEY, J.M. (2013). Increasing the accuracy of nitrogen dioxide (NO2) pollution mapping using geographically weighted regression (GWR) and geostatistics. *International Journal of Applied Earth Observation and Geoinformation*, 21, pp.374–383. Available at: http://search.ebscohost.com/login.aspx?direct=true&db=lah&AN=20133005342&site=eds-live&scope=site. [Accessed July 22, 2014].

SCHOLTZ, S.C. (1981). Location choice models in Sparta. In: Lafferty III, R.H., Otinger, J. L., Scholtz, S. C., Limp, W.F., Watkins, B., & Jones, R.D. (eds), *Settlement predictions in Sparta: A locational analysis and cultural resource assessment on the uplands of Calhoun County,* Arkansas. Fayetteville (AR): Arkansas Archaeological Survey 14, PP; 207-222

TOBLER, W.R., (1970). A Computer Movie Simulating Urban Growth in the Detroit Region. *Economic Geography*, 46(2), pp.234-240. Available at: http://search.ebscohost.com/login.aspx?direct=true&db=edsjab&AN=edsjab.10.2307.143141&site=eds-live&scope=site. [Accessed July 22, 2014].

VERHAGEN, P., KAMERMANS, H., LEUSEN, M. VAN, DEEBEN, J., HALLEWAS, D.P., ZOETBROOD, P. (2007). First thoughts on the incorporation of cultural variables into predictive modelling. In P. Verhagen, (ed.) *Case studies in archaeological predictive modelling.* Leiden: Amsterdam University Press.

VERMEER, J. A. M., FINKE, P. A., ZWERTVAEGHER, A., GELORINI, V., BATS, M., ANTROP, M., VERNIERS, J. and CROMBÉ, P. (2014). Reconstructing a prehistoric topography using legacy point data in a depositional environment. *Earth Surface Processes and Landforms* 39(5), pp.632–645. Available at: http://doi.wiley.com/10.1002/esp.3472 [Accessed July 22, 2014].

WACKERNAGEL, H. (2003). *Multivariate geostatistics* (3rd ed.), New York ; London: Springer.

WHEATLEY D. and M. GILLINGS. (2002). *Spatial technology and archaeology - the archaeological applications of GIS*. London: Taylor and Francis.

Chapter 8. 3D Archaeology and Virtual Archaeology

Measuring and Describing 3D Texture

Vera MOITINHO DE ALMEIDA, Juan Antonio BARCELÓ

Quantitative Archaeology Lab (LAQU), Dpt. of Prehistory, Universitat Autònoma de Barcelona (UAB), Spain
veramoitinho@gmail.com , juanantonio.barcelo@uab.cat

Abstract

In this paper we define an archaeological texture in terms of a formalized visualization of complex patterns within a surface composed of spatially organized subpatterns, which have a characteristic, somewhat uniform appearance. Unfortunately, there is no universal method of searching for informative visual marks. They can be extracted from any object almost ad infinitum, but one usually fails to formalize the significant criterion for what is intrinsically "textural". The insufficiency and lack of a clear consensus on the traditional methods of visual description – mostly descriptive and qualitative – have invariably led to ambiguous and subjective interpretations of its functions. In this paper we suggest some systematic, formal and standardized ways of describing texture, implementing objective, mathematical, quantitative, and whenever possible automated techniques.

Our approach is based on the use of non-contact close-range 3D scanners and advanced metrology methods and techniques. The real value of 3D digital data comes from the ability to be able to extract meaningful information from it. In this vein, we should extract a number of different texture descriptors from the generated 3D model. Many of these parameters are currently regulated by international standards, which describe terms, materials and methods, and how they can be used to consistently measure and analyze surfaces. The ISO 25178 series (2012) defines more than 40 areal parameters for measuring 3D surface texture. The advantage of applying 3D areal surface metrology methods and parameters is that these are calculated on the entire or sampled surface and not upon averaging estimation calculations derived from 2D profilometric methods and parameters.

This way of describing and measuring texture variation is applied to the study of the Neolithic stelae with horns from the Serra de Mas Bonet (Vilafant, Catalonia, Spain), and in a set of lithic tools that have been associated with them. This investigation is directed towards a better comprehension of the manufacturing procedures used mainly in the production of one of these stelae.

Keywords: 3D Scanning, ISO 25178, Manufacturing Procedures, 3D Texture, Quantitative Data.

1. Introduction

Texture is usually defined as those attributes of an object's surface having either visual or tactile variety, and defining the appearance of the surface (Tuceryan & Jain, 1998; Fleming, 1999; Mirmehdi, Xie & Suri, 2008; Barceló, 2009; Engler & Randle, 2009). Surface textures usually consist of a series of peaks and valleys that have characteristic form size and spacing, and normally of a periodic nature (Blunt & Jiang 2003). Surface texture is a key parameter in archaeological materials, where its study has been central to use-wear research, as well as in the understanding of manufacturing processes. Every manufacturing technique leaves a micro-scale 'fingerprint', or signature, on the surface which is unique to the manufacturing process. It is the surface that interacts with its working environment through some form of mechanical contact (Blunt & Jiang, 2003; Demkin & Izmailov 2010) – i.e., the mechanical surface – that is of interest here (Moitinho de Almeida, 2013).

Our starting point is that the process of measuring and technology used for quantifying texture data should be integrated with the purpose of analysis, that is, functional explanation. It is not just "documenting" how the artefact looks like, but to solve a problem expressed in formal and objective terms. The following case study describes an experiment based on the use of a sub millimetre resolution 3D scanner, in order to expand the analysis of surface texture features and patterns (i.e., the microtopography) of both experimental and archaeological materials, from the commonly and currently used image/pixel analysis method to a three-dimensional geometrical one, with the aim of extending the range of some important parameters for this kind of application (Barceló & Moitinho de Almeida, 2012; Moitinho de Almeida, 2013). 3D scanners can be used to capture information on areal surface layer through a single measurement, with enough data homogeneity and representativeness. A set of ISO 25178-2 surface texture parameters (ISO, 2012) are used to describe each 3D digital surface texture sample, bringing further quantitative methods to ancient manufacturing processes analysis.

The aim of this research is to understand how archaeological objects were manufactured in the past, and how they were used then in terms of features which are assumed to have been the consequence of an action (human, animal, natural) having modified the original appearance of the surface. Because archaeological objects are intrinsically three-dimensional, their surface texture should be able to be both characterized and analyzed in 3D, so as to improve the understanding of these phenomena. Therefore, we need to measure and describe the microtopography of 3D digital surfaces by using advanced metrology methods and techniques.

2. Measuring and describing 3d texture

Texture is usually defined as those attributes of an object's surface having either visual (i.e., colour, brightness, reflectivity, and transparency) or tactile (i.e., roughness, waviness, and lay) variety, and defining the appearance of the surface (Tuceryan & Jain, 1998; Fleming, 1999; Mirmehdi, Xie & Suri, 2008; Barceló, 2009; Engler & Randle, 2009). Surface textures usually consist of a series of peaks and valleys that have characteristic form size and spacing, and normally of a periodic nature (Blunt & Jiang, 2003). Surface texture is a key parameter in archaeological materials, where its study has been central to use-wear research, as well as in the understanding of manufacturing processes. Every manufacturing technique leaves a micro-scale 'fingerprint', or signature, on the surface which is unique to the manufacturing process. It is the surface that interacts with its working environment through some form of mechanical contact (Blunt & Jiang, 2003; Demkin & Izmailov, 2010) – i.e., the mechanical surface – that is of interest here (Moitinho de Almeida 2013).

In the course of this investigation we will focus on the real surface of an object, which is defined as a set of features which physically exist and separate the entire workpiece from the surrounding medium (ISO, 1996), where the texture of the surface – here, its microtopography, as a scale limited complex combination of spatial frequencies – is just one of its key features. In a mechanical sense, the surface texture of a part may affect its function. Hence, the real surface geometry of an object can be so complex and diverse that a small number of parameters cannot provide a full description. Yet, a more accurate description can be obtained if the set of parameters used is adequate or the number of parameters is increased. This is one of the reasons for introducing advanced metrology methods, techniques, and new parameters for assessment of surface texture (Gadelmawla et al., 2002).

Many of these parameters are currently regulated by ISO International Standards (technical committee TC213, working groups WG 15 "GPS (Geometric Product Specification) extraction and filtration technique" and WG 16 "Areal and profile surface texture"), which describe terms, materials and methods, and how they can be used to consistently measure and analyse surfaces. There are several 2D profile and 3D areal surface texture parameters and constituents. These can be characterized either from physical or digital surfaces, from macroscale to nanoscale, using advanced metrology methods and techniques, and by means of 2D profilers or 3D areal contact (e.g., stylus, atomic force microscopy) or non-contact instruments (e.g., optical interference, optical scatting, capacitance, ultrasound, Scanning Probe Microscope - SPM, 3D scanners), which span a wide range and resolution.

Because archaeological objects are intrinsically three-dimensional, their surface texture should be able to be both characterized and analysed in 3D so as to improve the understanding of functional phenomena. However, the correct measurement of the texture properties of a surface is fundamental for an archaeological explanation. The ISO 25178 series (2012) define more than 40 areal parameters for measuring surface texture, which are grouped in:

- *Height*: calculates the statistical distribution of height values along the z axis. Height parameters are dependent on the level of detail of the captured 3D data;
- *Spatial*: calculates the spatial periodicity of the data, specifically its direction. Spatial parameters are primarily dependent on the level of detail of the captured 3D data;
- *Hybrid*: calculates the spatial form of the data, based upon both amplitude and spatial information;
- • *Functional volume*: is calculated on the basis of the surface bearing area ratio curve (Abbott-Firestone curve). Indicated for bearing and fluid retention properties;
- *Features*: significant features are identified by a segmentation of the surface into valleys and peaks (watersheds algorithm and Wolf pruning). Then, parameters are calculated to quantify only the characteristics of the selected features.

The advantage of applying 3D areal surface metrology methods and parameters is that these are calculated on the entire or sampled surface and not upon averaging estimation calculations derived from 2D profilometric methods and parameters (Blunt & Jiang, 2003; Blateyron, 2006; Barceló, 2010; Blanc, Grime & Blateyron, 2011; Deleanu, Georgescu & Suciu, 2012).

Some of these parameters (or their 2D profile equivalent) have been for several years widely used in many industrial sectors – such as aerospace, automotive, cosmetics, electronics, energy, metallurgy, paper, plastics and printing – namely to investigate surface characteristics, use in product and process design, detect local defects, monitor manufacturing processes and tool performance, control surface quality, or predict surface behaviour during operational use (Blateyron, 2006). This quantitative approach is becoming increasingly used in archaeology (Zahouani, Guinet & Mathia, 1993; Astruc, Vargiolu & Zahouani, 2003; Astruc et al., 2011; Carcagni et al., 2005; Evans & Donahue, 2008; d'Errico & Backwell, 2009; Mélard, 2010; Faulks 2011; Stemp & Stemp, 2003; Stemp & Chung, 2011; Stemp, Lerner & Kristant, 2013), in palaeontology, in dental microwear analysis to reconstruct evolutionary animal adaptation, palaeodiets, and paleoenvironments (Ungar et al., 2003; Schulz & Kaiser, 2007; Schulz, Calandra & Kaiser, 2010; Kaiser et al., 2012), as well as for conservation and monitoring cultural heritage objects (Gaspar, Cummings & Hubbard, 2000).

Although there are different ways to assess surface texture (Scott, 1988; Muralikrishnan & Raja, 2008; Li & Flenley, 2011; Qi, Jiang & Scott, 2012), it generally comprises the following steps (Figure 1):

a) *Choice of 3D data system*
It consists of choosing the adequate measurement instrument, knowing that each type has its own strengths and weaknesses, and in accordance to the complexity and other characteristics of the measuring surface, required level of detail, scale of analysis, sampling method, analytical methods and parameters.

b) *Digital data acquisition / Sampling*
It consists of using a 3D scanning workflow (Moitinho de Almeida 2013) to digitize just a sample area or the whole surface and then extracting surface samples. Sampling should be adequate and consistent, in the sense that it represents

the whole surface area of evaluation. This captured data is a digital representation of the real surface in certain condition produced at a certain phase of a measurement.

c) *Digital surface pre-processing*
It may include surface levelling, form removal, data manipulation (e.g., truncation, rotation, inversion, sub-area extraction), and filtering.

d) *Digital surface description, comparison, and analysis*
This step includes using texture parameters to describe quantitatively, compare, and analyse the surface. Characterization techniques can be scale-dependent (i.e. the results depend on the measurement scale) or scale-independent (e.g., using a fractal dimension).

When planning the surface texture measurement strategy it is fundamental to consider measurement error and uncertainty of measurement results issues (Li & Flenley, 2011) (Figure 2).

Depending on the object of study – including the type of surface (layered, opaque, translucent, transparent, reflective) that is being measured, specific measurement scale and desired level of accuracy, scale of interest (e.g., cm, mm, microns) – and the archaeological question behind, one should select the most suitable metrology strategy, system, technique, and parameters. For details on technical procedures see Whitehouse (2002), Varadi et al. (2004), Masad, Al-Rasan & Little (2007), ASME (2010), ISO 25178 series (2012).

Figure 1: 3D Surface texture workflow: from choosing a 3D data acquisition system to describing the 3D digital surface texture.

Among the archaeological artefacts we have studied using this method, we will discuss here the case of a Late Neolithic zoomorphic lithosculpture. We are interested in understanding how this kind of object was manufactured in the past, namely by: differentiating texture patterns and by associating them with possible gestures, carving techniques, and used tools.

3. Case study

The construction of a large railway infrastructure in 2008 led to the discovery of a prehistoric settlement in the Serra del Mas Bonet (Alt Empordà, Catalonia). During the fieldworks numerous negative structures of various functions and types were documented, as well as a broad diachrony ranging from the 5th to the 2nd millennium cal BC. The best represented occupation phase is the late Neolithic (late 4th millennium cal BC), characterized by different negative structures, as well as a good preservation of its deposits. The most interesting finds of such deposits are a set of sculpted elements: 1 menhir and 6 stelae with carved horns on blocks of fine and medium-grained sandstone – an unprecedented element of megalithic art within the Empordà, where the zoomorphic figures were hitherto unknown (Rosillo *et al.*, 2010) (Fig.3).

4. Materials and methods

In this chapter, we describe the 3D digital geometric surface texture of a given archaeological object and experimental surfaces, in order to associate them with possible gestures, carving techniques, and used tools. To this end, we started by doing a controlled manual experimental replication of the archaeological surface textures. Next, 3D scanning both archaeological and experimental surfaces. Then, selecting a set of digital samples, representative of the surface´s textures. After that, characterizing and differentiating texture

Figure 2: Example of sources of measurement error in surface texture that lead to uncertainty of measurement results.

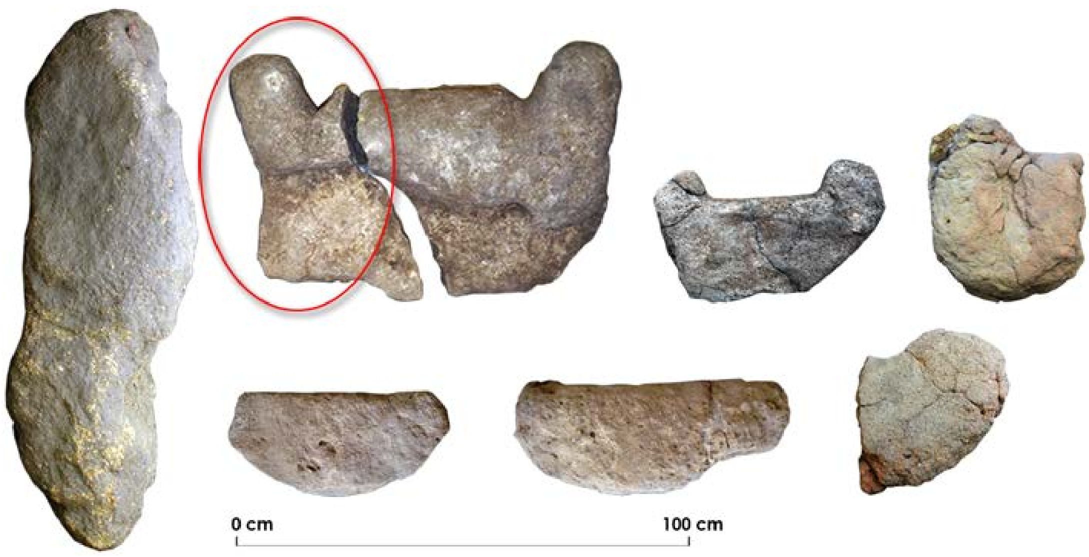


Figure 3: Menhir SMB/08 E-17/8/943 (left), fragments A and B of stela SMB/08 E-17/5/958 (top left), stela SMB/08 E-52/1/22 (top centre), fragment SMB/08 E-1/3/956 (top right), fragment SMB/08 E-48/1/45 (bottom left), fragment SMB-08 E-185/2/5 (bottom centre), fragment SMB/08 E-1/3/957 (bottom right) (Rosillo, Palomo & Moitinho de Almeida, 2013).

patterns. Finally, quantitatively analysing and comparing the archaeological texture features and patterns with the experimental ones.

Experimental Replication of Archaeological Surfaces

The stelae display several technological marks, which have enabled to generate hypotheses about used tools and manufacturing procedures. In addition, a number of lithic tools associated with the stelae have been recovered that could produce evidence to different manufacturing stages (Rosillo, Palomo & Moitinho de Almeida, 2013).

For this study, we selected the best-preserved stela (SMB/08 E-17/5/958, hereinafter referred to as "stela E17", fragment A: 422 mm width x 607 mm height x 221 mm depth, 17974404 mm3, 509046 mm2), as it retains a large amount of macro-traces, and its finding is associated to the products resulting from the configuration of the early form and also to macrolithic tools. After performing a visual inspection on stela E17, selected blocks of sandstone coming from the same bedrock were used to manufacture experimental replicas (Figures 4 & 5) (Moitinho de Almeida, Rosillo & Palomo, 2012, 2013; Palomo & Moitinho de Almeida, 2013) of:

• Six surface textures present in the stela, which were labelled from E_s1 to E_s6;

• Approximately half a stela, which included half body and one protuberance. In this case, the objective was twofold: (i) to continue testing different processes, tools, and gestures; (ii) to perceive the amount of effort a craftsperson could have put into making one of these medium-sized sculptures.

As for the experimental lithic tools, labelled from L1 to L4, these were replicated from limestone and quartz pebbles coming from the river Manol, which is very close to this prehistoric settlement. Even though copper is not present in the archaeological record, the stelae's chronological context allowed this possibility. Therefore, copper punches were also included as tools in this experimental work.

This experimental work was conducted by Antoni Palomo (UAB, Arqueolític) and Rafel Rosillo, and is described in detail in (Moitinho de Almeida, 2013).

In order to describe and compare archaeological and experimental 3D digital surface texture patterns, we proceeded with the method described previously:

4.1. Choice of 3D data system

This step consisted of choosing the adequate measurement instrument. The archaeological stelae had already been 3D scanned for previous research (Moitinho de Almeida, Rosillo & Palomo, 2012, 2013; Rosillo, Palomo & Moitinho de Almeida, 2013), as well as for conservation and monitoring purposes. Therefore, we decided to use the same 3D

FIGURE 4: EXPERIMENTAL REPLICATION OF ARCHAEOLOGICAL SURFACE TEXTURES. FROM LEFT TO RIGHT: ORIGINAL BLOCK OF SANDSTONE, EXPERIMENTAL LITHICS, MANUFACTURE OF SURFACE TEXTURES 1, 2, AND 3 (TOP); ORIGINAL BLOCK OF SANDSTONE, EXPERIMENTAL COPPER PUNCHES, MANUFACTURE OF SURFACE TEXTURE 4 (CENTRE); ORIGINAL BLOCK OF SANDSTONE, EXPERIMENTAL LITHIC, MANUFACTURE OF SURFACE TEXTURES 5, AND 6 (BOTTOM) (ADAPTED FROM ROSILLO, PALOMO & MOITINHO DE ALMEIDA, 2013) (PHOTOS BY RAFEL ROSILLO AND ANTONI PALOMO).

Exp. surface	Exp. tool	Tool type and technique	Direction of movement (approximate angle)	Removed surface area	Produced type of surface
E_s1	lithics L1, L2	chopper used as hammerstone, and hammerstone; pitting	perpendicular (90°)	small and precise; efficient for details (*)	coarse (*)
E_s2	lithics L1, L2	chopper used as hammerstone, and hammerstone; pitting	tangential (20°-45°)	larger and less precise; efficient for general reduction (*)	coarser (*)
E_s3	lithics L1, L2	chopper used as hammerstone, and hammerstone; pitting	perpendicular (90°) + tangential (20°, superficially)	intermediate (with respect to E_s1 and E_s2)	intermediate (with respect to E_s1 and E_s2)
E_s4	copper	punches; pitting	perpendicular (90°) + tangential (20°, dragged movement)	smaller, more precise and quick; more efficient for details (*)	coarsest (*)
E_s5	lithic L2	hammerstone; incision through pitting	perpendicular (90°)	small and precise (**)	coarser (similar to E_s1) (**)
E_s6	lithic L2	hammerstone; incision through pitting and abrasion	perpendicular (90°) + longitudinal abrasion in both directions	small and precise + small (**)	coarse with longitudinal smoothness (groove) (**)

FIGURE 5: QUALITATIVE DESCRIPTION OF THE EXPERIMENTAL REPLICATION OF ARCHAEOLOGICAL SURFACE TEXTURES: IDENTIFICATION OF EXPERIMENTAL SURFACES AND TOOLS, TYPE OF TOOL USED AND APPLIED TECHNIQUE; VISUAL INSPECTION OF APPROXIMATE MOVEMENT DIRECTION OF BOTH HAND AND TOOL TOWARDS THE EXPERIMENTAL SURFACE, AMOUNT OF REMOVED SURFACE AREA, AND TYPE OF PRODUCED SURFACE. COMPARISONS MADE BETWEEN SURFACES: (*) E_S1, E_S2, E_S3, AND E_S4; (**) E_S5 AND E_S6.

scanner (a non-contact close-range 3D structured light scanner: SmartSCAN3D Duo System, Breukmann), field-of-view (FOV: 450 mm stereo; resolution, according to the manufacturer: 280 µm), and overall 3D data capturing and processing procedures on the experimental surfaces, so as to make comparisons more consistent (for further details the reader is referred to Moitinho de Almeida, 2013).

4.2. Digital data acquisition / Sampling

It consisted in making use of the 3D digital models of the archaeological objects previously scanned, and further scanning the experimental surfaces with the same procedures. Next, in selecting a set of texture samples from both archaeological and experimental surfaces (Figure 6). The samples were either contiguous or separated (Blateyron, 2013), each measuring 50x50 mm.

Initially, 10 samples were selected from the experimental surfaces (labelled E_s1a, s1b, s1c, s2a, s3a, s4a, s4b, s5a, s6a, s6b), 6 samples from the experimental stela (labelled from Est_s01 to Est_s06), and 27 samples from the archaeological stela (labelled from Ast_s01 to Ast_s27). The squares in Figure 6 indicate the areas sampled. The small extension of the experimental surfaces did not permit to obtain a larger number of samples per region, more representatives of the applied gestures and techniques.

Samples E_s5a, E_s6a, E_s6b, Est_s2, Est_s5, Ast_s12, Ast_s26, and Ast_s27, were excluded from the present analysis, because they visibly contained two different types of textures. Hence, smaller samples will be required, at least on these regions.

4.3. Digital surface pre-processing

It consisted of 3D digital surface levelling and form removal. Before proceeding with the surface's waviness, roughness, and microroughness measurements, it was necessary to 1. level the surfaces by removing the slope using the least squares plane method and 2. remove the surface's general form by polynomial approximation (Figure 7). We have explored the use of Mountainsmap 7 software trial version (Digital Surf) not only to pre-process the samples, but also to perform the 3D areal surface texture measurements.

4.4. Digital surface description, comparison, and analysis

It consisted in quantitatively describing and analysing archaeological 3D digital areal surface texture patterns and features and compare with the experimental ones. The natural scale of the data was used, and the origin of the z axis (depth) was set at the highest value of the sampled data.

The following set of ISO 25178-2 parameters (ISO 2012) were used as metric variables to describe each sample: Sa (average roughness, mm), Sk (core roughness depth, mm), Smr (areal

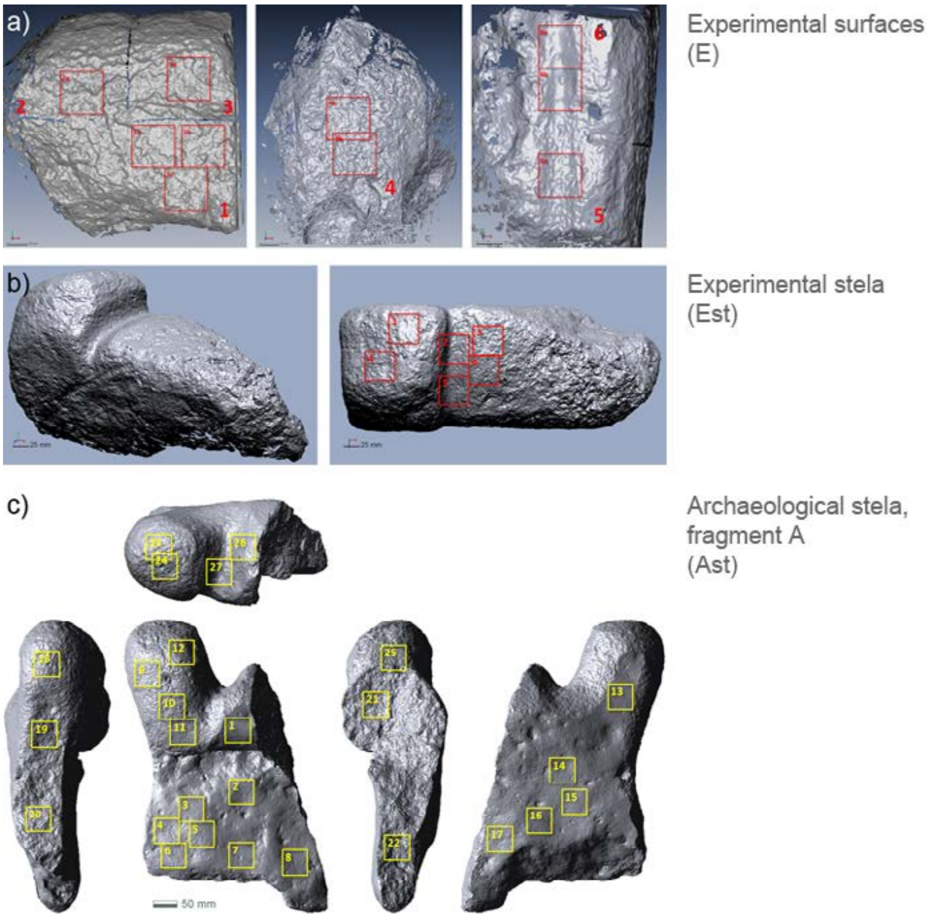


FIGURE 6: 3D DIGITAL SURFACE MODELS OF : (A) 6 EXPERIMENTAL SURFACE TEXTURES (LABELLED E_S1A, S1B, S1C, S2A, S3A, S4A, S4B, S5A, S6A, S6B); (B) EXPERIMENTAL STELA, HALFWAY SCULPTED (FROM EST_S01 TO EST_S06); (C) FRAGMENT A OF STELA SMB/08 E-17/5/958 (FROM AST_S01 TO AST_S27), FIVE ORTHOGRAPHIC VIEWS (TOP, LEFT, FRONT, RIGHT, BACK) IN ACCORDANCE TO FIRST ANGLE PROJECTION, MEANING THAT THE OBJECT IS LOCATED IN QUADRANT I. SQUARES INDICATE THE AREAS SAMPLED (50X50 MM).

material ratio, %), Spd (density of peaks, 1/mm2), Vmc (core material volume, mm3/mm2), Vvc (core void volume, mm3/mm2). Additionally, the texture's isotropy (%), as well as the first, second, and third main directions (º) were taken into account. An isotropic surface has values near 100 % and presents identical characteristics regardless of the direction of the measurement. Conversely, an anisotropic surface has values near 0 % and presents either oriented or periodic structure. The calculation of the isotropy value is similar to the ISO 25178 Str (texture aspect ratio) parameter, with the difference that in the first form is removed automatically before the autocorrelation function is applied (Blateyron, 2006; MountainsMap, 2013).

Variability and similarities between sampled textures and their constitutive features were statistically analysed in JMP® 10 software (SAS Institute). The results of the Sa, Sk (Gaussian filter: 0,8 mm), Smr (c=0,001 mm under the highest peak), Spd (pruning=5%), Vmc (p=10%, q=80%), Vvc (p=10%, q=80%) parameters, plus the isotropy (0,2 threshold) and the three main directions of the sampled surfaces textures (Figure 8), enabled the distinguishing of three major groups of all sampled surface areas (Figure 9):

1) Fine – only natural polished (in the present case, due to natural erosion) archaeological samples (Ast_s01-Ast_s08, Ast_s13-Ast_s17). No experimental surfaces were associated with this, which makes sense, as we did not introduce in this analysis any polished or other fine experimental surfaces. However, it was possible to record inter-individual variations among them. Even though all samples from this group showed a wide range of texture directions (Figure 10, left), samples Ast_s01 to Ast_s08, located on the front surface of the stela, indicated a predominance of oblique directions, and low isotropy; whereas samples Ast_s13 to Ast_s17, located on the back surface of the stela, indicated a predominance of perpendicular and near to perpendicular directions, and medium isotropy (Figure 10, right). These differences between front and back surfaces may suggest evidence on depositional (e.g. pressure of the sandy soil towards the stela) or post depositional processes affecting the natural eroded texture (e.g. wide range of directions, with a preponderance of oblique directions).

2) Coarse and regular – most of the remaining manufactured archaeological samples were here directly associated with the texture signatures of the experimental stela (Est_s03, Est_s06), thus suggesting that they may share similar manufacturing strategies. This pattern is characteristic of manufactured surfaces, modified by impacts made by chopper used as hammerstone and hammerstone, whose working direction can be determined as perpendicular and tangential to the surface (Figures 4 & 5). Even though experimental surface 4 (copper tools used) was included in this group, at this stage we believe this may be indicating that these sampled textures are just coarse and regular, and not that they are associated with the manufacturing of the archaeological stela (besides the resulting traces being much deeper than the archaeological ones, the efficiency of the experimental copper tools was lower when compared with the chopper's edges).

3) Coarse and irregular - this group included samples from fragmented surfaces without evidence of human modification (Ast_s21, s22) or just slightly manufactured (Ast_s19, s20, s23-25). The latter were to some extent associated with Est_s01 and Est_s04.

Even though the K Means NCluster analysis clearly distinguished sub-groups of archaeological texture patterns

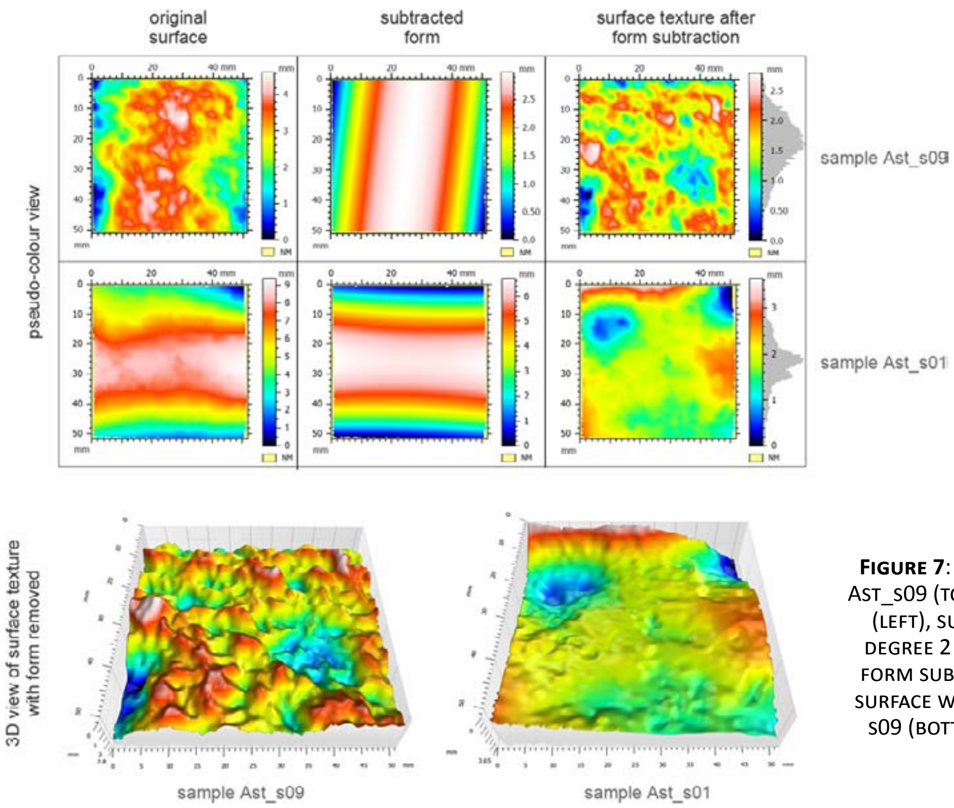


Figure 7: Pseudo-colour view of samples Ast_s09 (top) and Ast_s01 (centre): surface (left), subtracted form, polynomial of degree 2 (middle), surface texture after form subtraction (right). 3D view of the surface with form removed: samples Ast_s09 (bottom left), and Ast_s01 (bottom right).

Sample	*Sa*	*Sk*	*Smr*	*Spd*	*Vmc*	*Vvc*	Isotropy	1st	2nd	3rd
Ast_s01	0,354	0,0042	0,0069	0	0,3605	0,5647	24,3	153,5	0,2	26,5
Ast_s02	0,4131	0,0046	0,0181	0,0004	0,4648	0,5949	21	0,2	17,1	39,2
Ast_s03	0,1705	0,0034	0,0244	0,0015	0,1838	0,2505	22,6	135,3	63,2	153,5
Ast_s04	0,2455	0,0025	0,0076	0,0019	0,283	0,3259	24,1	116,8	0,2	153,5
Ast_s05	0,1183	0,0039	0,0092	0,0049	0,1206	0,1671	41,3	153,5	0,2	26,5
Ast_s06	0,161	0,0028	0,0155	0,0049	0,1744	0,2378	30,5	153,5	26,5	17
Ast_s07	0,2253	0,0037	0,0073	0,0064	0,2629	0,2886	44	26,8	45,3	0,2
Ast_s08	0,3162	0,0049	0,0085	0,0023	0,3459	0,4145	40,1	63,5	26,8	56,5
Ast_s09	0,3491	0,0112	0,0106	0,0133	0,3991	0,4954	67,2	90,1	116,8	141,4
Ast_s10	0,3031	0,0099	0,0103	0,0111	0,3521	0,4639	54,5	135,1	90	116,5
Ast_s11	0,3568	0,0085	0,0056	0,0059	0,403	0,5528	54,6	0,1	135	45
Ast_s13	0,1267	0,0031	0,0077	0,0061	0,1358	0,1668	48,4	90,1	71,5	26,5
Ast_s14	0,2971	0,003	0,0092	0,0011	0,3527	0,432	21,1	90,1	153,5	63,2
Ast_s15	0,2322	0,003	0,0076	0,0022	0,2796	0,347	80,9	44,3	90	62,7
Ast_s16	0,2771	0,0035	0,0095	0,0011	0,311	0,3923	52,9	45	116,5	63,5
Ast_s17	0,2624	0,0033	0,007	0,0026	0,3129	0,3631	53,2	45,6	26,8	71,8
Ast_s18	0,506	0,0098	0,0054	0,0049	0,6193	0,722	72,2	63,5	128,5	0,2
Ast_s19	0,6206	0,0104	0,0061	0,0038	0,6716	0,9786	65,3	63,5	90,1	116,5
Ast_s20	0,6131	0,0107	0,0053	0,0046	0,6693	0,9722	58,5	63,3	147,8	26,5
Ast_s21	1,0232	0,0129	0,0052	0,0027	1,2391	1,4778	28,2	116,8	90,1	63,3
Ast_s22	0,8583	0,0128	0,0048	0,0038	1,0412	1,1609	46,3	90,1	153,5	44,8
Ast_s23	0,5368	0,0118	0,0053	0,0068	0,6463	0,8459	40,8	0,2	153,3	26,8
Ast_s24	0,7611	0,0183	0,0098	0,0061	0,8681	1,122	45	0,2	63,5	116,5
Ast_s25	0,5634	0,01	0,0073	0,0046	0,682	0,7428	41,7	90	116,5	153,5
E_s1a	0,2661	0,0177	0,0181	0,0072	0,2866	0,4339	61	63,5	123,5	57,8
E_s1b	0,4351	0,0196	0,0177	0,0023	0,4706	0,6317	52,9	90,1	45	141,2
E_s1c	0,305	0,0186	0,0177	0,0057	0,3337	0,4716	72	90	153,5	132,5
E_s2a	0,5457	0,0212	0,0183	0,0011	0,6297	0,8347	24,1	116,5	94,2	8,7
E_s3a	0,2469	0,0151	0,0189	0,0042	0,2665	0,4055	55	171,4	90,1	11,9
E_s4a	0,4179	0,0109	0,0058	0,0065	0,4572	0,6194	75,4	90,1	153,2	146,4
E_s4b	0,4717	0,0109	0,0051	0,0072	0,5199	0,6848	74,3	116,5	0,2	90
Est_s01	0,6175	0,0099	0,0073	0,0038	0,7408	0,9617	85,9	0,2	45	26,8
Est_s03	0,3895	0,0102	0,006	0,0094	0,4365	0,613	32,1	0,1	45	63,5
Est_s04	0,6153	0,0102	0,0046	0,0053	0,6853	0,8418	56,2	0,2	45,1	26,5
Est_s06	0,3697	0,0099	0,0048	0,0061	0,414	0,5341	31,5	157,2	128,5	0,2

FIGURE 8: ISO 25178-2 3D DIGITAL AREAL SURFACE TEXTURE PARAMETERS (SK, SMR, SPD, SQ, VMC, VVC), ISOTROPY,1ST, 2ND AND 3RD DIRECTIONS DATA OF THE ANALYZED ARCHAEOLOGICAL (AST) AND EXPERIMENTAL (E, EST) SAMPLES.

(Figure 9), the small number of experimental surface samples did not enable further conclusions on linking experimental to archaeological 3D digital textures.

In some cases, variability in the surfaces' signatures may not express different types of manufacturing technique (namely, gestures, working position, tools, abrasion and lubrication materials, number of manufacturers). Instead, it may express varying degrees of wear (namely, due to time, applied force, or stela-tool interfacial bonds and real contact area). In addition, some wear traces/patterns may also result from shearing and rupture of rubbing materials inside and around the contact region of the carved surface (Myshkin, Petrokovets & Kovalev, 2005). Consequently, the association of these measurements must be interpreted with some caution, since it is determined by several technical issues, the skills of the craftsperson (Baena & Cuartero, 2009), and depositional and post depositional processes.

4. Conclusions

The approach herein presented suggests the interest of the proposed framework towards the understanding of ancient manufacturing procedures, by bringing quantitative 3D digital methods and techniques to ancient manufacturing processes analysis. The inclusion of physical experimental work in this study has permitted the raising and testing of interesting hypothesis about the traces of ancient labour procedures that can be detected with the appropriate analysis of surface 3D micro-variation. Results have enabled us to characterize morphologically the surfaces considered.

The ability of 3D scanners – as the type of one used here – to document geometric surfaces is well known. Although the 450 mm FOV has a lower resolution (280 µm) and LOD when compared with the other set of lens available on this scanner, it was sufficiently capable of acquiring meaningful texture data from sandstone objects. To ensure data consistency and best results, it was of paramount importance to apply the same methods, techniques, and measurement schemes within each stage and step of the framework. 3D areal parameters and specialized metrology software provided a new insight into the quantitative description and analysis of the three-dimensional digital surface textures of stela E17 and experimental surfaces. On these subjects, we understand that:

- The experimental replication of surface textures can be wider.
- If there is no other interest besides studying 3D digital surface textures, samples can be individually 3D scanned. This will, on the one hand, avoid the need to digitize the entire object, taking into account eventual hardware-software and time issues. On the other hand, it will permit scanning of samples more efficiently and with a higher level of detail. As mentioned, in this case the archaeological objects had been earlier fully

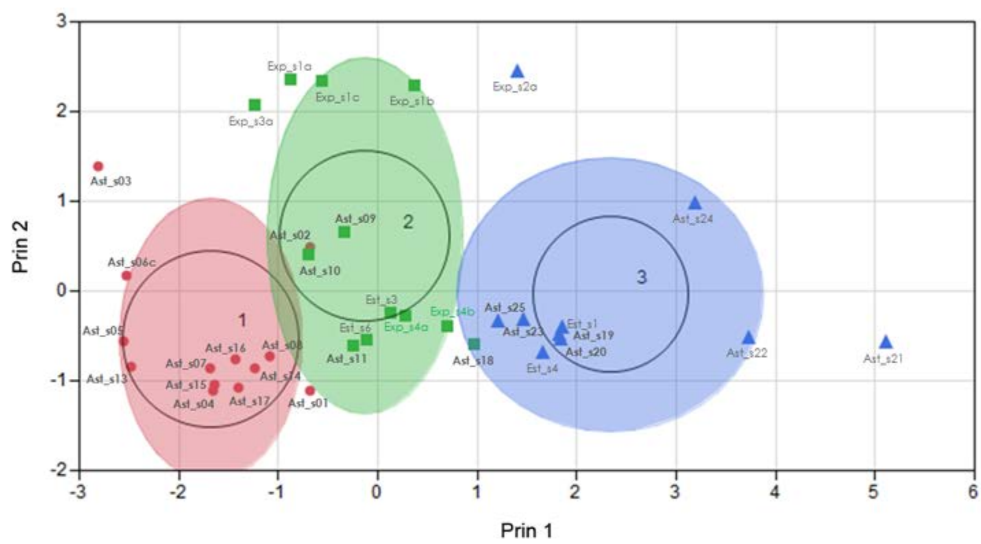


FIGURE 9: RESULTS OF THE K MEANS NCLUSTER=3 ANALYSIS OF THE SURFACE TEXTURE SAMPLES. GROUPS 1) FINE, 2) COARSE AND REGULAR, 3) COARSE AND IRREGULAR.

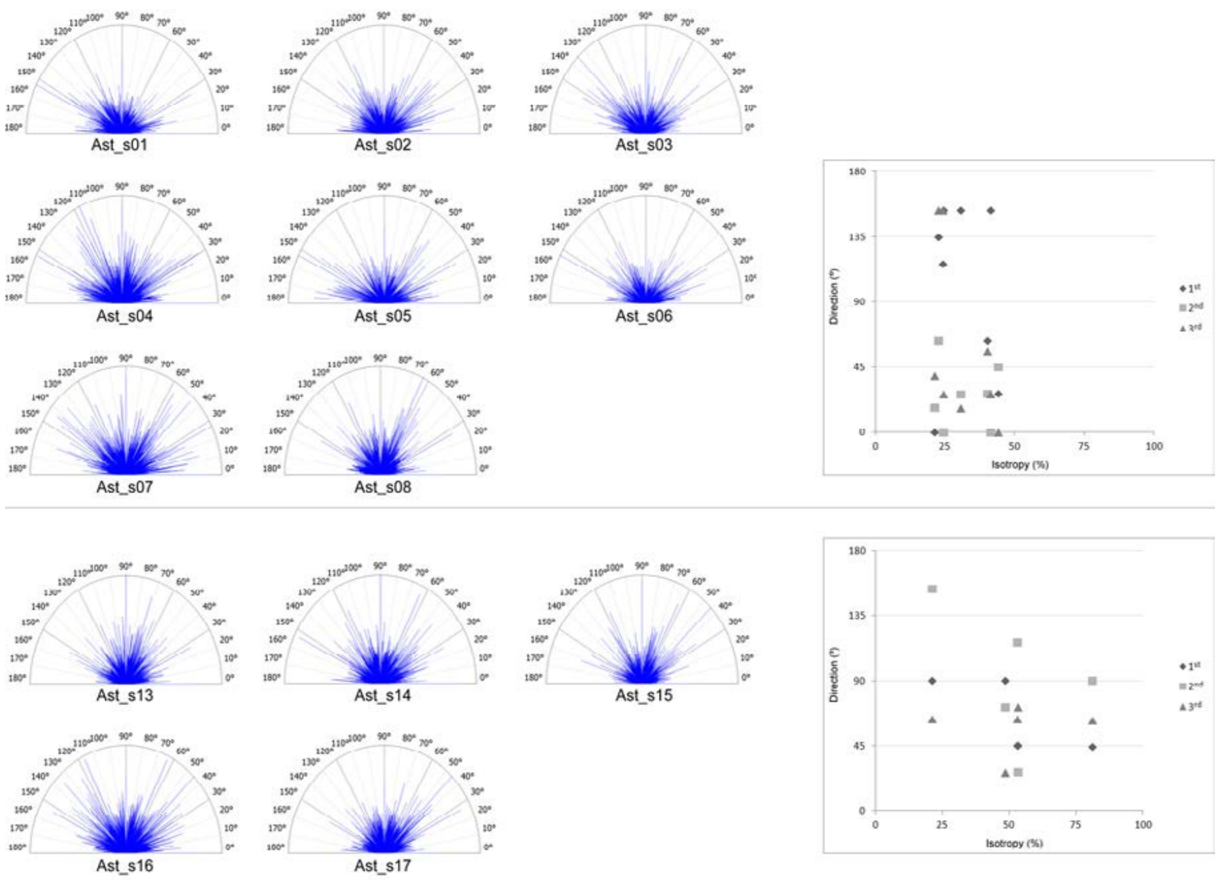


Figure 10: Texture direction and intensity graphics of the archaeological sampled surfaces (left). Isotropy and three main directions scattered plot of the sampled surface textures (right). Samples Ast_s01 to Ast_08 are located on the front view surface of the stela, whereas samples Ast_s13 to Ast_17 are located on the back view surface of the stela. A way of visualizing texture variability among sampled surfaces is through graphics of texture direction and intensity. In this way, we can evaluate whether texture variation within a sampled surface maintain the original isotropy of the working gesture.

scanned, in order to generate 3D digital replicas for multiple uses.

- Surface metrology applications have the potential to greatly contribute to our understanding of ancient manufacturing procedures, among other archaeological issues.
- The power of the ISO 25178-2 surface texture parameters and measurement of texture direction is vast, and can undoubtedly shed light on distinct archaeological issues. On the basis of these analyses, a set of 3D texture parameters which enabled discrimination of the current working surfaces with the best probability were identified: Sa, Sk, Smr, Spd, Vmc, Vvc. Of course, every selection of parameters and variables should always depend on the archaeological question behind and be adequate to describe the texture features of the objects of study.
- The analysis of the data should be done with caution as its interpretation depends on many factors, which include, but are not limited to, all steps of the method, measurement parameters/variables scheme, and performed type of analysis.

In order to pursue further the investigation towards the understanding of archaeological manufacturing procedures through three-dimensional digital surface texture analysis, future work may include:

- Comparing 3D surface texture data between the carved object and the instruments for carving and sculpting;
- Investigating 3D digital sampling issues, in order to allow a better representativeness of the surface textures. This would comprise 3D scanning the sampling areas with higher resolution and testing smaller sample sizes, so as to analyze to what extent different levels of detail and sample size determine the identification of combined texture signatures.

We are looking for creating a 3D digital reference collection of quantified material surface textures, towards the understanding of mechanisms of texture formation, as well as the identification of wear patterns. We are also investigating the possibilities for applying computer simulation (FEA) to a tribological system, hence extending the potentialities of the 3D digital reference collection.

Acknowledgements

The authors gratefully acknowledge: Antoni Palomo and Rafel Rosillo, for conducting the physical experimental work; the Institució Milá i Fontanals, Consejo Superior de Investigacion Científica (IMF-CSIC), for providing the 3D scanner; Digital Surf, for providing a free trial version of Mountainsmap® software; the Museu d'Arqueologia de Catalunya (MAC) and Arqueolític, for both making available the artefacts to be scanned. Special thanks extended to Gavin Alexander and

Carmen Cuenca-Garcia for revising the English language of this paper. This research was part of the project PADICAT (Patrimoni Digital Arqueològic de Catalunya), funded by the Obra Social la Caixa and the Asociació d'Universitats Catalanes (Programa RecerCaixa, RECER2010-05), as well as of the project "Experimenting and developing advanced computational intelligence techniques for simulating social dynamics and historical evolution", funded by The Spanish Ministry of Science and Innovation in the period 2013-2015 (Grant HAR2012-31036). This research also benefitted from Vera Moitinho's Ph.D. grant from the Fundação para a Ciência e Tecnologia (FCT), Portugal, co-funded by the European Social Fund.

References

ASME (2010). ASME-B46-1-2009 – Surface Texture (Surface Roughness, Waviness, and Lay). *American Society of Mechanical Engineers (ASME)*, New York (NY).

ASTRUC, L., VARGIOLU, R., BEN TKAYA, M., BALKAN-ATLI, N., ÖZBAŞARAN, M. & ZAHOUANI, H. (2011). Multi-scale tribological analysis of the technique of manufacture of an obsidian bracelet from Aşıklı Höyük (Aceramic Neolithic, Central Anatolia). *Journal of Archaeological Science* 38. pp. 3415-3424.

ASTRUC, L., VARGIOLU, R. & ZAHOUANI, H. (2003). Wear Assessments of Prehistoric Instruments. *Wear* 255. pp. 341-347.

BAENA, J. & CUARTERO, F. (2009). Perspectives metodològiques de l'experimentació en arqueología. *Cota Zero* 24. pp. 21–30.

BARCELÓ, J.A. (2009). *Computational Intelligence in Archaeology*. New York (NY): The IGI Group, Hershey.

BARCELÓ, J.A. (2010). Visual Analysis in Archaeology. An Artificial Intelligence Approach. In: Elewa, E.M.T. (ed.). *Morphometrics for Nonmorphometricians*. Springer Verlag, Berlin. Lecture Notes in Earth Sciences 124. pp. 92-156.

BARCELO, J.A. & MOITINHO DE ALMEIDA, V. (2012). Functional Analysis from Visual and Non-Visual Data. An Artificial Intelligence Approach. *Mediterranean Archaeology and Archaeometry* 12(2). pp. 273–321.

BLANC, J., GRIME, D. & BLATEYRON, F. (2011). Surface characterization based upon significant topographic features. *Journal of Physics: Conference Series* 311-012014. pp. 1-6.

BLATEYRON, F. (2006). New 3D Parameters and Filtration Techniques for Surface Metrology. *Quality Magazine*. Available from: http://www.qualitymag.com/ext/resources/files/white_papers/New3DParametersandFiltrationTechniquesforSurfaceMetrology-DigitalSurf.pdf [Accessed: 1 April 2014]

BLATEYRON, F. (2013). The Areal Field Parameters. In: Leach, R. (ed.). *Characterisation of Areal Surface Texture*. Springer. pp.15-43.

BLUNT, L. & JIANG, X. (eds.) (2003). *Advanced Techniques for Assessment Surface Topography*. London: Elsevier.

CARCAGNI, P., DAFFARA, C., FONTANA, R., GAMBINO, M.C., MASTROIANNI, M., MAZZOTA, C., PAMPALONI, E. & PEZZATI, L. (2005). Optical Micro-profilometry for Archaeology. In: Salimbeni, R. & Pezzati, L. (eds.). *Optical Methods for Arts and Archaeology. Proceedings of the SPIE*. 5857. pp. 118-128.

DELEANU, L., GEORGESCU, C. & SUCIU, C. (2012). A Comparison between 2D and 3D Surface Parameters for Evaluating the Quality of Surfaces. *The Annals of "Dunărea de Jos" University of Galati*. [Online] 5. p. 5-12. Technologies in Machine Building. Available from: http://freepdfdb.com/pdf/a-comparison-between-2d-and-3d-surface-parameters-for-57799261.html [Accessed: 1 April 2014]

DEMKIN, N.B. & IZMAILOV, V.V. (2010). The Relation between the Friction Contact Performance and the Microgeometry of Contacting Surfaces. *Journal of Friction and Wear*. 31(1). p. 48-55.

ENGLER, O. & RANDLE, V. (2009). *Introduction to Texture Analysis: Macrotexture, Microtexture, and Orientation Mapping*. Boca Ratón (FL): CRC Press.

D'ERRICO, F. & BACKWELL, L. (2009). Assessing the Function of Early Hominin Bone Tools. *Journal of Archaeological Science*. 36(8). p. 1764-1773.

EVANS, A.A. & DONAHUE, R.E. (2008). Laser Scanning Confocal Microscopy: a potential technique for the study of lithic microwear. *Journal of Archaeological Science* 35. pp. 2223-2230.

FAULKS, N.R., KIMBALL, L.R., HIDJRATI, N. & COFFEY, T.S. (2011). Atomic Force Microscopy of Microwear Traces on Mousterian Tools from Myshtylagty Lagat (Weasel Cave), Russia. *Scanning*. 33. pp. 304-315.

FLEMING, B. (1999). *3D Modeling and Surfacing*. San Francisco (CA): Morgan Kaufmann Publishers.

GADELMAWLA, E.S., KOURA, M.M., MAKSOUD, T.M.A., ELEWA, I.M. & SOLIMAN, H.H. (2002). Roughness Parameters. *Journal of Materials Processing Technology* 123. pp. 133-145.

GASPAR, P., CUMMINGS, A. & HUBBARD, C. (2000). Topographical Studies in the Conservation of Statuary Materials. *Conservation Journal.* [Online] 36. Available from: http://www.vam.ac.uk/content/journals/conservation-journal/issue-36/topographical-studies-in-the-conservation-of-statuary-materials/ [Accessed: 1 April 2014]

INTERNATIONAL ORGANIZATION FOR STANDARDIZATION (1996). *ISO 14460-1:1996 - Geometrical Product Specifications (GPS) - Geometric Features - Part 1: General terms and definitions*. Geneva: ISO.

INTERNATIONAL ORGANIZATION FOR STANDARDIZATION (2012). *ISO 25178-2:2012 - Geometrical Product Specifications (GPS) - Surface texture: Areal - Part 2: Terms, definitions and surface texture parameters*. Geneva: ISO.

KAISER, T. M., MÜLLER, D. W. H, FORTELIUS, M., SCHULZ, E., CODRON, D. & CLAUSS, M. (2012). Hypsodonty and tooth facet development in relation to diet and habitat in herbivorous ungulates: implications for understanding tooth wear. *Mammal Review* 43. pp. 34-46.

LI, P. & FLENLEY, J.R. (1999). Pollen Texture Identification using Neural Networks. *Grana*. 38. pp. 59-64.

MASAD, E., AL-ROUSAN, T. & LITTLE, D. (2007). *Test Methods for Characterizing Aggregate Shape, Texture, and Angularity*. NCHRP Report 555, National Cooperative Highway Research Program. Washington D.C. (WA): Transportation Research Board of the National Academies.

MELARD, N. (2010). L'étude microtopographique et la visualisation 3D dans l'analyse de gravures préhistoriques – L'exemple des pierres gravées de La Marche. *In Situ, Revue des Patrimoines*. [Online] 13. Available from : http://insitu.revues.org/6837 [Accessed: 1 April 2014]

MIRMEHDI, M., XIE, X. & SURI, J. (eds.) (2008). *Handbook of Texture Analysis*. London: Imperial College Press.

MOITINHO DE ALMEIDA, V. (2013). *Towards functional analysis of archaeological objects through Reverse Engineering processes*. PhD thesis. Barcelona: Dpt. of Prehistory, Universitat Autònoma de Barcelona (UAB).

MOITINHO DE ALMEIDA, V., BARCELÓ, J.A., ROSILLO, R. & PALOMO, A. (2013). Linking 3D Digital Surface Texture Data with Ancient Manufacturing Procedures. In: A.C. Addison, G. Guidi, L. De Luca, S. Pescarin (eds.). *IEEE Proceedings of the Digital Heritage International Congress, Marseille*. Volume 1. pp.735-738.

MOITINHO DE ALMEIDA, V., ROSILLO, R. & PALOMO, A. (2012). 3D Macrowear analysis of sculpture techniques in the manufacture of the Neolithic stelae with horns of the Serra del Mas Bonet (Vilafant, Catalonia). *Poster presented at the Computer Applications and Quantitative Methods in Archaeology Congress (CAA)*, Southampton.

MOITINHO DE ALMEIDA, V., ROSILLO, R. & PALOMO, A. (in press). Analysis and production. In: Sèrie monogràfica excavacions arqueològiques a Catalunya. *La Serra del Mas Bonet (Vilafant) i Els Banys de la Mercè (Capmany), dues ocupacions a l'aire lliure emmarcades en la prehistòria recent a l'Alt Empordà*. Barcelona: Museu d'Arqueologia de Catalunya.

MOUNTAINSMAP (2013). *MountainsMap manual*. Besançon: Digital Surf.

MURALIKRISHNAN, B. & RAJA, J. (2008). *Computational Surface and Roundness Metrology*. Springer.

MYSHKIN, N.K., PETROKOVETS, M.I. & KOVALEV, A.V. (2005). Tribology of Polymers: adhesion, friction, wear, and mass-transfer. *Tribology International* 38. pp. 910-921.

QI, Q., JIANG, X. & SCOTT, P.J. (2012). Knowledge Modeling for Specifications and Verification in Areal Surface Texture. *Precision Engineering* 36(2). pp. 322-333.

ROSILLO, R., PALOMO, A. & MOITINHO DE ALMEIDA, V. (2013). Técnicas escultoricas en la fabricación de las estelae con cuernos Neolíticas de la Serra del Mas Bonet (Vilafant, Alt Empordà-Nordeste Peninsular). In: Palomo, A., Piqué, R. & Terradas, X., (eds.). *Experimentación en arqueología. Estudio y difusión del pasado*. Girona: Museu d'Arqueologia de Catalunya, Sèrie monográfica, pp. 141–147.

ROSILLO, R., TARRÚS, J., PALOMO, A., BOSCH, A. & GARCÍA DE CONSUEGRA, R. (2010). Les Esteles amb Banyes de la Serra del Mas Bonet (Vilafant, Alt Empordà) dins de l'Art Megalític de Catalunya. *Cypsela* 18. pp. 43–59.

SCHULZ, E., CALANDRA, I. & KAISER, T.M. (2010). Applying Tribology to Teeth of Hoofed Mammals. *Scanning*. 32. p. 162-182.

SCHULZ, E. & KAISER, T.M. (2007). Feeding Strategy of the Urus 'Bos primigenius' BOJANUS 1827 from the Holocene of Denmark. *Courier Forschungsinstitut Senckenberg* 259. pp. 155-164.

SCOTT, P.J. (1988). Developments in Surface Texture Measurement. *Surface Topography* 1. pp. 153-163.

STEMP, W.J. & CHUNG, S. (2011). Discrimination of Surface Wear on Obsidian Tools Using LSCM and RelA: Pilot Study Results (Area-Scale Analysis of Obsidian Tool Surfaces). *Scanning* 33. pp. 279-293.

STEMP, W.J., LERNER, H.J. & KRISTANT, E.H. (2013). Quantifying Microwear on Experimental Mistassini Quartzite Scrapers: Preliminary Results of Exploratory Research Using LSCM and Scale-Sensitive Fractal Analysis. *Scanning* 35. pp. 28-39.

STEMP, W.J. & STEMP, M. (2003). Documenting stages of polish development on experimental stone tools: surface characterization by fractal geometry using UBM laser profilometry. *Journal of Archaeological Science* 30(3). pp. 287-296.

TUCERYAN, A. & JAIN, A.-K. (1998). Texture Analysis. In: Chen, C.H., Pau, L.F. & Wang, P.S.P. (eds.). T*he Handbook of Pattern Recognition and Computer Vision*. World Scientific Publishing, p.207-248.

UNGAR, P.S., BROWN, C.A., BERGSTROM, T.S. & WALKER, A. (2003). Quantification of Dental Microwear by Tandem Scanning Confocal Microscopy and Scale-sensitive Fractal Analyses. *Scanning*. 25. pp. 185-193.

VÁRADI, K., PALÁSTI-KOVÁCS. B., CZIFRA, Á., NÉDER, Z. & KOVÁCS, K. (2004). 3D Characterization of Engineering Surfaces. Available from: *Budapest Tech Jubilee Conference*. http://conf.uni-obuda.hu/jubilee/varadi.pdf [Accessed: 1 April 2014]

WHITEHOUSE, D.J. (2002). *Handbook of Surface Metrology*. Bristol: Institute of Physics.

ZAHOUANI, H., GUINET, A. & MATHIA, T. (1993). Analysis of 3D-pattern Topography of Manufactured Surfaces in Regard to Tribological Situations. *Proceedings of the 6th International Congress on Tribology*, Budapest. pp.284-289.

Old Versus New – Introducing Image-Based 3D Modeling into the General Documentation Workflow of Archaeological Rescue Excavations. Case Studies: the Čachtice and Bratislava castles, Slovakia

Jan ZACHAR,[1] Seta ŠTUHEC[2]

[1] VIA MAGNA Ltd., Vrútky, Slovakia
[2] University of Ljubljana, Slovenia
jan.zachar@yahoo.com, seta.stuhec@gmail.com

Abstract

This paper presents the possibilities of implementation of image-based modeling procedures combining the Structure from Motion (SfM) approach with dense Multi-View Stereo (MVS) algorithms in a documentation workflow of rescue archaeological excavations. The primal focus is the evaluation of its advantages as well as its drawbacks with respect to the special needs and demands of documentation practice of rescue excavations. Further, the paper tries to pose questions on the relevance of 3D data acquisition and their subsequent exploitation possibilities. At the on-going rescue excavations on the courtyard of the Bratislava castle and the area of Čachtice castle (Slovakia), such an image-based modeling approach was introduced into the existing documentation workflow. To assess the benefits and drawbacks of these relatively new techniques and compare them to the recording strategy used hitherto, the complete field documentation was carried out in a traditional way (i.e. total station measurements, still photography and 2D photogrammetry) as well as by applying an SfM+MVS approach using the software package PhotoScan Professional from Agisoft. Afterwards, both datasets were compared in terms of accuracy and interpretative value. Additionally, the time and financial costs needed to acquire basic data sets and produce the final archaeological documentation, were taken into account. With these comparisons, it is hoped that the benefits and drawbacks of image-based 3D modeling in the context of commercial archaeological rescue excavations will become clearer.

Keywords: Archaeological Documentation, Photogrammetry, Computer Vision, Image based Modeling, Structure from Motion

1. Introduction

Currently, rescue excavations make up a very big part of all archaeological activity, especially in the private sector. Archaeologists working at such excavations face financial limitations on the one hand and the demand for the most complete and efficient execution of the dig on the other. Because of its destructive nature, high-quality archaeological documentation is of utmost importance in archaeological excavations. Given the characteristic time pressure and resulting lack of on-site interpretations during rescue excavations, accurate and exhaustive documentation becomes even more important as post-excavation interpretations will be entirely based upon them. Although technologies such as terrestrial laser scanning enable a fast, accurate and rather complete acquisition of archaeological data for subsequent analyses, the lack of financial means and technical knowledge of rescue archaeologists generally hamper their incorporation into the documentation workflows established over the years (which are generally based on total station measurements combined with photographs and pencil drawings).

However, the recent boom of computer vision programs that combine the Structure from Motion (SfM) approach with dense Multi-View Stereo (MVS) algorithms changed this situation. These packages enable the generation of accurate three-dimensional (3D) models from a collection of photographs, in a straightforward and cost-effective manner, without the need for extensive photogrammetric and computer vision knowledge of the user not information on the geometrical properties of the scene. As such, they became a common part of the documentation practice in many research and commercial fields, including cultural heritage monitoring and preservation.

At the on-going rescue excavations on the courtyard of Bratislava castle (Slovakia), such an image-based modelling approach was introduced into the existing documentation workflow. Within a reconstruction project at Čachtice castle, the complete castle was documented using this method.

2. The traditional documentation approach

Documentation of rescue excavations in particular, has to be quick and efficient. Ideally, an archaeologist should be able to inspect the results of his/her measurements, photographs and drawings on-site, which would help him/her to decide on a further course of action in the excavations. The nature of archaeological excavation methodology that is based on recognizing separate stratigraphic units, demands a sufficient overview of the archaeological situation of the site. Therefore, spatial data and especially the consequently derived plans and maps, are of the upmost importance already during the excavation process. Furthermore, the course of rescue excavations usually have to be adjusted to the new building specifications, as the new constructions are usually built simultaneously on the already excavated part of the site. This means that the archaeological units are often documented part by part, not only because of the nature of gradual excavation methods, but also because of the separation of the site into sectors, which are subordinate to the new construction plans.

The traditional documentation of archaeological sites is based on drawings, photographs, measurements, description forms, lists and diaries. In order to create the previously mentioned plans of the archaeological situation on the site, photographs are taken from above as vertically as possible using a ladder or a pole (if possible also an unmanned aviation vehicle (UAV) can be employed). The photographs are then rectified and georeferenced using ground control points measured by a total station or a GPS device. The archaeological units captured on such photographs are afterwards vectorized and the result made into an orthophotoplan.

When taking a photograph, the analogue signal that is acquired and turned into a digital image is degraded in various ways. It has to be taken into account that geometrical distortions occur in every image because the three-dimensional (3D) properties of the scene are mapped to a two-dimensional (2D) sensor plane. The mapping result (i. e. the final image) is influenced by a wide variety of factors of which the geometric errors induced by the optics, the tilt of the camera axis and the topographical relief contribute most to image deformations. Compensating for these deformations through some kind of geometric correction is essential for the extraction of accurate information by means of vectorization.

Basically, there are two solutions which deal with these geometric deformations; simple rectification and rigorous orthorectification (Verhoeven et al. 2013).

2.1. Simple rectification

For tilted images, the scale will vary with direction: the background of tilted photographs features a smaller scale than the scale of objects in the foreground. The projective transformation of a tilted image to a horizontal plane can remove these tilt displacements (and thus scale differences). Such a procedure is called a (planar) rectification and the result is a rectified photograph. Often, first and second order polynomials are also used to obtain (semi-)rectified photographs, although these algorithms differ to a varying degree of the aforementioned projective transformation.

In the absence of lens distortions and perfect flatness of the imaged scene, the rectified tilted image will be identical to a vertical image and the result will be a true orthophotograph. However, lens distortions are always present and truly flat surfaces are seldom found at archaeological excavations. Since any difference in height will cause topographic (or elevation) displacements, any feature lying above or below the horizontal reference surface will be misplaced in a planar rectification. More advanced algorithms are thus required, if accurate mapping from photographs is needed (Wolf and Dewitt 2000).

2.2. Rigorous orthorectification

When the geometric image correction aims to compensate also for topographically-induced deformations and lens distortions it is called orthorectification or differential rectification. The result of such a correction is a planimetrically correct true orthophotograph. Until a few years ago, true orthophotographs could only be achieved with advanced photogrammetric packages such as the Trimble INPHO Photogrammetric System. Besides cost, these packages were also limited for excavation documentation as photogrammetric skills were essential. However, due to the ever increasing technological improvements in computer hardware and the serious advances made in the past fifteen years in the scientific development of computer vision, cost-effective approaches exist today. Additionally, they allow a straightforward workflow and can still produce metrically accurate true orthophotographs (Verhoeven et al. 2013).

3. 3D image-based documentation

To overcome the above mentioned rectification problem an accurate 3D model of the scene can be used.

Such a 3D model can be created using a 3D laser scanner or a collection of overlapping photographs and a computer vision based software package. 3D laser scanners are much faster in obtaining the data, but image-based modelling is more cost-efficient. For this reason image-based 3D documentation is generally more suitable for rescue excavation. Furthermore, computer vision based software packages that combine Structure from Motion (SfM) and Multi View Stereo (MVS) algorithms allow a very straightforward generation of accurate 3D models. To generate such 3D models, specific computer vision knowledge is not needed, as the process is semi-automatic. Additionally, the data acquisition is relatively quick, as no preliminary camera calibration is needed. Using SfM algorithms, the software calculates internal (focal length, lens distortion coefficients, principal point) and external (camera position and orientation) camera parameters automatically and during this process generates a sparse point cloud. Afterwards the MVS algorithm is used to generate a dense point cloud. The next step is connecting these points, creating a solid mesh onto which texture can afterwards be applied. Some software packages (such as Photoscan Professional) enable also georeferencing of the 3D model (Verhoeven et al. 2012).

The best result can be achieved with distinctively coloured and textured objects. Finally a correct photo capture approach is essential to achieve the best possible results.

3.1. Data acquisition

Because not a lot of metadata is needed, the aforementioned approach enables the generation of 3D content from almost any set of photographs. Nevertheless, to successfully generate a 3D model, one should hold on to the certain rules when taking the photographs.

Till recently there was a rule that every photograph should be taken from another location to prevent plausible misalignments. In a recent publication it was suggested that not only normal pictures (i.e. with the optical axis about 90° to the scene, both vertical and horizontal) should be taken, but also some convergent images should be acquired (more on the so called one-panorama-step in Wenzel et al. 2013, see also Nocerino et al. 2014).

The photographs should be taken with a sufficient overlap. There should be at least a 60 % overlap, although 80 % often works better. An easy rule to hold on to is a 'rule of three', which states that each feature should be visible on at least three photographs. Because only the features that are visible on the photographs can be reconstructed, it is necessary to capture the object properly from all possible positions. In order to reach also the higher structures, pole aerial photography (PAP) or unmanned aerial vehicles (UAV) can be used in addition to the ground based photography.

An important part is also the image quality. Photographs should be sharp and acquired with good quality optics. If possible, it is advised to take the complete set of photographs with the same focal length and focus distance, as it normally gives better results. Image- based modelling does not set any requirements concerning the image resolution. However, the resolution of the input data influences the quality of the processed results. That is why it is recommended to employ a camera with at least 5Mpx resolution. In order to produce professional quality orthophotoplans, it is better to opt for 12Mpx resolution.

In terms of archaeological excavations this documentation technique means a relatively easy-to-use, cost-effective alternative to 3D laser scanning that literally brings a new dimension to the documentation archive. Integration of this approach into the standard documentation workflow is not too demanding from the acquisition point of view; as digital photography and ground control point measurements are already employed, the only thing left to do is to accordingly acquire a larger amount of photographs.

3.2. Data processing

Today there are plenty of SfM+MVS packages available, ranging from free and open source to commercial ones. Some of them (such as ARC3D, 123D Catch) provide an online service that automatically carries out the entire procedure. These online approaches are very easy to use. After the upload of the photos, all the computation is done server side and the final result (3D mesh) is delivered to the user mostly via e-mail or is downloadable on a particular web page. However, the number of images is very limited and additionally, hardly any manipulation of the parameters is possible. Web based service 123D Catch was tested for the documentation of the entrance gate to the northern forecastle of the Čachtice castle. 123D Catch enables the possibility of user interference in the alignment procedure as well as in defining the level of mesh density. The basic disadvantage of most free web-based image based modelling services including 123D Catch is also the absence of the straightforward procedure for orthophotoplan and digital elevation model (DEM) creation, which have to be created in other software (SW) options.

From the commercial SW packages, Photoscan Professional from the Russian provider Agisoft Ltd. has lately become a widely spread image-based modeling tool with a broad range of application fields including cultural heritage and archaeological documentation (De Reu et al. 2013, De Reu et al. 2014, Doneus et al. 2011, Verhoeven et al. 2012a, Verhoeven et al. 2012b). Photoscan Professional offers a user friendly interface and easy- to- learn semi-automatic workflow which does not require deep theoretical knowledge of photogrammetry and computer vision. In that sense it is a useful tool for photogrammetric non-professionals. It is a SfM+MVS based SW that enables the creation of a georeferenced and textured 3D mesh. Furthermore it offers other functionalities such as scaling (in case we do not possess GCPs for georeferencing) and the orthophotoplans and DEM extraction, which can be done in an easy straightforward manner. For the purpose of archaeological documentation the mentioned functionalities are of crucial importance. SW enables manual marker placement on the signalized GCP and subsequently one can assign the measured values of the coordinate reference system to them. Consequently after georeferencing a 3D model, direct extraction of georeferenced orthophotoplan and DEM through a pop up window with various setting possibilities is available.

3.3. Data deliverables

Documentation used for archaeological interpretation and preservation is (still) focused on 2D documentation. Rather than employing 3D models as such, georeferenced 2,5D DEM, orthophotoplans and cross-section plans are used (Fig. 8, 9). These deliverables are generally also of prime importance for the excavation´s report.

Georeferenced plans and maps derived from the 3D model are more accurate than the plans acquired by traditional techniques; rectification normally results in pseudo-orthophotographs, which do not account for all possible relief-related deformations and lens distortions that are induced in the photograph. On the other hand 3D models can be manipulated in a way that derives a bird´s eye view without any further distortions and inaccuracies. Furthermore, georeferenced 3D models allow generation of different types of views, cross-sections, analytical calculations, etc. in the post-excavation period that might help to answer newly arisen questions.

4. Case studies: archaeological excavations of the Čachtice and Bratislava castle

Traditional as well as 3D image-based modeling documentation approaches were tested on rescue archaeological excavations of the Čachtice and Bratislava castles in Slovakia. Both of them yielded remains of still standing architecture which was very suitable for testing the efficiency of 3D image-based modeling procedures for the purposes of archaeological documentation.

4.1. The historical overview

4.1.1 The Čachtice castle

The fact that we are willingly entering the an in Today, the Čachtice castle is a ruin situated on the top of a limestone peak (375 m) within the Čachtice carstic hills in Male

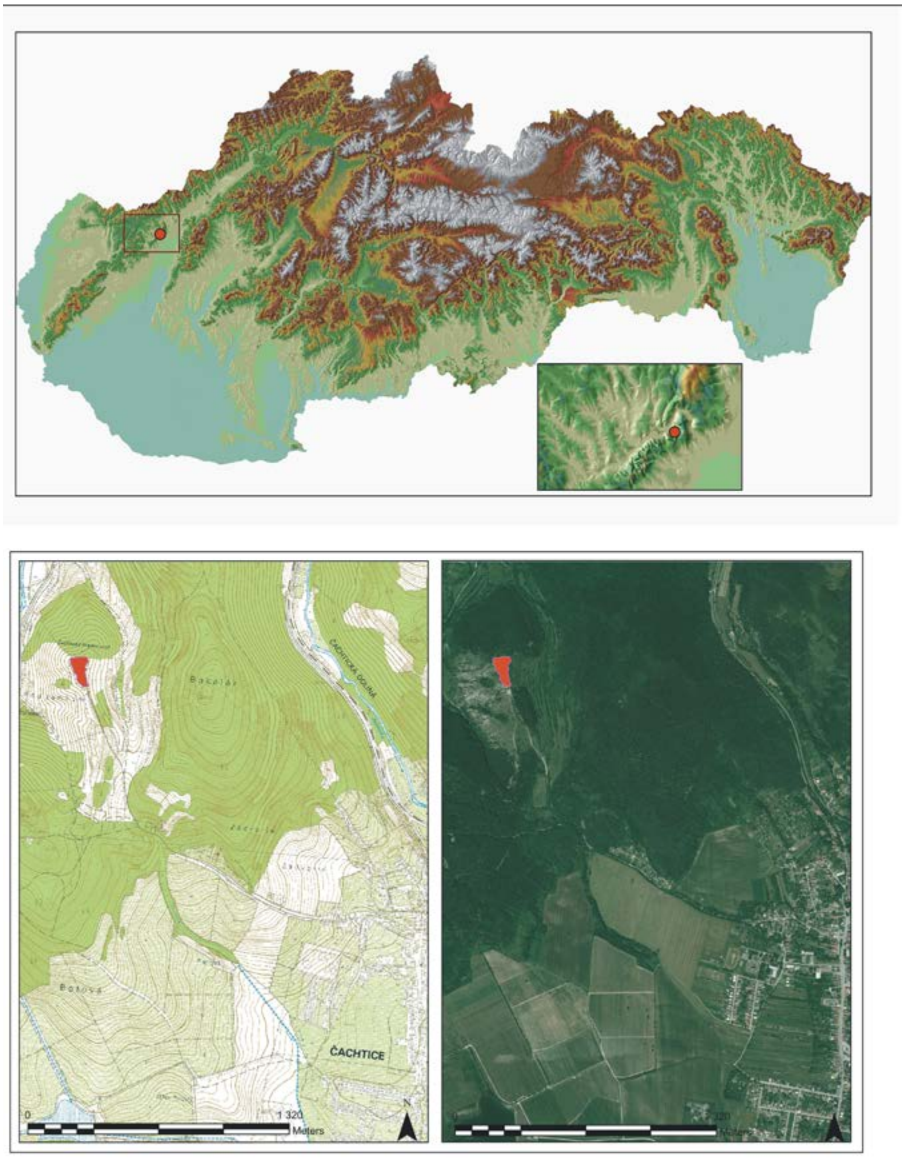


FIGURE 1: LOCATION OF THE ČACHTICE CASTLE.

Karpaty mountains between the villages Čachtice and Višňové (Fig. 1).

The former frontier castle served as a control point on the cross way from the Myjava valley to the Váh valley.

The castle was presumably built after the year 1260 under the patronage of Kazimir of the Hunt-Poznan house. The oldest part of the castle is the pentagonal residential tower situated in the beginning of the terrain disposition and it is dated to the 13th century (Fig. 2). During the 14th and 15th century the building activity was intense and it resulted in a new defense area north of the main tower which consisted of new fortification features (cannon bastions), fore-castle parts and masonries. The last major building activities date back to the 16th century when a complete reconstruction of the upper part of the castle was constructed and new architectural features were added to the fore-castle. The castle has been known mainly due to the famous Elizabeth Bathory, wife of Ferencz Nadasdy, who was later known as the 'Blood Countess'. During upcoming centuries mainly maintenance services were carried out. In 1708 the castle was captured and plundered by the rebels of Ferenc II Rákoci. Since the 18th century the castle has been deteriorating (Plaček and Bóna 2007, 92).

In 2012 the project of the castle restoration and static treatment was launched. Restoration works made archaeological excavations of the concerned castle parts inevitable. Excavations were limited to specified areas subjected to the masonry static treatment. The remediation project, together with archaeological and architectural research, was finished January 2014. Archaeological and historical-building research took place mostly in the upper castle. Furthermore, the lower castle and its courtyard were researched partially. In total, 21 trenches were excavated. Other areas where earthworks took place were observed for potential archaeological finds. Additionally to the great amount of artefacts, bottom pavement levels of existing architectures, as well as architectural fragments of earlier castle building phases were discovered.

4.1.2 The Bratislava castle

The Bratislava castle is situated on the dominant peak in the western part of the city center (Fig. 4). The castle hilltop has been intensively settled since late Aeneolithicum. During the late LaTene period the acropolis of the *oppidum* was situated here. Recent excavation unearthed remains of Roman architecture from the 1st century BC, which represents a unique find north of the Danube. The new discoveries significantly changed our current understanding of the historical development of the site.

At this moment archaeological excavations are still ongoing; therefore the archaeological assessment has not yet been completed.

4.2. Traditional documentation

The basic measurements were taken with GNSS Rtk Rover prior to the excavation begin in order to obtain the point field in national grid system (S-JTSK). The point field served as the basis for further measurements carried out by the total station for obtaining ground control points (GCP) used for

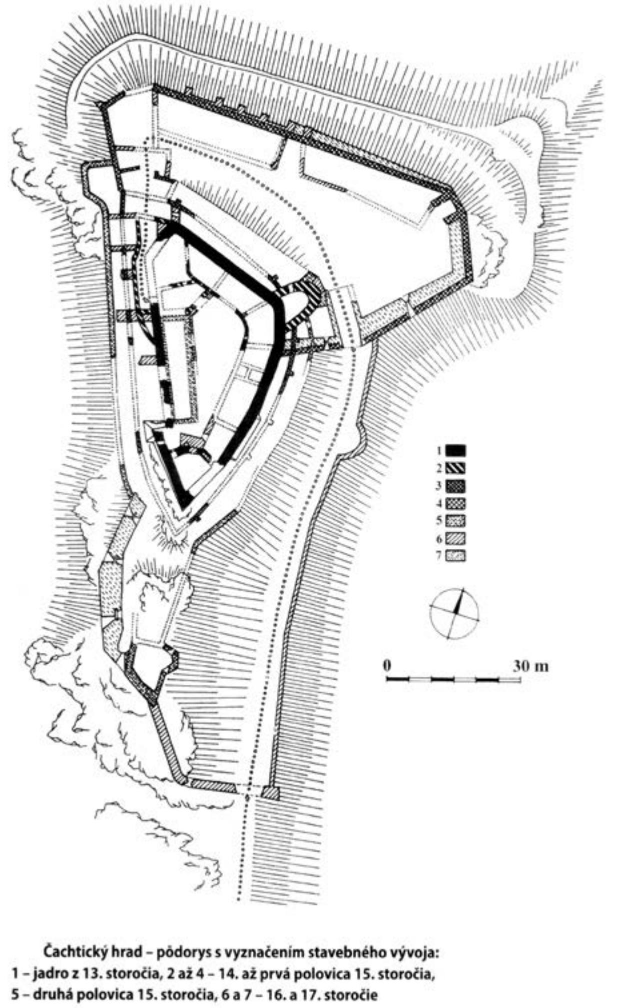


Figure 2: The layout of the castle with analytical visualisation of building phases: 1 - core of the castle (13th century); 2,3,4 - from 14th to the 1st half of 15th century; 5 - the 2nd half of 15th century; 6,7 - 16th and 17t century (according to Plaček - Bóna 2007,92).

Figure 3: 3D model of the Čachtice castle with the surrounding landscape prior to begin of the restoration project. Raw data acquisition was done via ultra light drone. Number of photos used for processing: 80. Applied SW: Photoscan Pro.Version 1.0.0 build 1734.

rectification and georeferencing of the photographs. This operation was carried out in SW Autodesk Civil 3D 2012 with Raster Design extension, which offers several polynomial rectification options. The polynomial transformation of the 1st grade turned out as the most suitable choice in most cases. It is characterized by independent scaling of the x and y axis with no local rectification. In SW terminology, the procedure is referred to as 'rubbersheeting'. Beside 1st grade polynomial transformation the SW offers polynomial of the 2nd and 3rd grade as well (similarly to ArcMap georeferencing options) (Fig. 6). For checking the accuracy root mean square error (RMS) is calculated. The prime goal of this photogrammetric process was to obtain georeferenced orthophotoplans, which were afterwards vectorized in CAD SWs (Fig. 5).

Application of the above mentioned procedure in documenting the masonry structures with significant height distinction between the top and bottom parts of the walls yielded significant distortions and inaccuracies. This proved to be a big problem in application of 2D photogrammetry for documenting still standing architectural structures which represented major parts of archaeological features on Čachtice as well as on Bratislava castle.

4.3. 3D documentation

In the process of documenting Čachtice and Bratislava castle, pole aerial photography (PAP) as well as UAV were applied. Before the excavation was launched, the whole area of Čachtice castle and its broader landscape neighbourhood was documented via ultra-light drone. The photos were subsequently used for the creation of the 3D model of the castle landscape (Fig. 3). For photo acquisition of particular architectural parts and archaeological trenches, PAP or hexacopter were applied. In some cases photos that were not shot with the intention of being raw data for image based modelling were tested as well. In this sense especially vertical photos done with PAP with enough overlapping turned out to be very useful. On contrary oblique ones done by hand held camera very often caused 'doubled' structures by 3D model generation which was mainly due to big angle distinctions between particular photo sequences although overlapping was sufficient.

Figure 4: Location of the Bratislava castle. A-open street map, B-Google earth.

The photographs were processed using Photoscan Professional 1.0.0 and 1.0.2 and post-processed in MeshLab 1.3.2 and a trial version of Geomagic Studio 2013. In the case of Čachtice castle, all data of the particular archaeological trenches and structures was processed as unique sets. Special attention was dedicated to the eastern palace, which underwent the most visible restorational changes. This area was documented in specific stages of archaeological excavations as well as during the restoration work (Fig. 10).

The excavation of the northern terrace of Bratislava castle had to be documented in a different way. Documentation of the excavation was, in this case, carried out in small segments with the prime intention to generate georeferenced orthophotoplans as soon as possible. Due to processing time, particular masonries were segmented into fragments (e.g. facades) that were processed separately. The documentation was carried out in segments that later had to be merged together in the post-processing stage in order to see the complexity of the structures (Fig. 12, 13, 14) .

4.3. Documentation deliverables

Already during the excavations, georeferenced 2,5D DEM, orthophotoplans and cross-section plans were used (Fig. 7, 8, 9). These deliverables were also of prime importance for the on-site interpretation and excavation report. They were generated by Photoscan Professional (versions 1.0.0 build 1734 and 1.0.2 build 1824) and trial version ArcMap version 10.0 with Spatial analyst extension. The interpretations were drawn as plans in the Civil 3D 2012 software package from Autodesk (Fig. 11, 14). For visualization purposes of DEM (analytical hillshading, hypsometry, and contour visualisation) and cross-section presentations, a trial version of Global Mapper 15 and ArcMap 10.0 with Spatial analyst extension were employed.

5. Conclusion

Because 3D image-based data acquisition is a relatively new documentation approach at archaeological rescue excavations, the question of necessity of such documenting arises, especially as there are still many limitations and problems to be solved. For example, as already discussed above, the prime output of excavations is a printed 2D report. Sharing and dissemination of 3D data are still very limited and have not yet become commonly used (e.g. 3DPDF or web shared platforms such as Sketchfab). 3D models have not yet fulfilled their full potential, as most of them above have mainly a presentational role. They are, however, successfully used to extract 2D orthophotoplans and cross-sections.

The main advantage of the full geometrical documentation in 3D on the site is to gain the most accurate and complete recording of the situation. Such documenting means that every individual unit does not have to be measured separately, since the necessary measurements can be then later extracted from a georeferenced 3D model. This implies that we get more data in a shorter time period. Using the Brown model for qualitative lens distortion correction and the 3D model itself bring the most accurate and comprehensive documentation approach that can be used at rescue excavations today. Its advantage is especially noted when documenting complicated 'volume' structures, such as masonries. Furthermore, 3D models significantly assist archaeological interpretation and give a new notion about the explored site.

On the other hand, a larger amount of data requires efficient storing and structuring of the data. Furthermore, it is almost impossible to post-process everything on-site, since the process is very time-consuming. That is why a large part of the post-processing is postponed to the 'cabinet' stage of archaeological works. This speeds up the documentation time on-site, but prevents the possibility of

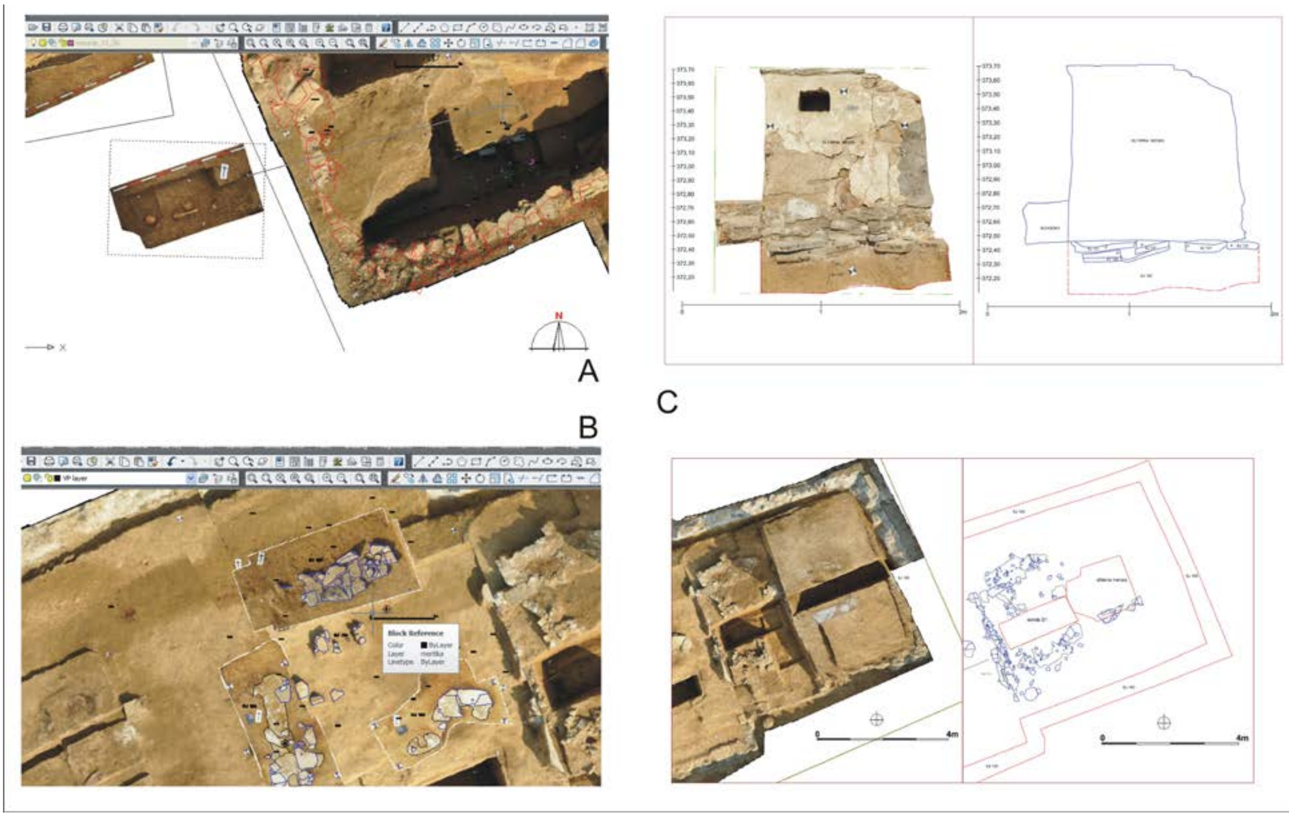


Figure 5: 2D Photogrammetry, basic workflow: A-Georeferencing a photogram according to measured ground control points (orthorectification), B-Vectorization of a photogram, C-Print layout.

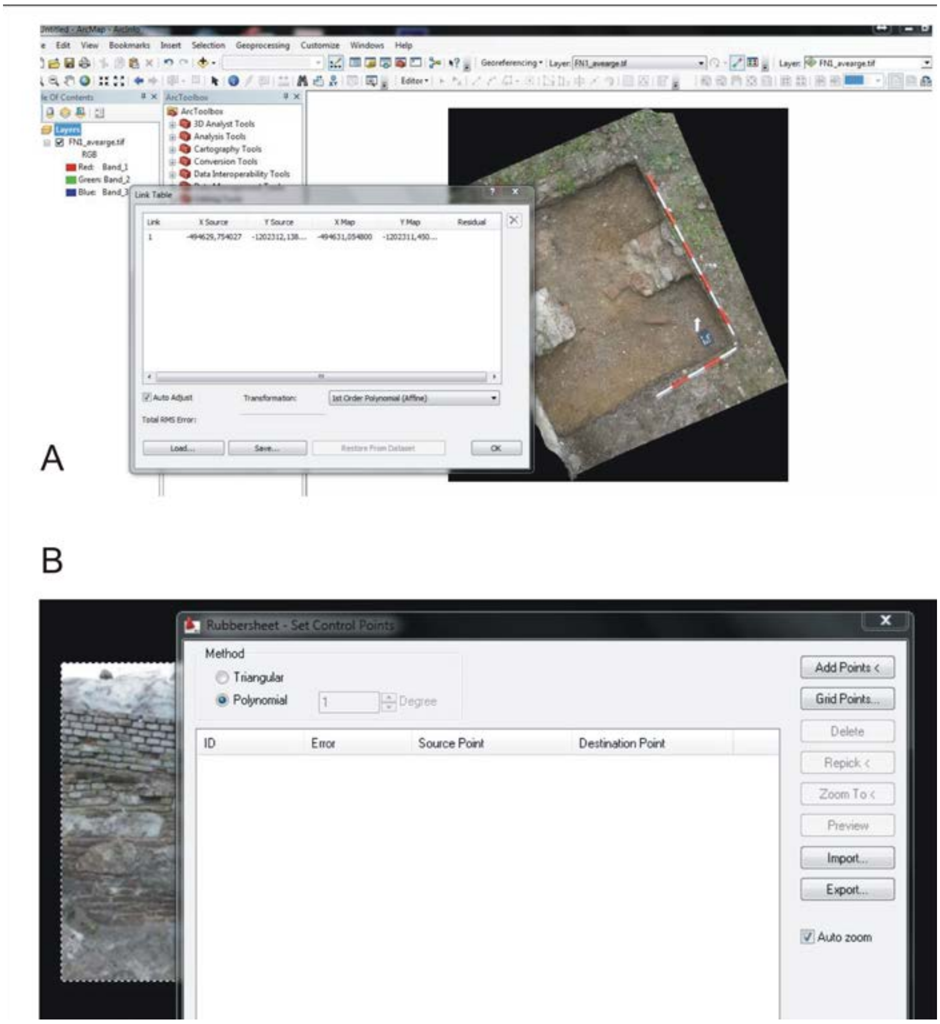


Figure 6: 2D Photogrammetry, orthorectification procedure. A-ArcMap interface, B-Autodesk Civil 3D- Raster Design interface.

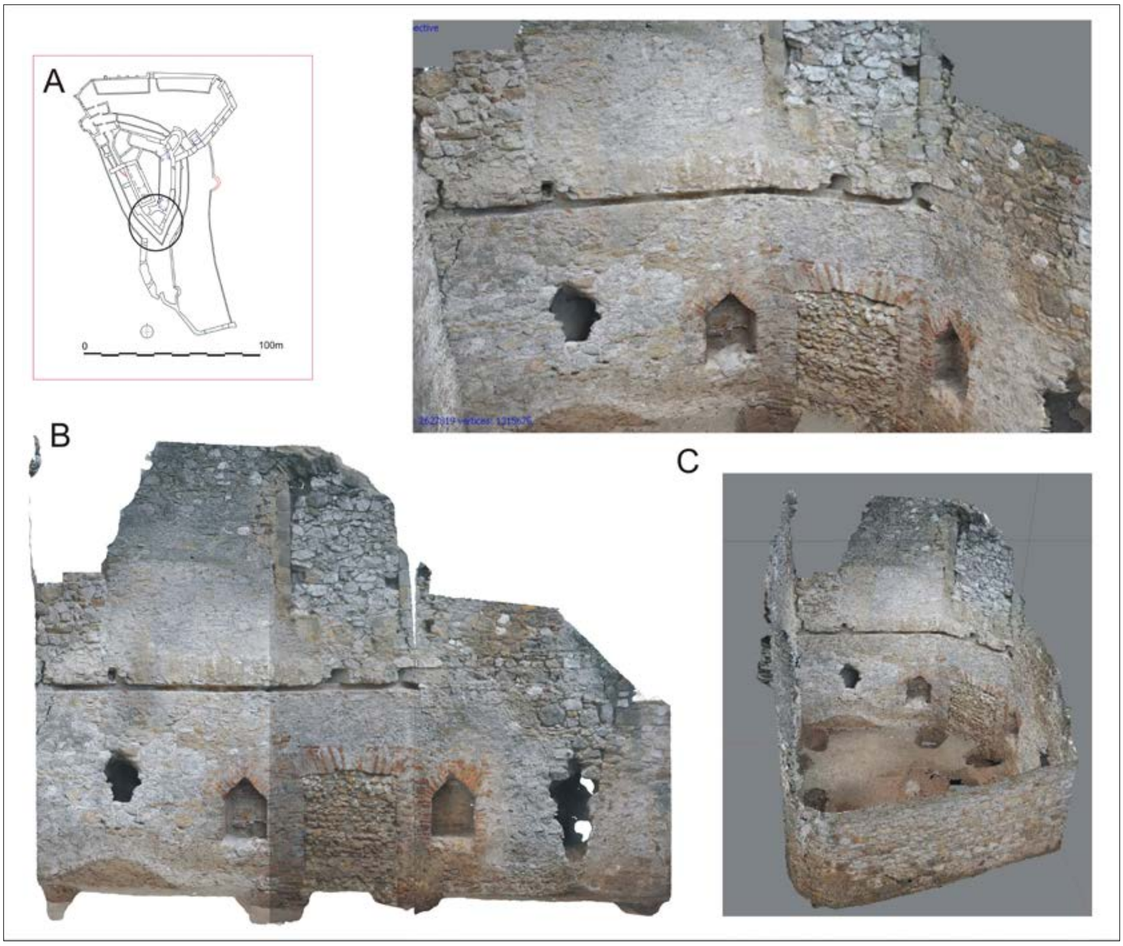


Figure 7: The Čachtice castle. Main castle tower.
A- Location of the tower within the castle layout,
B-Composite unrolled ortophotoplan derived out of the 3D model, C- Solid 3D model of the tower interior with texture map. Photos: 20 (hand handled camera). Mesh: 2 millions faces.

reassessment of the documentation on-site. Moreover, the results are unforeseeable, which together with the lack of time prevents the pursuit of the ideal workflow procedure, which goes as follows: 3D model – orthophotoplan – printed – plan drawing on-site – vectorization of the plan on computer.

Moreover, to successfully integrate the new technique into the existing documentation workflow, certain storage standards for created 3D models have to be met. As all digital media, also 3D models run into the same problems on storage, interoperability and long-term preservation. For the purpose of archaeological excavation documentation, it is important that the 3D models are stored in a structured, organized way. This means an efficient 3D database and registry should be created that would allow for an overview of the 3D models and their metadata (size, location, key words, references...). To the authors´ knowledge there is no such database at the moment that would be tailored to the needs of archaeologists. However, some 3D software packages (such as aSPECT 3D from ArcTron 3D) offer a 3D environment linked to a database with a storage and data retrieval option. Another option is a Geodatabase that can be created using the ArcGIS software package. In any case there is a lot of room for improvements, which should ensure an efficient and flexible structure of a database as well as a simple manipulation of 3D models and their metadata. It is important that this kind of a database is user-friendly and does not require any special preliminary knowledge. Namely, archaeologists usually lack such technical expertise, or, under the constant time pressure of rescue excavations, cannot afford to spend more time on it.

On the other hand it is also necessary to build storage deposits of raw data (in this case photographs used for the 3D image-based modelling) that can be used in the future for comparison studies as well as the source for possible future processing SWs with new yet unknown processing possibilities.

Acknowledgements

The article was written within the CONPRA project, funded by the European Commission funding under the 7th Framework Programme, Marie Curie actions.

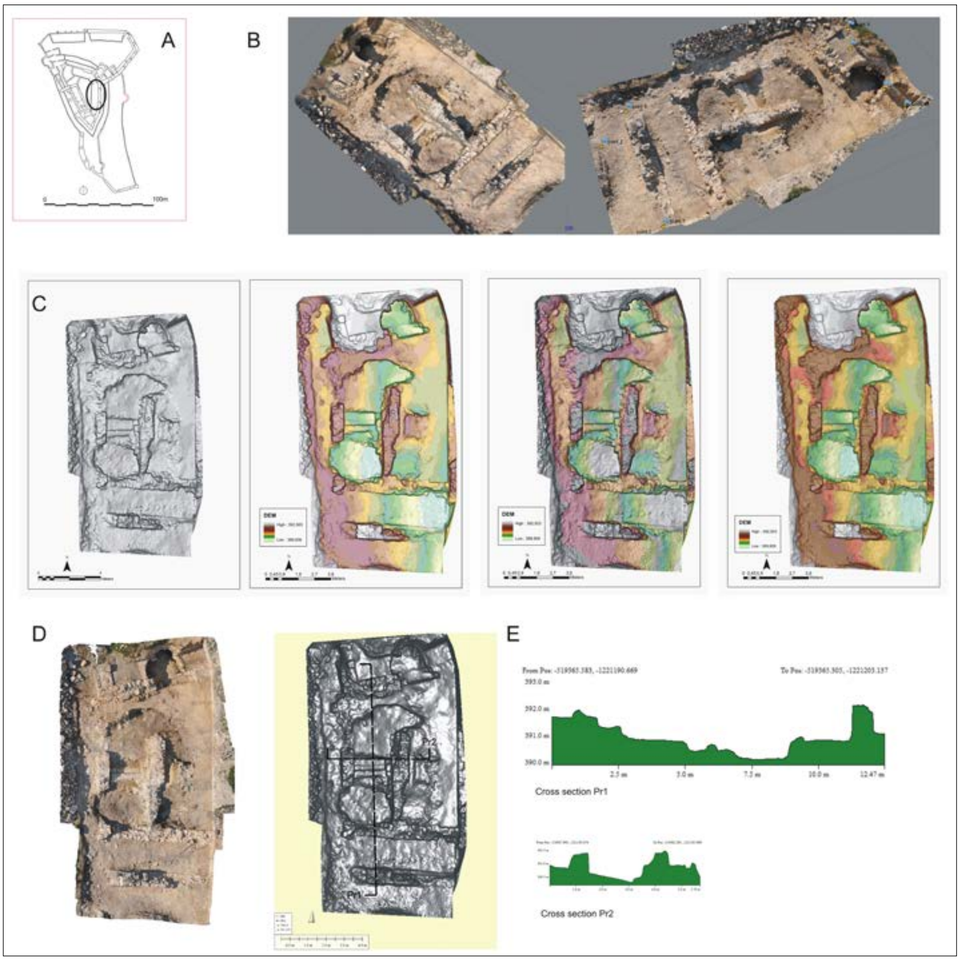


Figure 8: The Čachtice castle. Trench K25 in eastern palace. A-Location of the trench within the castle layout, B- Solid 3D model of the trench with texture map, C-DEM with different forms of visualisation (analytical hillshading and hypsometry), D- Ortophotoplan, E- Cross sections. Photos: 65 (monopod stative). Mesh: 2,5 millions faces.

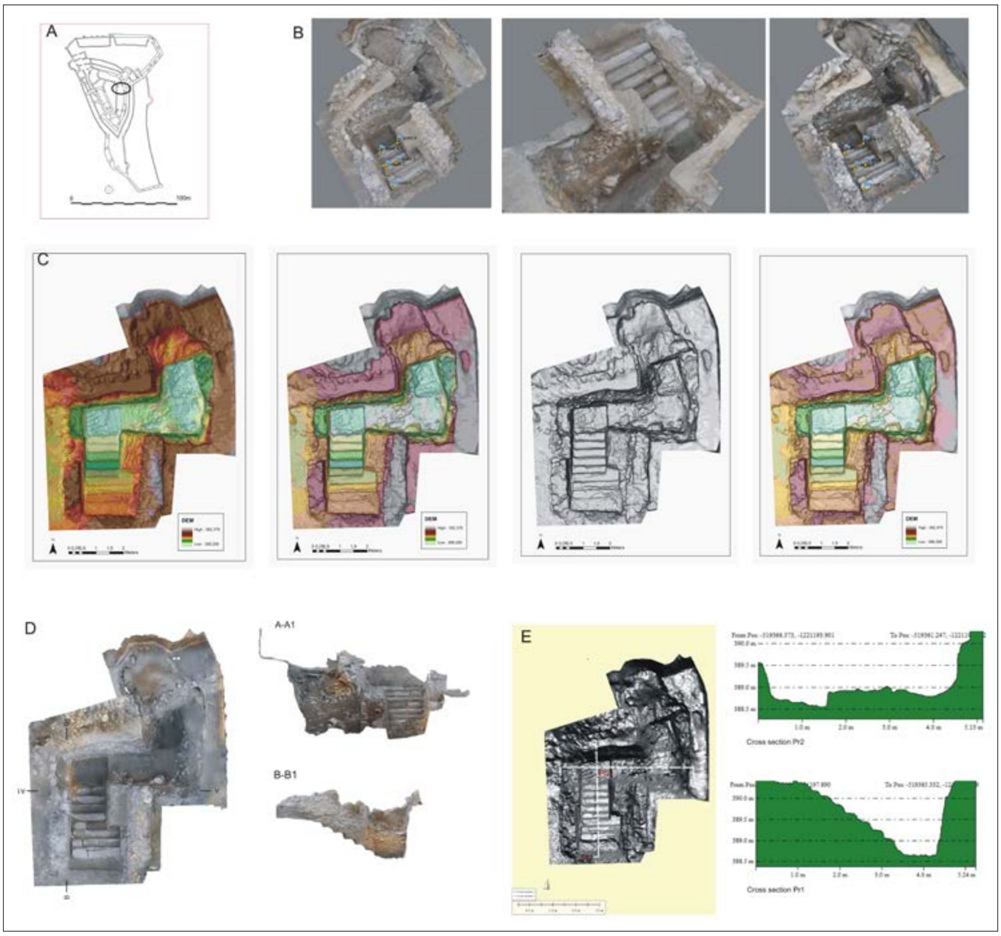


Figure 9: The Čachtice castle. Trench S3 in eastern palace. A-Location of the trench within the castle layout, B-Solid 3D model of the trench with texture map, C- DEM with different forms of visualisation (analytical hillshading and hypsometry), D- Ortophotoplan of the layout and profiles, E- Cross sections. Photos: 28 (hand handled camera). Mesh: 1,1 millions faces.

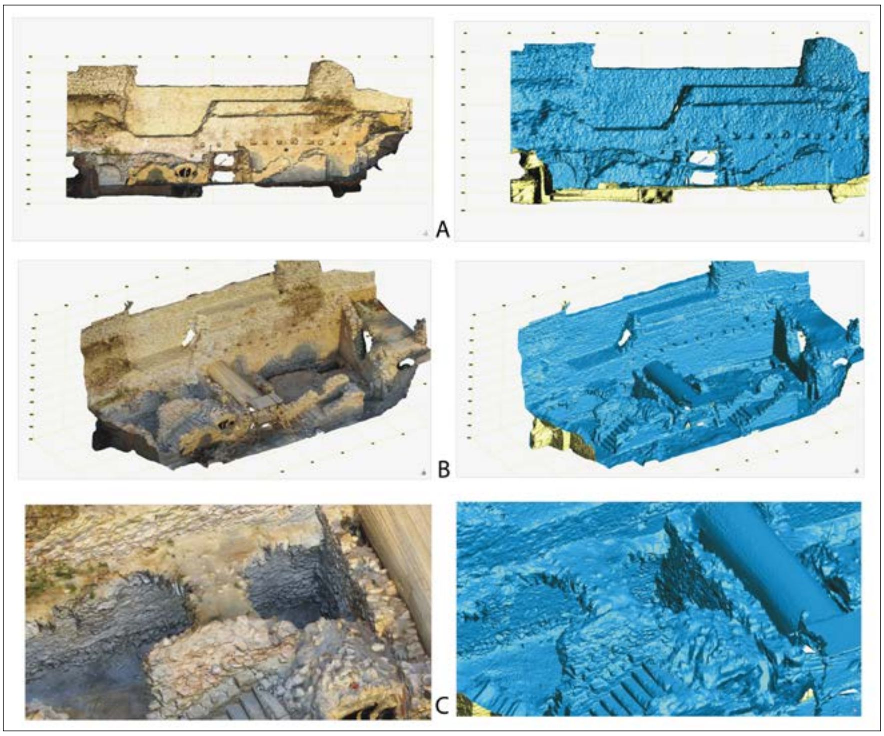


Figure 10: The Čachtice castle. 3D model of the eastern palace during the restoration process.
A-Vertical view towards eastern interior facade,
B-Isometric view,
C-Detail view.
Photos: 250 (hand handled camera).
Mesh: 35 millions faces.

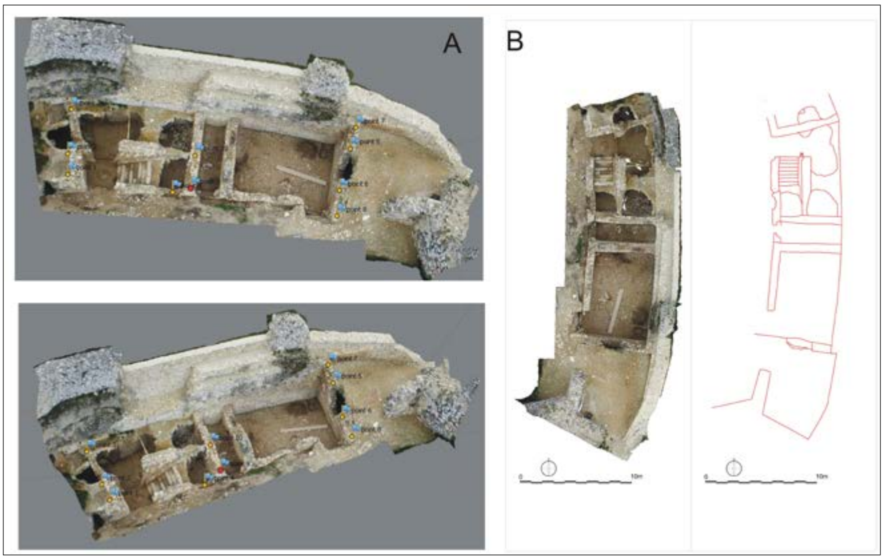


Figure 11: The Čachtice castle. Eastern palace.
A-Solid 3D model with texture map,
B-Vectorised ortophotoplan.
Photos: 45 (unmanned aerial vehicle). Mesh: 1,7 millions faces.

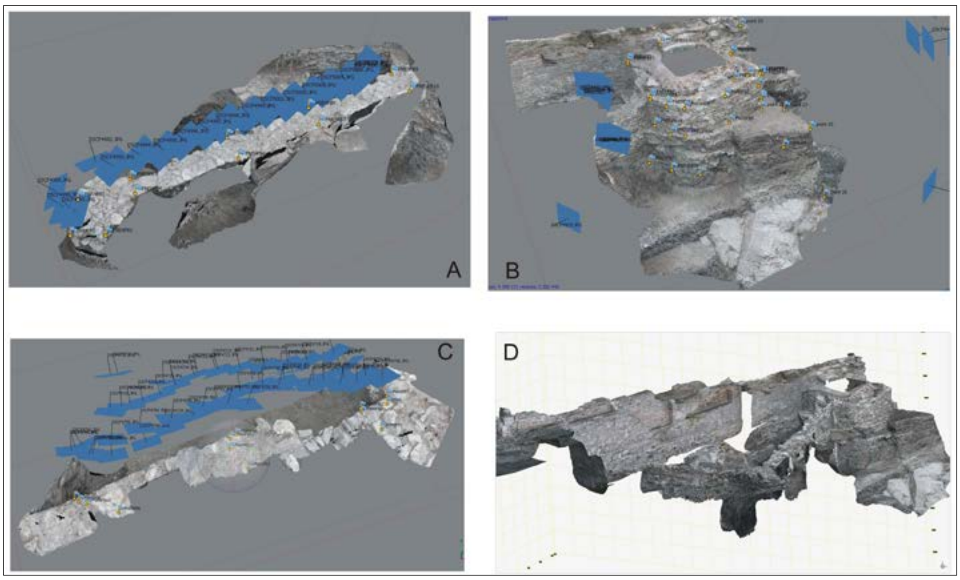


Figure 12: The Bratislava castle - northern terrace. Baroque terrace wall.
3D model was combined of 8 particular models that were generated by stages as excavations advanced.
A,B,C-particular models,
D-combined model.

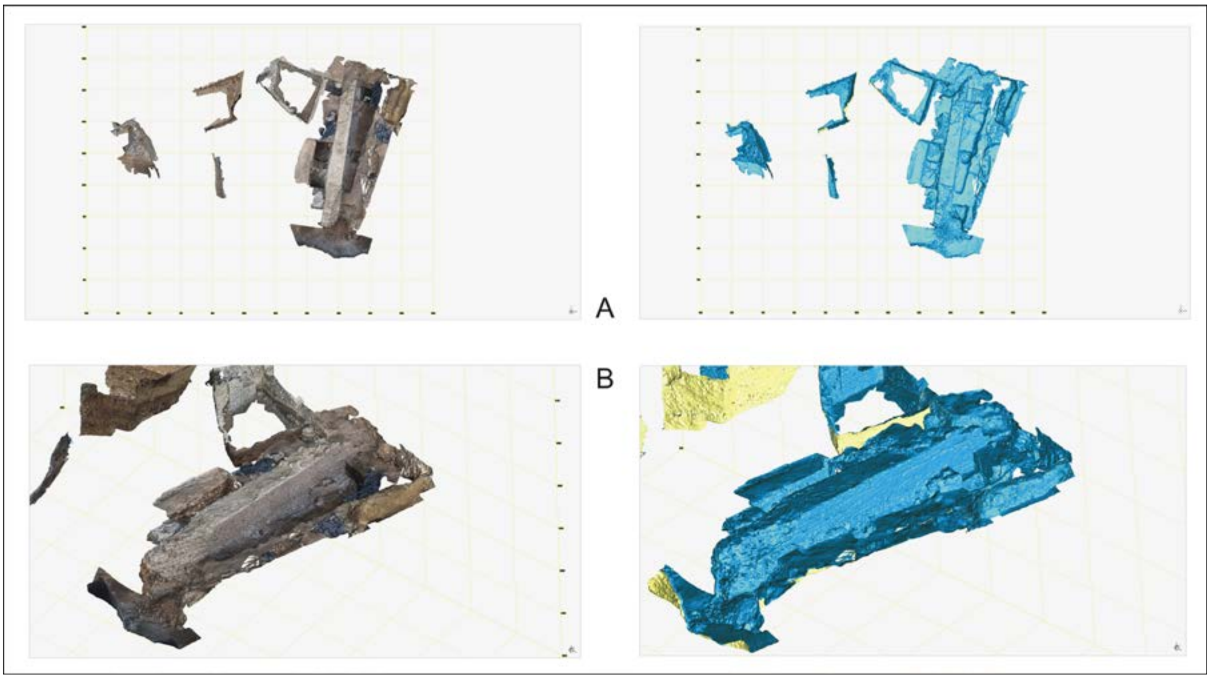


Figure 13: The Bratislava castle - northern terrace. Remains of the romanart architecture.
3D model was combined of 10 particular models that were generated by stages as excavations advanced.
A-top view, B-isometric view.

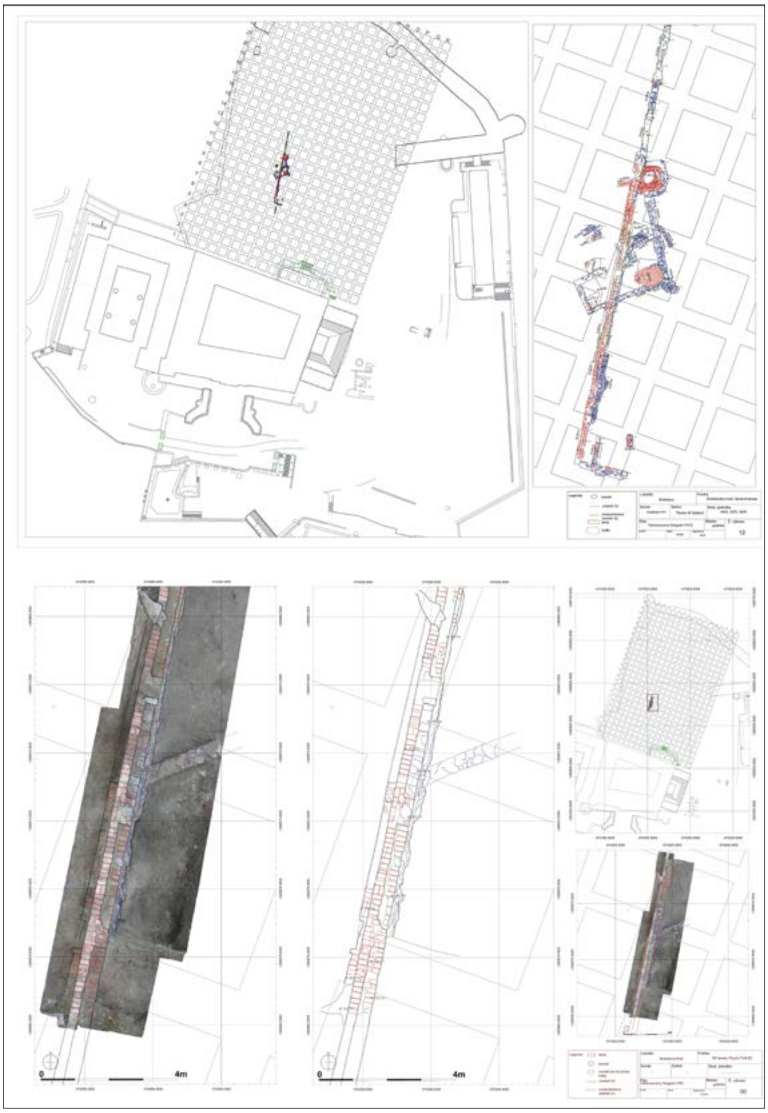


Figure 14: The Bratislava castle - northern terrace. Sample of vectorization of orthophotoplans generated via image- based modeling procedure.

References

DE REU, J., DE SMEDT, Ph., HERREMANS, D., VAN MEIRVENNE, M., LALOO, P. and DE CLERCQ, W. (2014). On introducing an image-based 3D reconstruction method in archaeological excavation practice. *Journal of Archaeological Science* 41, pp. 251–262.

DE REU, J.; PLETS, G.; VERHOEVEN, G. ; DE SMEDT, Ph.; BATS, M.; CHERRETTÉ, B. ; DE MAEYER, W.; DECONYNCK , J.; HERREMANS, D. ; LALOO, P. ; VAN MEIRVENNE, M. and DE CLERCQ, W. (2013). Towards a three-dimensional cost-effective registration of the archaeological heritage. *Journal of Archaeological Science* 40 (2), pp. 1108–1121, checked on 11/29/2012.

DONEUS, M.; VERHOEVEN, G.; FERA, M.; BRIESE, Ch.; KUCERA, M. and NEUBAUER, W. (2011). From deposit to point cloud – A study of low-cost computer vision approaches for the straightforward documentation of archaeological excavations. In *Geoinformatics* 6 (XXIIIrd International CIPA Symposium), pp. 81–88. Available from: http://geoinformatics.fsv.cvut.cz/pdf/geoinformatics-fce-ctu-2011-06.pdf. [Accessed: 4th July 2014].

NOCERINO, E.; MENNA, F. and REMONDINO, F. (2014) Accuracy of typical photogrammetric networks in cultural heritage 3D modeling projects. In *Int. Arch. Photogramm. Remote Sens. Spatial Inf. Sci*. XL-5, p. 465–472.

PLAČEK, M. and BÓNA, M. (2007). *Encyklopédia slovenských hradov*. Bratislava : Slovart.

VERHOEVEN, G.; DONEUS, M.; BRIESE, Ch. and VERMEULEN, F. (2012a). Mapping by matching: a computer vision-based approach to fast and accurate georeferencing of archaeological aerial photographs. *Journal of Archaeological Science* 39 (7), p. 2060–2070.

VERHOEVEN, G.; TAELMAN, D. and VERMEULEN, F. (2012b). Computer vision-based orthophoto mapping of complex archaeological sites: the ancient quarry of Pitaranha (Portugal-Spain). *Archaeometry* 54 (6), pp. 1114–1129.

VERHOEVEN, G.; SEVARA, Ch.; KAREL, W.; RESSL, C.; DONEUS, M. and BRIESE, Ch. (2013). Undistorting the Past: New Techniques for Orthorectification of Archaeological Aerial Frame Imagery. In: Corsi, Ch.; Slapšak, B. and Frank Vermeulen (Eds.). *Good Practice in Archaeological Diagnostics*. Natural Science in Archaeology. Cham: Springer International Publishing, pp. 31–67.

WENZEL, K.; ROTHERMEL, M.; FRITSCH, D. and HAALA, N. (2013). Image acquisition and model selection for multi-view stereo. In: *3D Virtual Reconstruction and Visualization of Complex Architectures. 5th ISPRS International Workshop. Trento, Italy, 25-26 February. 3D-ARCH2013*, pp. 251–258.

WOLF, R. and DEWITT, A. (2000). *Elements of photogrammetry with applications in GIS*. 3rd. Boston: McGraw-Hill.

Beyond Spreadsheets: Digitising the Archaeological Artefact Inventory Process

Arianna TRAVIGLIA,[1] Stephen WHITE,[2] Andrew WILSON[1]

[1] University of Sydney, Australia
[2] Independent consultant
arianna.traviglia@sydney.edu.au, stephenawhite57@gmail.com, andrew.wilson@sydney.edu.au

Abstract

Registration and cataloguing of artefacts has almost completely moved from paper to computers in a variety of forms and with the introduction of powerful mobile devices comes the opportunity to digitally record this process in highly linked and flexible ways.

This paper showcases the process of automated artefact inventory, cataloguing and study applied within 'Beyond the city walls (BCW): the landscapes of Aquileia', a landscape archaeology project based in Aquileia (Italy). Routine collection of artefact samples was undertaken during the field walking survey seasons and a large amount of manifold archaeological artefacts has been collected over the four years of the project. To handle the cataloguing and storing of a considerable number of artefacts an efficient process was set up to minimise the time spent in primary inventory procedures and to automate them as much as possible, as well as provide a valuable analysis tool in the following artefact study phase. The system uses Android mobile devices, Open Data Kit Collect XForms application and an eResearch Data Modelling and Management system, and one of its fundamental aspects is the decoupling of the data collection workflow from the ultimate data model.

As part of the initial development, a conceptual data model was created. The inventory and research workflows were then identified and a data collection form created for each workflow. The form – with the existing data model integration mapping – was then downloaded into the ODK Collect Android application running on Lab-dedicated tablets. Data collected were then uploaded, tracked and ingested seamlessly into the data model of the system.

Choosing a flexible delivery system that allows applets to capture data in a customised way enables adaptation of the system for a variety of data collection needs and to further extend the information gathered to support later eResearch.

Keywords: Archaeological artefacts, digital recording methods, artefact management, museum studies, mobile devices.

1. Introduction

Mobile digital capture of archaeological data is currently the focus of a variety of projects and as the collection of born-digital data becomes increasingly widespread and sophisticated, fundamental issues in processing, ingestion and data modelling come to the fore. Numerous archaeological projects are now taking advantage of reliable and sophisticated hand-held devices to capture data. Successful implementations exist (for field data: Fisher, Terras, and Warwick, 2010; May and Crosby ,2010; Powlesland *et al.*, 2010; for lab data: Nyers and Vollmer 2014), mainly or exclusively in the domain of field data, often with just limited dedicated artefact inventory modules, but many are very project specific or use a variety of tools for different recording tasks, leaving manual integration of the data until later.

The work presented here is a test case in the development of a data collection and ingestion system to speed up, validate and automate laboratory-based inventory and cataloguing of archaeological artefacts for subsequent management and storing within permanent storage facilities. The system uses Android mobile devices (mainly tablets), Open Data Kit Collect XForms application and an eResearch Data Modelling and Management system to implement an integrated approach that supports a seamless flow from form creation to database ingestion dedicated to handle the artefacts accumulated while running the '*Beyond the city walls (BCW): the Landscapes of Aquileia*' archaeological project (2010–ongoing).

This paper reports on the results of this case study and discusses issues including the collection and augmentation of different data types using appropriate tablet applications, support for flexible workflows and decoupling workflows from data models.

1.1. From field to lab

In many archaeological projects, the inventory and study of large numbers of artefacts represent a major challenge. BCW, a landscape archaeology project based in Aquileia (NE Italy) and focused on the landscape transformations occurred in the area as result of the Romanisation process, also faced this issue. The project entails systematic field walking survey in order to verify remotely detected ground traces and anomalies that could signal the existence of as-yet-unknown settlements in the hinterland of this major Roman metropolis. During the period between 2010 and 2012, six field-surveys have been undertaken following conventional procedure of field-walking in (virtual) grids or line transects (Traviglia, 2011). Verifications have been conducted in March-April and October-November periods, with repetition of ground-truthing activities at the same location, in order to ensure systematic coverage over the investigated areas in varying environmental, climatic and seasonal conditions.

Routine collection of artefact samples was undertaken during the field-walking survey seasons and a large amount of manifold archaeological artefacts (>26000 fragments, including ceramic, glass, metals, building material etc.) has been collected over the three years of the field activities. To handle the cataloguing and storing of such considerable number of artefacts within the storing facilities of the National Archaeological Museum of Aquileia and to ensure reliable catalogue data recording and retrieval, an efficient process had to be set up to minimise the time spent in primary inventory procedures and to automate them as much as possible in order to ensure data consistency, as well as provide a valuable analysis tool in the following artefact study phase.

1.2. From paper to digital

In the past the adoption of digital data capture was piecemeal and was often triggered by the availability of specific technology or expertise, the need to solve a particular recording problem, the necessity to add new capabilities to an existing system or the desire to enhance workflows and to speed delivery of results – either traditional or digital. More recently the challenge of integrating and linking captured data to provide a digital research environment has been recognised. Systems such as the Integrated Archaeological Database system (IADB) and Intra-Site Information System (IntraSIS) are notable attempts to build such systems (Rains, 2011; IntraSIS, 2014).

It must be recognised that, in order to be adopted, new systems must build on existing traditions in scholarship and practice and therefore a single universal system may not be practical. This is especially true when one considers the wide range of circumstances in which archaeology is practiced and the variety of requirements imposed by different legal jurisdictions and particular cultural concerns.

The solution provided for BCW is a system that meets a number of needs that were identified at the beginning of the development. The primary requirement was to ensure the collection of systematic data resulting from the inventory and study of archaeological artefacts and to automate time-consuming repetitive data entry typical of lab data capture. Secondly, the application had to be capable of handling several simultaneous users working on separate devices at the same time and to be able to support linking to a number of multimedia files such as images. Further requirements included the need for a device that could provide a more efficient and faster solution than simply typing on a laptop: this to minimise the time spent in recording information and to provide a manageable tool that could face the inconvenience (carrying a laptop) of moving around a large storage facility (including several non co-located buildings) with no or limited desk space. That meant also that the device had work independent of any Wi-Fi or wireless connection. Such devices had to be: affordable as more than one operator was working on the inventory and several of them had to be acquired; sturdy in order to be used in dusty environments; and powerful with capacity of lasting for several hours without need of battery replacement or recharge due to logistics of the work.

The work on which we report here is not the creation of a major piece of digital data collection infrastructure for artefact inventory and cataloguing; rather it is a light, inexpensive and flexible means of using readily available robust digital devices to collect any kind of field and laboratory data and generate an XML output suitable for whatever database management system is being used.

2. Development

The specific aim of this development (Fig. 1) was to automate the capture and ingestion of laboratory data into the existing BCW Data Modelling and Management system without changing existing workflows so that it would conform to an existing complex and highly linked data model. Therefore a fundamental aspect of the project was the decoupling of the data collection workflow from the ultimate data model. As part of the initial development, a domain specific conceptual normalised data model was created and implemented in a server database.The workflow analysis that followed involved division of workflows into independent units of work (*flowlets*) to best facilitate the integration of technology into the work process. The laboratory data items from these *flowlets* were organised into automatic, linked, controlled vocabularies, freetext, photographic and drawing categories and then mapped to the data model which allowed the creation of a data collection form definition that matched the *flowlet* and the database model. These definitions relied on the linked data and controlled vocabularies that were extracted from the database of the Data Modelling System.Finally, a two-phase import process was created for traceability, where raw XML data files and resource files (pictures, drawings etc) were wirelessly upload from tablets and

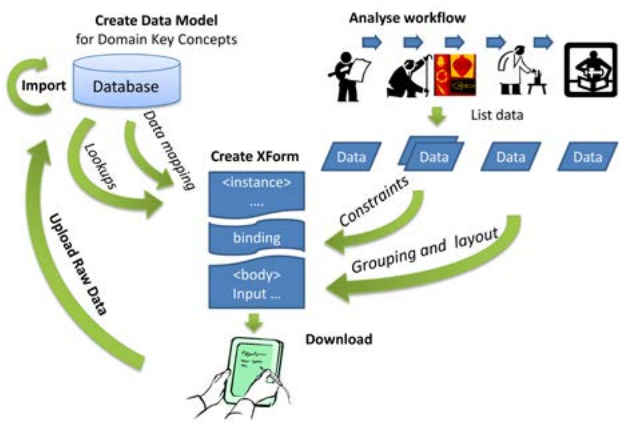


FIGURE 1: A DIAGRAM OF THE GENERAL WORKFLOW.

Item	Type	Support Data	Multiplicity	Dependencies/Constraints
Inventory No.	string		singular	unique
Artefact Type	integer	Lookup Vocab	singular	controlled vocabulary of types
Survey	integer	Lookup FKey	singular	ids of survey records
Quantity	integer	n/a	singular	n/a
Description	string	n/a	singular	n/a
Photo	binary	n/a	multiple	n/a

TABLE 1: *TABLE OF DATA ITEMS WITH TYPE, SUPPORT DATA, MULTIPLICITY AND DEPENDENCIES.*

ingested into the database system as archival XML data records and media files. An import service completed the final step by parsing the XML data to create new records or update existing ones.

2.1. Work Flowlets

Although there are several BCW workflows that were automated, this paper focuses on the processing of artefacts which extends from preparation (including washing, pre-sorting, preliminary typological classification, labelling artefacts), through preliminary inventory (including classification to subcategories, broad periodisation, economic valuation, preliminary description etc.), finishing with thorough study and cataloguing (including refining the typological classification by comparison to published collections, taking measurements, photographs and sampling, etc.). The data model needed for this aspect of the BCW project involved the definition of an *Inventory Item entity* and of an *Artefact entity*, where the *Inventory Item entity* consists of the metadata required for tracking the inventory of artefacts, a preliminary classification level for immediate storage use, and *Artefact entity* is the set of metadata to fully describe all research data observed for a particular artefact. *Artefact entity* was broken up into various subtypes on the basis of physical material employed in the artefact construction (e.g. Ceramics, Glass, Metal...), the metadata requirement changed based on the material. Since the inventory metadata was useful during artefact research *Artefact entity* was defined as a proper subclass of inventory item with the aim to facilitate the extending of inventory items into *artefacts entities* during the research phase.

The data recorded during the inventorying of finds (*Inventory Item*) fell into broad categories of quantification, preliminary classification, administration (including national inventory number temporary and final storage location within Museum premises), valuation and documentation as defined by the project goals and requirements of local CH authorities. Among the more refined data collected for the artefact study and cataloguing (the metadata for *Artefact entity*) were artefact class, artefact subclass, description, provenance, sampling, measurements, drawings and photographs. Defining the metadata for all concepts includes identifying their type, support data, multiplicity and dependencies, as this has a direct impact on the design of the form as shown in the table below (Table 1).

2.2. Data collection Form

Once the data model and workflow had been identified, a data collection XForm XML definition was created for each workflow for use with Open Data Kit (ODK) Collect – a forms data collection tool able to upload multiple data files to a server using a defined and public protocol (JavaRosa). ODK Collect is a component of the open source XML based ODK Suite, a set of tools that allows data collection using Android mobile devices and data submission to an online server (Opendatakit 2014). ODK Collect turns an XForm definition into a sequence of input prompts that apply form logic, entry constraints, and repeating sub-structures. Users work through the prompts and save the response data at any point. The strength of using XForm technology resides in the fact that one can define the structure of the response data and that the definition of the flow of input prompts is separate from, and not dependent on, the order of that structure. Adding an import service which supports import of partial data and update of existing data enables the complete separation of data storage model from data collection workflow, allowing the collection of data object metadata across any number of forms and in any order.

Creating a form involves writing an XML file that conforms to the XForms W3C standard (W3C 2014). This project found it necessary to restrict the creation of forms to the level of ODK/JavaRosa implementation of that standard. An XForm definition has five parts (2 optional) divided between two nodes, <model> and<body>, as shown in Fig. 2.

The parts can be briefly described as follows:

i. The first section in the model element is an <**instance**> element, which defines the structure of the output XML file known as a 'form instance', which will be uploaded to the server.
ii. The second section is optional and contains **Named**<instance id='uniqueName'>**instance** elements for defining Lookup tables of name-value pairs for selection type inputs, which allows for offline data linking.
iii. The third section is optional and contains multilingual label definitions, which was not used for this project.
iv. The forth section defines the **binding** of types, constraints, relevance and processing to nodes in the <instance> data structure.

```
<?xml version="1.0"?>
<html xmlns= "http://www.w3.org/2002/06/xhtml2/"
xmlns:xf="http://www.w3.org/2002/xforms"
xmlns:xsi= "http://www.w3.org/2001/XMLSchema-instance"
xmlns:ev="http://www.w3.org/2001/xml-events"
xmlns:xsd="http://www.w3.org/2001/XMLSchema"
xmlns:jr="http://openrosa.org/javarosa"
xsi:schemaLocation="http://www.w3.org/2002/06/xhtml2/ http://www.w3.org/MarkUp/SCHEMA/xhtml2.xsd
http://www.w3.org/2002/xforms  http://www.w3.org/MarkUp/Forms/2002/XForms-Schema.xsd " >
<head>
<title>BCW Inventory Item Flowlet</title>
<xf:model>
<instance>
      <!--Definition for output xml format that maps to the data base model-->
</instance>
<instance id="lookup list names">
      <!--Definition for lookup list-->
</instance>
<itext>
      <!--Definition for labels for language localization list-->
</itext>
<!—bindings definition section -->
</xf:model>
</head>
<body>
<!— definition for grouping of ordered prompts shown  only upon meeting relevance constraints -->
</body>
</html>
```

FIGURE 2: STRUCTURE OF AN XFORM DEFINITION FILE.

v. The last section is the **user interface** definition, which is a linear ordered definition of prompts or groups of prompts (possibly repeating), which sets up the user experience. This definition of the input elements on the form also defines the layout, paging, labels, help text and selection option data sources. The inputs are linked to nodes in the <instance> data model.

There are several tools available to create XForms definition files, but at the time of development these were not robust enough to describe the workflows, mappings and constraints needed for this project. In order to make fine adjustments to match the workflows, the solution was to create forms by hand and automate the creation processes where possible.

The creation of an XForm definition starts with description of the *Instance model*, followed by the *Binding* for that model, and concludes with the definition of the *User Interface Prompts* in succession with the *named Instance* that they rely on.

2.2.1 Instrance Model

This <instance> element must contain a single child node that becomes the root node of the output response XML, called the 'form instance' file. Defining a 'version' attribute for this node allows ODK Collect to compare updated forms and identify newer versions. The node structure contained within this root node can be any valid XML structure; for this project the nodes fall into categories of: provenance with creator and originating database; recording session metadata including device id, start time and logged in user; state variables to initialise and control the presentation of input prompts; response data structure for import of new records and update information for existing records. It is important to understand that this structure defines what can be in the response file and should include any information necessary for readability of this file for archival purposes, even if this information is not germane to the import process (field descriptions for response data for example). While the response data are stored in the text of the associated nodes only those that have a valid relevance as defined in the binding section will end up in the response file. Lastly, ODK collect allows the definition of templates making it possible for several nodes to be treated as a response unit of user inputs. A simplified model of the inventory item is shown in Figure 3.

The project database had Record Type IDs (rtIDs) and field Detail Type IDs (dtIDs) which were integrated into the response structure for use in the import service as well as the mapping to the workflow.

2.2.2 Binding

The binding section is where one defines the type, importance, automation, constraints, calculation, interdependency and relevance of the input responses. This project found it useful to implement the notion of sub-forms into the workflow that would be presented dependent on some primary selection the user makes. In the instance designed, this equates to a hierarchy of nodes where the parent node is used simply as a wrapping node carrying no data value. Making the parent node 'not relevant'– by setting the state variable used to calculate the node relevance – removes the corresponding input prompts from the workflow as well as the parent node from the output response XML along with its child nodes. For example in Fig. 3 *dt57* contains a record template for a measurement record *rt83*. By making the *dt57*'s binding relevance dependent on the state variable node called 'takeMeasurements' having a value of 'yes' and by adding a workflow control prompt that directs the user to select 'Yes' if they want to enter a measurement, the sub-form of measurement input prompts and *dt57* (with all of its descendent nodes) are added when the user responds with 'Yes' and omitted when the user responds 'No'.

```
<instance>
<formImportXML id="BCWinventory" version="20131005">
<database id="25" urlBase="http://178...65/database/">aquileia</database>
<generatedByuserID="2">Stephen White</generatedBy>
<createdBy/>
<deviceID/>
<createTime/>
<takeMeasurements>No</ takeMeasurements>
<records>
<rt45 depth="0">
<type id="45" conceptID="25-45">Inventory Item</type>
<id/>
<dt28 conceptID="10-352" type="Artefact type" name="Item type" />
<dt84 conceptID="10-54" type="Alt name" name="Inventory Number" />
<dt27 conceptID="2-15" type="Creator" name="Recorder" isRecordPointer="true" />
<dt67conceptID="2-10" type="Date" name="Check-in Date" />
<dt51 conceptID="3-224" type="Photo" name="Photograph" />
<dt57 conceptID="25-457" type="reference" name="Measurement" isRecordPointer="true" />
 <rt83 jr:template="">
  <type id="131" conceptID="25-131">Measurement</type>
  <dt54 conceptID="25-54" type="Dimension type" name="Dimensions type"/>
  <dt55 conceptID="25-55" type="measurement" name="measurement" />
  <dt65 conceptID="25-65" type="note" name="measurement note" />
 </rt83>
</dt57>
</rt45>
</records>
</formImportXML>
</instance>
<bind nodeset="records/rt45/dt57" relevant="selected(/formImportXML/takeMeasurements, 'Yes')" />
```

FIGURE 3: SAMPLE INVENTORY ITEM INSTANCE MODEL.

2.2.3 User Interface Prompts

The prompts section is where one defines the ordering, appearance, choices, grouping, multiplicity and type of input prompts. The ODK collect has built in input widgets for string, numbers, date/time, select (multi select), select1 (single select) and binary (sketch or photo) with the ability to launch any Android application for extensibility in data collection. This project used grouping for sub-forms along with repeatability for things like measurements and photographs, select input prompts for controlled vocabularies and linked data, and simple input for everything else. In the case of multiple mutually exclusive sub-records, the workflow inputs were divided into pages where all inputs for the same page number, regardless of sub-record, were placed on the same page and relevance was used to ensure that only the correct set would display.

2.2.4 Named Instance.

Lookups - Controlled vocabularies

The XForm standard allows the definition of Named instances, which can be used to supply the name-value pairs for a select input widget. The select input prompt definition can reference a Named instance as the source for the option-set by using an XPath like parameterised expression. For example, a set of terms for categorising artefact type (major classification) shown in Figure 4 are linked to the select prompt definition (located in the user input section) by assigning the nodeset attribute of the 'itemset' node the value of 'instance('artefactTypeChoices')/artefactTypeGroup/artefactType'. The mapping inside the 'itemset' ensures that ODK Collect places all artefactType 'name' nodes' text as labels displayed in the dropdown list and 'value' nodes' text as values stored in the model.

```
<instance id="artefactTypeChoices">
<artefactTypeGroup>
<artefactType>
<name>Ceramic</name><value>25-5816</value>
</artefactType>
<artefactType>
<name>Metal</name><value>25-5817</value>
</artefactType>
</artefactType Group>
</instance>
</xf:model>
</head>
<body>
...
<select1 appearance="minimal" ref="/fhml/records/rt45/dt28">
<label>Artefact Type</label>
<hint>Specify the artefact type</hint>
<itemsetnodeset="instance('artefactTypeChoices')/artefactTypeGroup/artefactType ">
<label ref="name"/>
<value ref="value"/>
</itemset>
</select1>
...
</body>
```

FIGURE 4: EXAMPLE OF NAMED INSTANCE FOR ARTEFACT CLASS WITH SELECT INPUT PROMPT DEFINITION.

```
<instance id="workerChoices">
<workerGroup>
<worker>
<name>Dr. A Traviglia</name><value>34955</value>
</worker>
<worker>
<name>Stephen White</name><value>34981</value>
</worker>
<worker>
<name>Andrew Wilson</name><value>34953</value>
</worker>
<worker>
</workerGroup>
</instance">

<select appearance="minimal" ref="/ formImportXML /records/rt45/dt27">
<label>Recorder</label>
<hint>Person responsible for recording inventory items</hint>
<itemsetnodeset="instance('workerChoices')/workerGroup/worker">
<label ref="name"/>
<value ref="value"/>
</itemset>
</select>
```

FIGURE 5: FOREIGN KEY USAGE FOR LINKED DATA.

The Named instances can be subdivided into a refinement hierarchy of related terms or concepts by simple nesting. This also works for cascaded selection, where subsequent choices are defined by the previously selected choice, as one might find, for example, in a country-province-city selection form.

Lookups – Linked Data

In a graph database, as in most conceptual domains, data tend to be highly linked where most records have at least one link to another record. This project chose to tackle the issues around collection of linked data and seamless import of that data. The problem can be broken down into two sub-problems: (i) representation of links for new and existing records in the collected data and (ii) resolution of those link representations during the import process.

After analysing the workflow patterns, it was clear that data collected in one workflow needed to be available in another, requiring cross-form data sharing. This project first investigated the ability to have an external lookup widget or to launch multiple instances of ODK collect, where both solutions proved to be too costly in terms of time and programming resource. The team decided to factor these cases back into the Data Modelling system by prefabricating placeholder records (records with minimal data like temporary title and unique key) and use these in a lookup for cross-form usage. An example of this is the use of the lookup 'workers' across forms, as seen in Fig. 5: the names will be used as labels and the values are actual database foreign keys that will link the inventory record with the person record chosen. Having this list of Title/FKey pairs supports linking new records to existing records in the database. If this list is created for a set of empty inventory items and it is used on two or more forms where the value is be mapped to the <id> node for the record then the import process treats the data from each form as a record update rather than a creation thereby enabling different workflows to contribute partial data to a single record.

2.3. Testing and using the form

The creation of the form is best done in an XML editor that can validate the form creation identifying any syntax issues immediately. After creation, several tests are important before releasing this to actual workflows. The first test consists of compiling the form: this can be done in an Android emulator or by downloading the form to ODK Collect on an Android device. The next level of testing is a basic UI flow test verifying that the forms, sub-forms and prompts display in the correct order and under the correct constraints. Following this, it is necessary to verify that the created response file maps correctly to the input choice and that any irrelevant data nodes are not present for the given form input. The final steps consist of validating that the form uploads to the server, running the import process and checking that the correct records and fields have been populated.

2.4. Ingestion into the data model

ODK collect follows the JavaRosa defined protocol for the submission of form instance data and posts a multi-part HTTP form which contains the response XML file and any related resources, such as photos. From the start, the BCW project saw these submissions akin to pages in a lab notebook – an integral part of the archaeological record – and therefore adopted the two stage import process, uploading and archiving of Lab data and importing these data into the Data Model.

2.4.1 Uploading Forms

The project developed a submission service that accepts the ODK upload of XForm instances and stores those as raw data files creating metadata records in the system for each file. The XML document is linked to a 'form instance' record type and all resources are linked with 'media' record types for archival purposes. The form instance record is given a metadata status of 'uploaded' and a one-shot import URL (preloaded to the import service). In this way, the researcher can upload singularly or in bulk and then process the forms in any order by clicking the URL to ingest the Lab data seamlessly into the Data Model.

2.4.2 Importing Linked Data

An import service was written to read the response file, in the format defined by the Instance model, then constructing new and/or updating existing records. The service makes a single pass through the response XML and when the service reads an 'rt' node it validates the record type and determines an update process by the existence of a valid <id> node. For each record it validates each 'dt' detail type and imports its value. Any fields that are foreign key fields that contain the value from a corresponding lookup already populated with record primary keys are imported direct. For the case where the field is a foreign key which contains a nested 'rt' node, the service uses recursion to import the nested record data (as a new record) returning the record's primary key.

3. Achievements

The most notable achievements of this development is the complete separation of workflow from the data model, which facilitates (i) the integration of data collection into the workflow and (ii) the division of the workflow at any point in the process, thus allowing the record data to be collected across multiple forms.

Another noteworthy achievement is the capability to import fully linked data by exporting needed foreign keys as lookup lists and by structural nesting of the data model for new linked records. For workflows where the data collection is affected by dynamically changing choices, form templates with routinely generated lists are a valid way to manage form changes.

The flexible architecture of this system allows the adaptation of forms to fit the task of the workflow while automating the collection of data. The system is essentially open, so that any new, useful or improved widget can be easily integrated. For example, a form created for inventory, which includes count of artefacts, is augmented with a specialised app/widget for tap counting with audio feedback allowing eyes to focus on artefact counting.

Figure 6: Tablet data entry.

Choosing a flexible delivery system that allows extension via applets to capture data in a customised way along with the decoupling of workflow from data model enables the adaptation of the system for a variety of data collection needs as they change and extend to support ongoing eResearch.

References

FISHER, C., TERRAS, M., and WARWICK, C. (2010). Integrating New Technologies into Established Systems: A Case Study from Roman Silchester. In: Frischer, B., Webb Crawford, J. and Koller, D. (eds), 2010. *Making History Interactive. Computer Applications and Quantitative Methods in Archaeology (CAA). Proceedings of the 37th International Conference, Williamsburg, Virginia, United States of America, March 22-26*. [Online]. Oxford: Archaeopress. Available from: http://www.caa2009.org/articles/Fisher_Contribution191_c%20(1).pdf [Accessed: July 10, 2014].

IntraSIS (2014). *Welcome to Intrasis*.[Online] Available from: http://www.intrasis.com/engelska/index_eng.htm. [Accessed July 10, 2014].

MAY, S.C., and CROSBY, V. (2010). Holy Grail or Poison Chalice? Challenges in Implementing Digital Excavation Recording. In: Nicolucci, F., and Hermon, S. (eds.) 2010. *Beyond the Artifact. Digital Interpretation of the Past. Proceedings of CAA2004, Prato 13–17 April 2004.* Budapest: Archaeolingua. pp. 49–54.

NYERS, A., and VOLLMER, K. (2014). Archie - The Development and Implementationof an Open Source Archaeological Database System. Poster presented at the 2014 SAA 79th Annual Meeting in Austin, Texas, April 23–27. Available from: http://oregonstate.edu/psal/sites/default/files/archie_-_the_development_and_implementation_1.pdf [Accessed: July 10, 2014].

OPENDATAKIT (2014) *Collect*. [Online] Available from: http://opendatakit.org/ [Accessed July 10, 2014]

POWLESLAND, D., MAY, K., RACKHAM, J., and TIPPER, J. (2010). Digit: Archaeological Summary Report and Experiments in Digital Recording in the Field. *Internet Archaeology* 27. Available from: http://dx.doi.org/10.11141/ia.27.2 [Accessed July 10, 2014]

RAINS, M. (2011). Creating a virtual research environment for archaeology.' In Kansa, E., Whitcher Kansa, S., and Watrall E. (eds). *Archaeology 2.0: New Approaches to Communication and Collaboration*, Los Angeles: Cotsen Institute of Archaeology Press. pp. 159–169.

TRAVIGLIA, A. (2011). Integrated archaeological investigations for the study of the Greater Aquileia area. In: Roberto, V. (ed.). *Proceeding of the 1st Workshop 'The New Technologies for Aquileia (Udine*

2011). Available from: http://ceur-ws.org/Vol-806/. [Accessed July 10, 2014]. C1–C11.

W3C (2014). XForms 1.1 *W3C Recommendation 20 October 2009*. [Online] Available from: http://www.w3.org/TR/xforms11/. [Accessed July 10, 2014].

Potentialities of 3D Reconstruction in Maritime Archaeology

Elisa COSTA,[1] Carlo BELTRAME,[1] Francesco GUERRA[2]

[1] Dipartimento di Studi Umanistici, Università Ca' Foscari Venezia, Italy
[2] Laboratorio di Fotogrammetria del Circe, Università IUAV di Venezia, Italy
elisacostas@libero.it, veneziabeltrame@unive.it, guerra2@iuav.it

Abstract

During excavation and documentation of a site many factors affect the choice of the type of technical documentation: features of the evidence, soil properties and the depth of the shipwreck. We describe three research projects to show how 3D reconstruction can be the final step of different types of documentations: 3D photogrammetry, 2D drawings, and 3D trilateration.

3D reconstruction is the best type of documentation, because it allows us to optimize time and work during the excavation, and to obtain complete data for post-excavation study. We have to evaluate the potential for innovation, and new prospects for research and study of this methodology for technical documentation. Furthermore, the 3D model allows us to better record and represent the position of the cargo or the wooden hull; and at the same time it offers an attractive display for the public, who can better appreciate the archaeological evidence.

Keywords*:* Maritime Archaeology, Photogrammetry, 3D Reconstruction, Survey

Introduction

In the field of maritime archaeology, experimentation of various techniques of surveying and documentation is developing rapidly. The aim of this paper is to compare these techniques, and in particular, the three-dimensional survey used for the documentation of underwater archaeological shipwrecks. In some sites, in particular where the topography presents an extensive vertical component, such as a rocky slope, some form of three-dimensional measurement is required (Green, 2004). Difficulties in using underwater technologies normally used in land archaeology, such as total-stations and laser scanners, have oriented the research mainly toward the photographic method (Drap, 2012a; McCarthy, Benjamin, 2014). Integrating this method with a technical survey and measurements is the best solution for obtaining a complete, exhaustive and precise study.

In planning the documentation, there are several factors which influence the choice of the particular underwater survey technology: the characteristics of the various archaeological contexts, that are the subject to be documented, the morphology of the seabed, the condition of the water, and the depth of the site (which determines the dive time limits).

The team of the Department of Studi Umanistici of the Ca' Foscari University, under the direction of Carlo Beltrame, has been experimenting various techniques in their underwater investigation project. We have examined the results of the use of these techniques to identify the most suitable one for each situation. All the technical methods employed for the study of the archaeological site are focused on three-dimensional documentation. The 3D model is the final product of processing which can be developed from a three-dimensional survey (such as triangulation or photogrammetry), from two-dimensional documentation (such as a simple map and sections), or directly from the measurement of each single find.

The technique of trilateration with linear measurements was first used to produce 3D plans for the excavation of the Mary Rose shipwreck in 1982 (Rule, 1982). Nick Rule at that time proposed the Direct Survey Method (DSM) technique (Rule, 1989). In the late 1990s, Peter Holt developed a software, which is now the most popular documentation technique in underwater archaeology (Holt, 2003). This technique consists in trilateration by measuring the points of the archaeological site from a number of external control points, using simple tape measures and a dive computer to measure the depth. All the readings are processed by a special 3D software to produce points in space with x, y and z coordinates. Those points can be employed to build a model of the objects documented, or to link a points cloud model produced by the photogrammetric technique.

Photogrammetry is of two types: the classical one, which consists of the production of a series of overlapping photographs following a regular scheme from a square grid; that is at a defined and constant distance from the subject. In this way, the camera is always orientated and perpendicular to the object that is being documented (Green, 2004). The other type consists of photographing from free positions without the use of a fixed grid, and without positioned targets in the site, as is necessary in the first type. These photographs must be made both from the

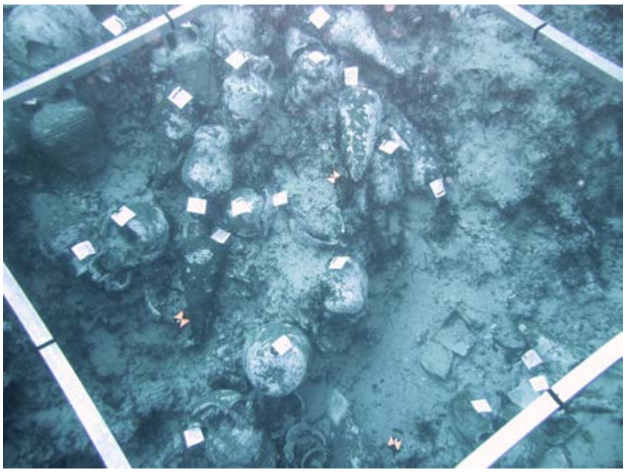


Figure 1: Cape Stoba (Isle of Mljet, Croatia) shipwreck (photo: S. Caressa).

top and from a 45 degree position to determine the lateral angles of the objects (Green, 2004; Skarlatos, Demesticha, Kiparissi, 2012). The photographs must be taken with a considerable overlap; in fact, the key factor of this method is redundancy: each point of measured space must be seen in at least three photographs (Drap, 2012 a; McCarthy, Benjamin, 2014).

We present three case studies of research projects which the Department of Studi Umanistici has carried out in the last two years on different archaeological and environmental situations.

The first one is a cargo of amphoras belonging to the Cape Stoba shipwreck, sunk in the 10-11th century AD off the Isle of Mljet in Croatia, and which is the object of an excavation carried out in collaboration with the Conservation Institute of Zagreb (Zmaic & Miholjek, 2012). This is quite a complex case study because the amphoras lay on a slope at depths from 20 to 28 m (Fig. 1). The 3D documentation appeared useful in allowing a much better perception of the site, which was not obvious from a simple plan with depths marked on it, but where the depths were not visually emphasized.

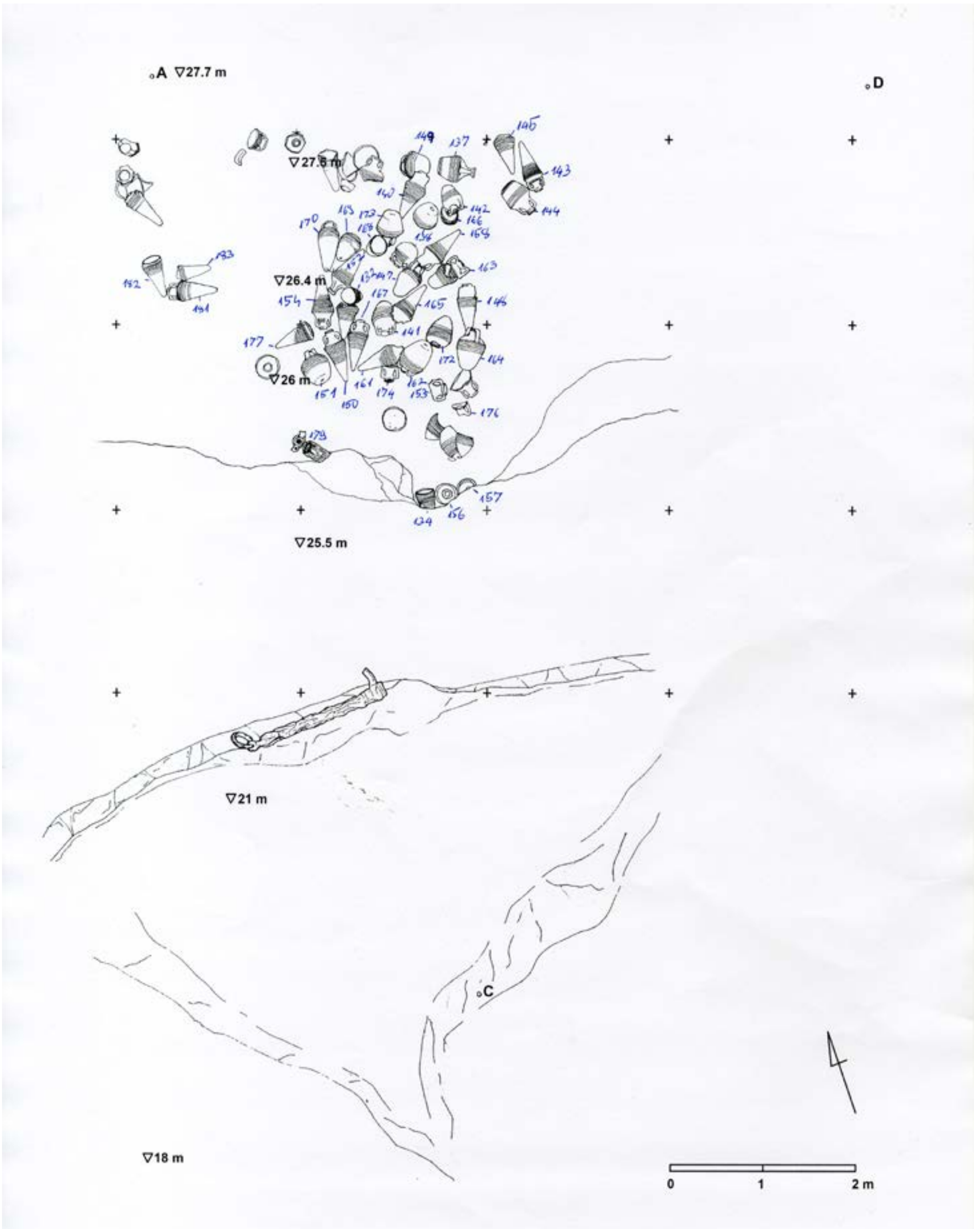


Figure 2: Plan of Cape Stoba wreck, documentation of the 2011 season (drawings: V. Zmaic).

Figure 3: Some of the 3D models of the amphoras (reconstruction: E. Costa).

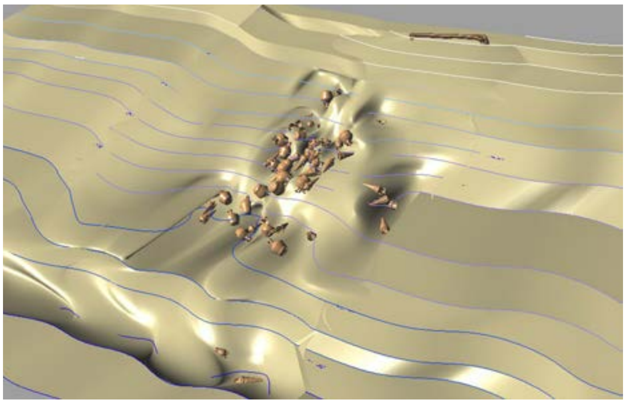


Figure 4: Some of the 3D model of the amphoras (reconstruction: E. Costa, S. Caressa).

Figures 5 and 6: 3D model of the central part of the cargo from two different perspective. (reconstruction: E. Costa, S. Caressa).

The seasons of excavation from 2009 to 2011, coordinated by the Conservation Institute of Zagreb, had produced detailed excavation plans of the cargo, which did not show the variation of the depth of the cargo, which was indicated only randomly on the plans (Fig. 2). At the end of the excavation, in 2012, we decided, therefore, to produce a three-dimensional documentation of the wreck, to show the complexity of the cargo layout in detail. In this situation, we decided to produce the 3D model using photogrammetry from free positions. This type of documentation produces measurable photography, that allows the calculation of the three-dimensional coordinates of each element, and consequently its precise position in space.

This documentation is based on three phases of survey:

- the measurement of the amphoras made in the laboratory after the excavation;
- the theoretical model of the amphoras, the geometry necessary for the graphic representation of the objects;
- the photogrammetric survey based on underwater photography (Drap, 2012 b).

From the two-dimensional drawings of the different types of amphora, we produced the 3D model, both of intact and fragmented amphoras recovered in the site,

Figure 7: Sveti Pavao (Isle of Mljet, Croatia) shipwreck (photo: D. Della Libera).

producing the outline of the object, which was rotated around the vertical axes of the amphora to create the solid wire-frame (Fig. 3) (Green, 2004). For this we used the software Rhinoceros, an application used to create solid figure, to which is possible to apply photographic textures. After the modeling according to the real dimensions, all the amphoras were scaled and positioned in space by the spatial coordinates obtained from photogrammetry. At a later stage, from the bathymetric survey, which was also obtained by photogrammetry , we produced the model of the seabed, a rocky ridge (Fig. 4).

To finish the model, all the surfaces created were rendered by the application of photographic textures that reproduced the colors and features of the original elements (Figs. 5, 6).

Because before the beginning of the Italian-Croatian collaboration on this site, the documentation was only two-dimensional, we are now working on the plans and the depths indicated on these to produce a 3D model of the entire site.

In this way, the 3D model of the entire cargo and of the morphology of the seabed allowed us to illustrate realistic views of the context from different points of view. This type of representation provided more detailed information than the conventional archaeological site plan (Green, 2004).

The second case study is a late 16th century AD Venetian shipwreck, armed with bronze artillery, with a cargo of Iznik pottery, which was discovered at the bottom of the so called Sveti Pavao cliff at the isle of Mljet (Beltrame, Gelichi & Miholjek, 2014). During the 2010 and 2011 seasons, the Croatian and Italian mission decided to excavate and document a section of a hull, preserved at 42 m, to study its construction technique (Fig. 7).

Due to the depth, beyond sport-diving limits, which, although the divers used Nitrox, did not allow more that 18 minutes' bottom-time, we had to use a very fast technique for the documentation. At the same time, this technique had to be precise due to the particular subject, that is a section of wooden hull where each small detail is important for understanding the technique of connection of each element, and where the edges of wooden element were not clearly visible.

Firstly, we decided to carry out photogrammetry from free positions and, secondly, to make a 3D model which could allow a better study of the construction details, also giving the possibility of obtaining measurements of the various elements directly from it. The depth of the hull and the perishability of the wood did not allow us, indeed, to leave the hull uncovered for future investigation, but required immediate covering, with no chance of returning to the site for further measurements.

For the production of the 3D model, made with the Rhinoceros software, the available documentation was only the two-dimensional drawings produced by the photogrammetry, one plan and three cross-sections (Fig. 8), photographs of details of each wooden element and some measurements. Transverse sections were made on only three of the ten frames individuated (Alfonso, 2014). Considering that the thickness of the other frames and of the longitudinal elements as keelson, maststep and *cosce* were not shown, to recreate the correct shape and position of these absent elements, we needed to use, in addition to the two-dimensional drawings, the photographic

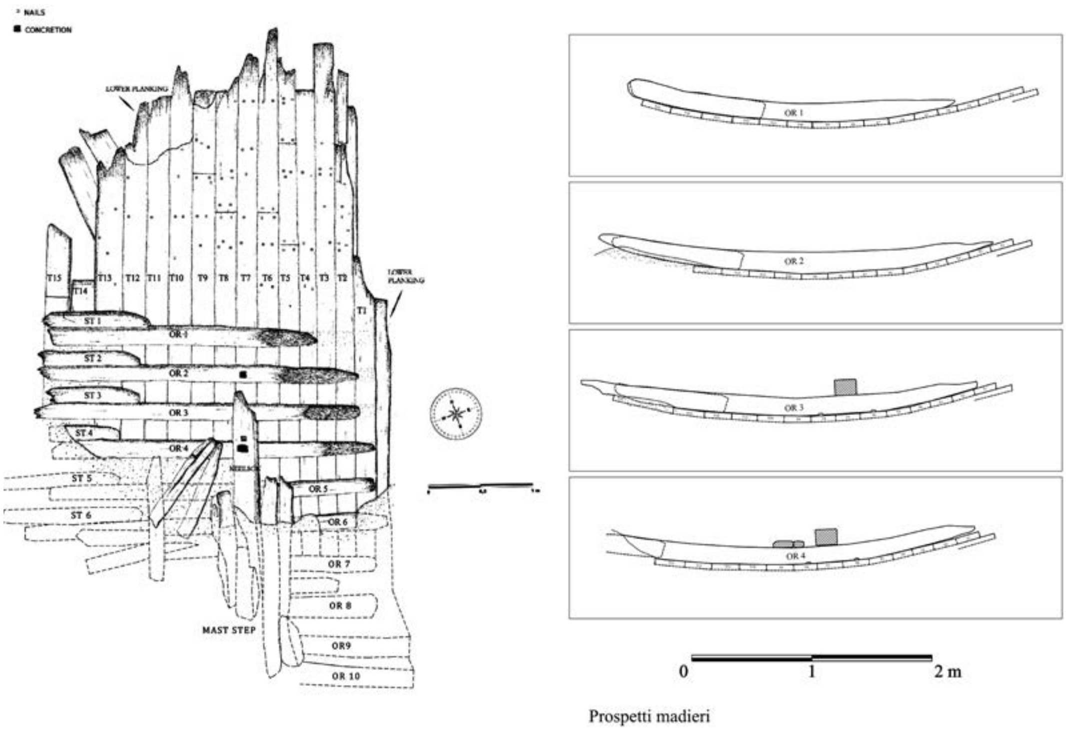


Figure 8: Plan and section of the wooden ship (reconstruction: C. Alfonso).

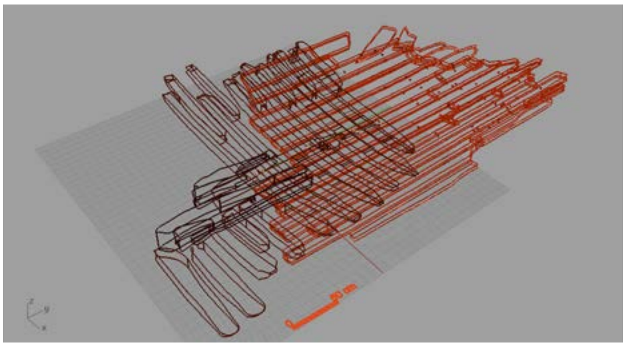


Figure 9: 3D reconstruction of the wooden ship, wireframe (reconstruction: E. Costa).

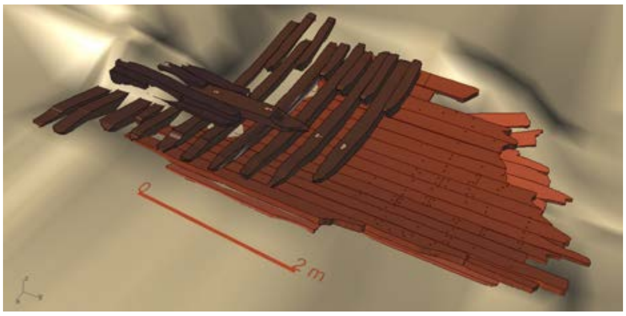
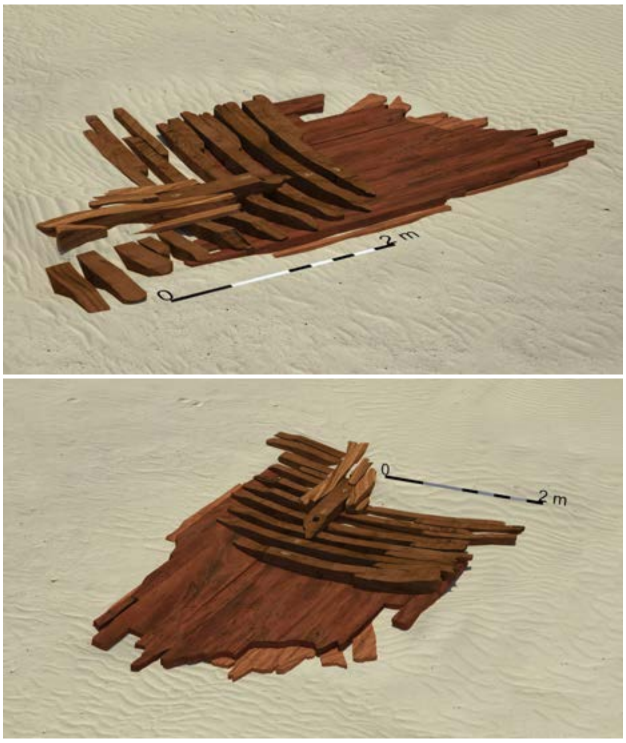


Figure 10: 3D reconstruction of the wooden ship, surfaces (reconstruction: E. Costa).



Figures 11 and 12: 3D reconstruction of the wooden ship, textures (reconstruction: E. Costa).

Figure 13: Punta Scifo D (Crotone, Italy) roman shipwreck (photo: D. Della Libera).

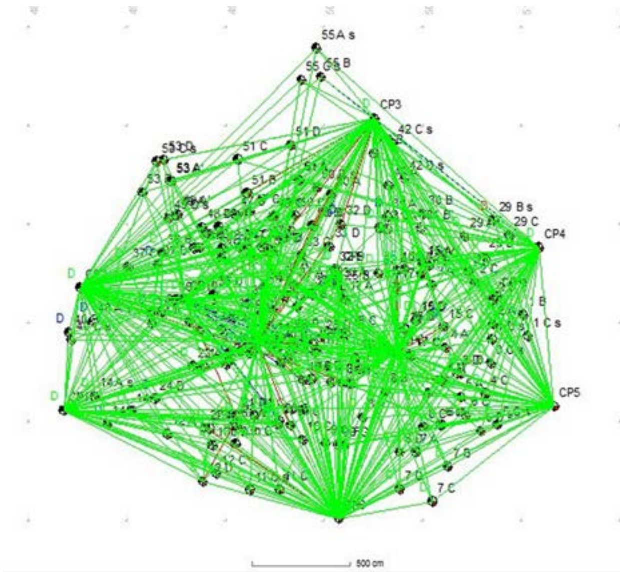


Figure 14: Triangulation of the points with DSM (processing: E. Costa).

Figure 15: 3D reconstruction of the marble cargo, Rhinoceros (reconstruction: E. Costa).

documentation and the measurements taken underwater by the naval archaeologist. The combination of these various types of information helped us to illustrate all the construction elements of the wooden ship, modeled on the actual dimensions of the shipwreck (Figs. 9, 10). For the rendering we used different textures for each species of wood, with specific wood grain and color, which is important to recreate realistic features of wood, essential for the study of the naval construction (Figs. 11, 12) (Costa, 2014).

This kind of documentation allows us also to analyze, study and appreciate technical details of the construction of the hull from different perspectives and allows us to measure construction elements every moments after excavation, without the necessity of re-exposing the underwater site.

Furthermore, this model will be used as the basis for a 3D reconstruction of the entire hull before the sinking.

The last case study is completely different; both in its environment, and the archaeological subject. It is a cargo of marble blocks and slabs sunk in the 3rd century AD in the shallow waters of Punta Scifo near Crotone, in Italy (Medaglia, Beltrame & Lazzarini, 2013) (Fig. 13). The depth of 6 m did not impose any limitation to bottom-time for the documentation, but the 20 x 15 m area of the shipwreck and, in particular, the dimensions and different depths of the marble blocks suggested the creation of a 3D model to show the third dimension, and to allow the next process of reconstruction, both of the formation processes, and of the original arrangement of the cargo aboard.

The possibility of working underwater for two hours, for each pair of divers, and the square shape of the blocks suggested the use of the DSM technique which requires the construction of a net of control points all around the site to take both linear measurements and depths (with a simple dive computer). The data were processed with Site Surveyor software to create a series of points with x, y, z coordinate (Fig. 14). It was then possible to import this net of points into 3D software to model and render the cargo (Fig. 15).

The photomosaic, made during the 2013 season, has recently been used to experiment a technique of documentation which is quite new in underwater archaeology.

Today a number of software solutions, based on the Structure-From-Motion (SFM) and Dense Multi-View 3D Reconstruction (DMVR) algorithms, can produce three-dimensional data even with a low cost equipment (McCarthy & Benjamin, 2014). They allow the production of high quality 3D models by using unordered image collections that depict a site or an object from different viewpoints. The main advantage of having uncalibrated cameras, or other types of image acquisition devices, is the easiness of the acquisition process (Dellepiane, Dell' Unto, Callieri, Lindgren & Scopigno, 2014).

These algorithms simultaneously estimate the parameters of internal and external orientation and the coordinates of the object. Some interesting tools for close range applications have been developed in the areas of computer graphics, computer science and automation to support 3D acquisition in a short time and at low costs (Remondino & Rizzi, 2010; Koutsoudis,Vidmar, Ioannakis, Arnaoutoglou, Pavlidis & Chamzas, 2014).

A Nikon D700 camera with a 20 mm lens was used for the photogrammetric survey of a double set of images. Approximately 120 photographs were taken acquired and oriented to model the wreck. The images were processed with the Photoscan software of Agisoft, one of the major commercial SFM-DMVR software currently available. Photoscan is an advanced image-based 3D modeling solution aimed at creating professional quality 3D content from still images; based on the latest multi-view

Figures 16 and 17: 3D reconstruction of the wooden ship, textures (reconstruction: F. Guerra).

3D reconstruction technology. It operates with arbitrary images, and is efficient in both controlled and uncontrolled conditions. Photographs can be taken from any position, providing that the object to be reconstructed is visible in at least two images; both image alignment and 3D model reconstruction are fully automated. The SFM method uses a number of unordered images that depict an object from arbitrary viewpoints and attempts to recover camera parameters and a sparse points cloud that represents the three-dimensional geometry of a scene.

Camera position and scene geometry are reconstructed simultaneously through the automatic identification of matching features in multiple images. These features are tracked from image to image, enabling initial estimates of camera positions and object coordinates, which are then refined iteratively using non-linear least-squares minimization. The procedure of image processing and 3D model construction of the bas-relief can be described as follows:

- The first stage is the camera alignment: the software searches for common points in photographs and matches them, and also finds the position of the camera for each picture and refines camera calibration parameters. As a result a sparse points cloud and a set of camera positions are formed;
- The second stage is building a dense points cloud based on the previously estimated camera positions and the pictures themselves;
- The third stage is obtaining the 3D polygonal mesh, representing the object surface based on the dense points cloud;

- The fourth stage is applying the texture to the polygonal mesh.

The model of the marble blocks produced by this technology is very satisfying, both in accuracy and appearance; indeed the view of the context is very realistic and very attractive (Figs. 16, 17).

From this model it is possible to create ortho-photos and sections, closely resembling the two-dimensional documents commonly used by archaeologists for recording data and publications (Drap, 2012 a).

This technique allows the use underwater of a solution which makes a product very similar to that of a laser scanner, which is impossible to use in this environment and which would cost much more. We must also consider that the points cloud is already photochromatic, contrary to the laser scanner which uses a default color: this characteristic allows us to match the original color to each item.

The technique of photogrammetrical documentation with points clouds appears to be the most promising one, especially in the practicability and cost, because it needs only a simple underwater camera and open-source software (Skarlatos, Demestiha & Kiparissi, 2012). Its appeal is great, and this is a characteristic which we cannot underestimate in the perspective of better diffusion of the archaeological research products.

Conclusion

considering the practical problems of the examples proposed, such as the great depth of the sites, the need to cover the subjects after the documentation and considering that the archaeological excavation is usually a destructive operation (Drap, 2012 b), we can say that photogrammetry and 3D reconstruction together give the best solution. This combination allow us to optimize the work underwater, and to acquire a more complete database than with two-dimensional drawings, affording a good basis for reconstruction studies of the ship and cargo. It is also obvious that three-dimensional visualization presents a global vision of an archaeological site that cannot be obtained by any other survey method (Drap, 2012 b). The 3D model allows a better view of the shipwreck in its context, because it is possible to move , rotate and view the site from any perspective (Green, 2004). Furthermore, it produces an image with great appeal for the public, who are able to appreciate much better the archaeological site with no experience of reading traditional archaeological documentation.

References

ALFONSO, C. (2014). The photogrammetric documentation of the hull. In: Beltrame, C., Gelichi, S., Miholjek, I. *Sveti Pavao shipwreck. A 16th century venetian merchantman from Mljet, Croatia.* Oxford: Oxbow Books. pp. 17-19.

BELTRAME, C., GELICHI, S., MIHOLJEK, I. (2014). Sveti Pavao shipwreck. A 16th century venetian merchantman from Mljet, Croatia. Oxford: Oxbow Books.

COSTA, E. (2014). 3-D reconstruction of the hull. In: Beltrame, C., Gelichi, S., Miholjek, I., *Sveti Pavao shipwreck. A 16th century venetian merchantman from Mljet, Croatia.* Oxford: Oxbow Books. pp. 22-24.

DELLEPIANE, M., DELL' UNTO, N., CALLIERI, M., LINDGREN, S., SCOPIGNO, R. (2014). Archaeological excavation monitoring using dense stereo matching techniques. *Journal of Cultural Heritage* 14(3). pp. 201-210.

DRAP, P. (2012 a). Underwater Photogrammetry for Archaeology. In: Carneiro Da Silva, D. (ed.). *Special applications of photogrammetry.* Available from: http://www.intechopen.com/books/special-applications-of-photogrammetry/underwater-photogrammetry-forarchaeology [Accessed: 15th March 2014].

DRAP, P. (2012 b). Archéologie sous-marine: relevé et réalité virtuelle pour l'étude de sites inaccessibles. *Archeologia e Calcolatori* Suppl. 3, pp. 427-439.

GREEN, J. (2004). *Maritime archaeology. A technical handbook.* Second edition. San Diego, California: Elsevier Academic Press.

HOLT, P. (2003). An assessment of quality in underwater archaeological surveys using tape measurements, *International Journal of Nautical Archaeology* 32(2). pp. 246-251.

KOUTSOUDIS, A., VIDMAR, B., IOANNAKIS, G., ARNAOUTOGLOU, F., PAVLIDIS, G., CHAMZAS, C. (2014), Multi-image 3D reconstruction data evaluation. *Journal of Cultural Heritage* 15(1). pp. 73-79.

MCCARTHY, J., BENJAMIN, J. (2014). Multi-image photogrammetry for underwater archaological site recording: an accessible, diver-based approach. *Journal of maritime archaeology* 03/2014. Available from: http://www.researchgate.net/publication/260795817_Multi-image_Photogrammetry_for_Underwater_Archaeological_Site_Recording_An_Accessible_Diver-Based_Approach. [Accessed: 1st May 2014].

MEDAGLIA, S., BELTRAME, C., LAZZARINI, L. (2013). La navis marmorium di età romana "Punta Scifo D" (Crotone). Risultati preliminari della prima campagna di indagini subacquee. *Rivista di Archeologia* 98, pp. 137-165.

REMONDINO, F., RIZZI, A. (2010). Reality-based 3d documentation of natural and cultural heritage sites – techniques, problems and examples. *Applied Geomatics* 2(3). pp. 85-100.

RULE, M. (1982). *The Mary Rose, the excavation and raising of Henry VIII's flagship.* London: Conway Maritime Press.

RULE, N. (1989). The Direct Survey Method (DSM) of underwater survey, and its application underwater. *International Journal of Nautical Archaeology* 18(2). pp. 157-162.

SKARLATOS, D., DEMESTICHA, S., KIPARISSI, S. (2012). An "open" methods for 3D modeling and mapping underwater archaeological sites. *International Journal of Heritage in the digital era* 1(1), March 2012. DOI: 10.1260/2047-4970.1.1.1. Available from: http://multi-science.metapress.com/content/q603163q0k861431/. [Accessed: 17th April 2014]

ZMAIC, V., MIHOLJEK, I. (2012). Medieval Byzantine shipwreck in the Eastern Adriatic, *Skyllis* 11. pp. 94-99.

Cultural Heritage Documentation in Cave Environnement Using Low-Cost Means for Archaeologists
Case Study of the Larchant Caves in the Fontainebleau Forest in France

Aurelia LUREAU,[1] Mehdi BELARBI,[2] Pascal RAYMOND,[2] Régis TOUQUET[2]

[1]Université Paris1 Panthéon-Sorbonne et UMR Trajectoires, Paris, France
[2]Institut National de Recherches Archéologiques Préventives (INRAP), Paris, France
aurelia@lureau.eu, mehdi.belarbi@inrap.fr, pascal.raymond@inrap.fr, regis.touquet@inrap.fr

Abstract

French archaeologists are nowadays often caught between the possibilities of advanced technologies such as photogrammetry and lasergrammetry, and the limits and drawbacks of their outdated tools (computers, digital cameras, etc.) and their limited means.

We mixed old and new technologies by using photogrammetry, free open softwares and an educational sculpting software, in order to obtain scientific, workable, and presentable 3D models for two caves.

We worked on the 'Painting's cave' and the 'Abri Marchais' of Larchant, two caves which are located on the west edge of the stoneware plateau of the Fontainebleau forest, at 70km south-east of Paris in France. Both caves have numerous geometric engravings dated from the Mesolithic period to the modern era.

In order to be able to repeat the test in other caves in Fontainebleau, we tried to create a method for the whole process: from survey to data exploitation to presentation of the results.

Keywords: Archaeology, Cave, Engraving, 3D, Photogrammetry

1. Geographic location of the project

1.1 Fontainebleau forest

The Fontainebleau forest is located in the North of France and occupying 25,000 ha at the South-East outskirts of Paris (Figure 1). It sits on a geological sandstone formation made of very pure and white sand which makes up 98% of the forest ground with almost no spring water. The oldest anthropological traces found in the forest dates back to the Upper Palaeolithic and since then the forest bears traces of human life.

The ONF[1] (French National Office of Forests) aims at reaching out to the population by using a website and a mobile app, by participating to the safeguard of the natural and cultural environment by organizing various public manifestations, and having an active place in research projects. This project was carried out inside a Research Program of the INRAP[2] (French National Institute for Research in Preventive Archaeology) focused at studying Prehistoric and Mesolithic sites in two French regions, directed by B.Souffi.

FIGURE 1: LOCATION OF THE FONTAINEBLEAU FOREST IN FRANCE.

We have worked on a cave and a shelter located in the Larchant district, in the south of the forest: The painting's cave and Marchais' shelter.

1.2 The painting's cave

The cave is 8 meters long with a horse shoe shape, it has an entrance with a high vertical wall overlooking the first chamber with a large engraved green stone seat, and a second chamber, smaller, slightly tilted and also engraved.

[1] ONF website: http://www.onf.fr/
[2] INRAP website: http://www.inrap.fr/

It was discovered in 1959 and excavated in 1980 by Jacques Hinout (Figure 2).

This cave is very important for two reasons:

- it allowed the dating of all caves in the forest thanks to a piece of stone fallen from the green stone seat on the Mesolithic ground with engravings on its hidden face and then covered with other Mesolithic levels of sediments: as it was never moved before the archaeologists dug it, it permitted the engravings' certain dating
- it contains a painting, the only one known today in the whole forest

The engravings of the cave are databled from the Mesolithic period, Middle-Age, modern era, and even from WWI.

The Mesolithic engravings are non-figurative and geometric, most of them are lengthwise, and they can sometimes be slightly curved. They are made with a tool by slotting: rubbing progressively in a straight line to go through the different sandstone layers, despite the hardness of the rock. Most of the engravings in this cave are very deep and wide.

The most frequently found engraving in this region is the 'grid': never identical, sometimes very irregular, sometimes very regular. It's in the Painting's cave that the largest grid of the forest can be found: 81cm long with 34 little furrows. This peculiar type of engraving has been studied but no mathematical pattern has been found to show a common ground for the placement of the furrows between them, their number or their length (Figure 3).

FIGURE 2: ENTRANCE OF THE PAINTING'S CAVE WITH THE GREEN STONE SEAT

FIGURE 3: TWO ENGRAVED MESOLITHIC 'GRIDS'

During the 1980 excavation, approximately 300 engraving tools were found in the cave. Most of them are of sandstone, but some are of flint, which is a material not found in the forest. These tools are the witnesses of long-distance exchanges, and of a will to use specific tools: they may indicate a particular importance attributed to the engravings.

1.3 The Marchais shelter

This shelter presents two major difficulties for a photogrammetric survey:

- the majority of the engravings are located on a curved wall and on an inclined flat 'table'. The curved wall is easily accessible, but the flat table faces the roof of the shelter with a distance of only 50-60cm which makes the survey very difficult. Furthermore, the shelter itself is inclined, with a sand ground and a low roof, and very narrow near the flat table: it is difficult to move and install a tripod.
- unlike in the Painting's cave, the engravings are tiny and not very deep, especially on the flat table which has a very smooth and almost uniformly

FIGURE 4: INTERIOR OF THE SHELTER: THE ENGRAVED FLAT TABLE AND CURVED WALL

FIGURE 5: DETAIL OF THE ENGRAVED FLAT TABLE AND CURVED WALL TO SHOW THE VERY FINE ENGRAVINGS

coloured surface. This flat surface was obtained, according to some archaeologists, by heavy sanding and bears the marks of a palimpsest: the surface has been reused a second time for new engravings or to change or complete some of the older ones (Figures 4 and 5)).

2. Project

Until now, the only data available for scientific use in both caves are drawings and photographs, this data, however, suffers from optical deformation and human interpretation. 3D data, as it was proved many times now, can be objective without interpretation. It moves the interpretation phase from the moment of the survey, to the study of the data.

Our goal was to obtain a 3D model scientifically accurate and usable for research. The goal wasn't to make a 3D model just for presentation or visual aesthetics, but to use the technology available today to archaeologists in order to obtain a survey usable for study and analysis in the process of archaeological research.

We want to clearly maintain this position because today, in the French archaeological world, 3D models are accepted for vulgarisation to the public, but very little for research. At the very best, it is viewed has a tool that can be used to obtain orthophoto and stratigraphic sections of the terrain, but still very little as a mean of research in itself. Our goal here was to show that even with very little efforts it's possible to obtain a scientific survey useful for archaeological research.

The survey was also done with a conservation goal: the rock lives and changes, and the visitors of the forest, sometimes without knowing, can damage the engravings. The inventory of the caves and engravings is in progress by the INRAP, ONF and GERSAR;[3] but the idea was to make a 3D mapping precise enough for researchers but also conservation and for the public.

Our goals were:

- to generate a representation of the engravings with a minimum of optical deformation, precise enough for detailed measurement, for interpretations and preservation diagnosis
- to obtain a 3D model that can be easily manipulated and viewed
- find a way to measure the size (length, width and depth) of the engravings
- to measure the necessary time for the survey of a whole cave plus some details, to evaluate the resources needed for a complete GIS-like mapping for rock art in the Fontainebleau forest

3. Means

We worked at the INRAP's Courneuve office with the tools available to the topographers: 5 computers, 3 digital cameras and a Leica theodolite for topography (Table 1).

[3] GERSAR: Group for the Study Research and Safeguard of Rock Art

Computer	CPU	RAM	GPU	OS
Dekstop1	Core2Duo E6550 2.3GHz	4Go	Radeon X1550	WinXP
Desktop2	Core2Duo E6550 2.3GHz	4Go	Quadro FX370	WinXP
Desktop3	Xeon E12220 3.10GHz	12Go	GTX460 SE	Win7 64bits
Laptop1	Core i5 3210M 2.50GHz	4Go	Intel HD Graphics 4000	Win7 32bits
Laptop2	Core i5 2410M 2.30GHz	8Go	Intel HD Graphics 3000	Win7 64bits
Desktop4	Core i5-2500K 3.30GHz	32Go	GTX470	Win7 64bits

Table 1: DIFFERENT CAMERAS AND COMPUTERS USED DURING THE PROJECT.

Unfortunately, these are not of the latest generation. So, for some of the heavier calculations, we used a personal computer, more powerful than the others. But these outdated tools are often the same as archaeologists have at their disposal everyday and it was important to test the effectiveness of the procedure using this standard equipment.

We used these software:

- Visual SfM[4] for the photogrammetric calculations
- MeshLab[5] for 3D data processing
- CloudCompare[6] to manipulate and compare meshes
- Photoscan[7] for the texture
- SfM Georef[8] for georeferencing
- Sculptris[9] and Mudbox[10] for 3D sculpting

We decided to use as much as possible free open-source software for practical reasons and because open tools and open data formats are, in our opinion, vital for the archaeological collaboration work.

But we used PhotoScan from Agisoft and Mudbox from Autodesk because no other software could perform in very specific tasks as these two did. We used PhotoScan for texturing because it is very simple to use and fast, and once the model is georeferenced with SfM Georef it modifies the file .OUT created by Visual SfM during its calculations. This file contains the coordinates in space of each picture used for the calculation, so their coordinates in the space around the object. PhotoScan understand this file and can create the texture directly. We used Mudbox because it can work with high resolution models. We tried Sculptris from Pixologic which is a free sculpting software, but it could not cope with high-resolution models.

4. Protocol

With our limited means and time (we only had a month to work), we knew it was neither possible, nor useful, to produce a very detailed model of the whole cave and shelter. We made low-resolution global models for both caves integrating them with some high-resolution models: the painting and one heavily engraved wall for the cave and the flat engraved table for the shelter. In order to take

[4] Visual SfM website: http://ccwu.me/vsfm/
[5] MeshLab website: http://meshlab.sourceforge.net/, VCG-Lab : http://vcg.isti.cnr.it/
[6] CloudCompare website: http://www.danielgm.net/cc/
[7] PhotoScan website: http://www.agisoft.ru/products/photoscan
[8] SfM Georef website: http://www.lancaster.ac.uk/staff/jamesm/software/sfm_georef.htm
[9] Sculptris website: http://pixologic.com/sculptris/
[10] Mudbox website: http://www.autodesk.com/products/mudbox/overview

measurements on the models and to integrate them in a GIS, they were georeferenced.

During the surveys of caves such as these, photographers must be very vigilant because, after some time, all the walls seem alike and one can forget what has already been covered. Also, some parts of the cave are very difficult to access, and here photogrammetry takes its whole meaning as a field survey.

5. Digitalizing the Painting's cave

5.1 Survey

We used 700 pictures taken with 3 different digital cameras: a Canon EOS 400D and two Canon EOS 500D, with flash, without tripod. Visual SfM, unlike OSM Bundler allows us to work with different cameras.

5.2 Calculations and point cloud

The 3D dense cloud was calculated in 1450min or 24h with Visual SfM and CMVS/PMVS. The result was a 70millions point cloud, which we could not directly manipulate on our computers at La Courneuve. We decided to decimate it in MeshLab in 3 models and put them together to obtain a complete 1million point cloud. The resolution was still acceptable and the quality very good. The cleaning step was long because there were a lot of parasite points, possibly caused by the sometimes non-optimal illumination and by the high-frequency stone pattern, almost identical on the whole surface.

The façade of the cave was surveyed with the theodolite and a topographical path was created through the woods to the road as far as two topographical nails of the IGN[11] in order to have a georeferenced environment.

5.3 The mesh and its problems

The mesh was generated in MeshLab with the 'Poisson Reconstruction' and the result was a 2millions triangles mesh.

While trying to mesh the cave we encountered problems: the cave, in its rear end has very little space between the ground and the ceiling. MeshLab's 'Poisson Reconstruction' tries to produce closed volumes: the entrance was easy to clean, but the interior of the cave was an issue as ground and ceiling joined together in many places. These issues are normal using the Poisson Reconstruction as this method follows the general direction of the largest number of points: this makes it resilient to noise, but for us it meant that in the less accessible parts of the cave, where we had less points, the reconstruction would often join ceiling and ground and not really follow the real geometry. Another problem during the reconstruction was the parasite points, which created 'blobs' that must be manually cleaned.

So we decided to cut the cave in two: the ground and the ceiling. Once we did this, it was very easy to create the

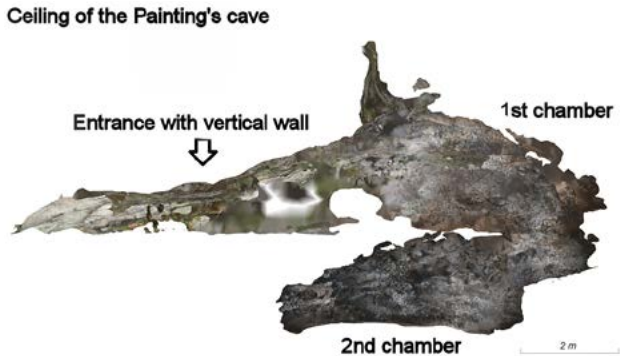


Figure 6: Mesh of the ceiling of the cave after the separation.

Figure 7: Mesh of the ground of the cave after the separation.

mesh, rapidly find the problems and clean them. Once they were cleaned, we put them back together again to obtain the model of the whole cave (Figure 6 and 7). The white parts in the meshes are areas where there wasn't enough information to compute a surface: the sand on the ground, places with little access, etc. We created a filling geometry to have a representation of the ground, but the proposition is just a hypothesis (Figure 8).

5.4 Assembling the various models

The complete model of the cave was done with the goal to have a background for our detailed walls, with a simplified geometry but still accurate. It gives a context, shows the orientation of the wall, and the position in the cave, of the engravings. The detailed models were then put in the correct positions (see Sections 8 and 9).

5.5 Model presentation

The complete model was decimated with MeshLab to obtain a 240.000point cloud and 475.000 polygons mesh, in order to create a 3D.PDF.

By using the newly released open-source tool 3DHOP,[12] it was then possible to generate an interactive web presentation (based on HTML5 and WebGL), using higher-resolution models.

[11] IGN : French National Geographic Institute

[12] 3D Heritage Online Presenter, 3DHOP website: http://vcg.isti.cnr.it/3dhop/

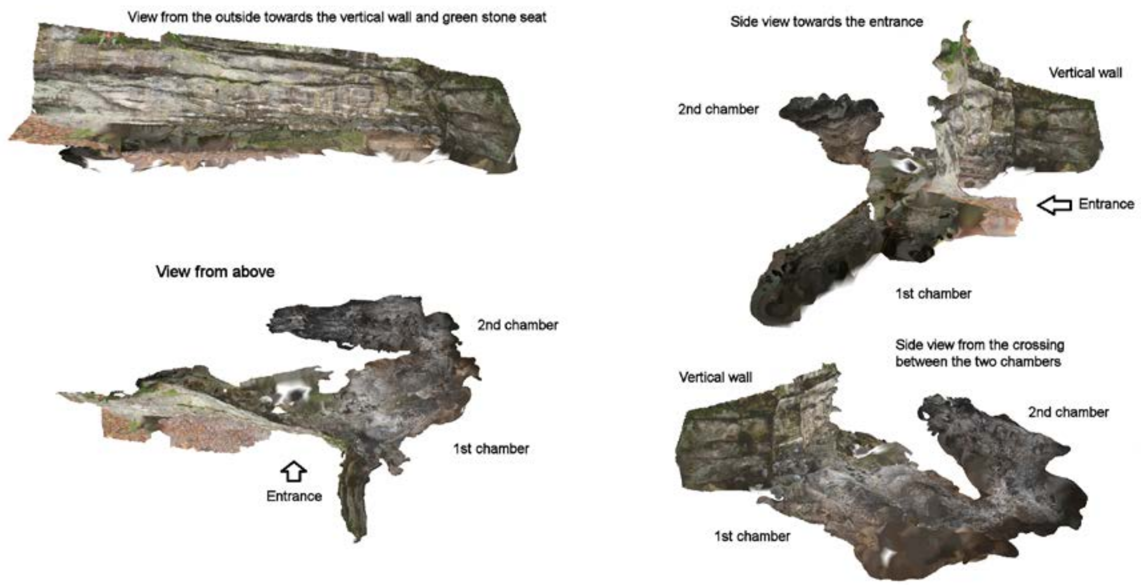


FIGURE 8: VARIOUS VIEWS OF THE PAINTING'S CAVE 3D MODEL.

6. Digitalization and creation of the 3D model of the Marchais shelter

6.1 Difficulties for the pictures

We took around 500 pictures with three digital cameras: a reflex Canon EOS 500D, a bridge Olympus 590-UZ and a compact Nikon Coolpix P310.

The first time in April, we took pictures of the two heavily engraved walls: the flat table and the wall near it. The reflex camera was too bulky for the tiny space, we tried with the compact and bridge. The photos were usable but the flash was flattening the engravings too much. We tried using a home-made diffuser using what was available on the field (a piece of handkerchief sustained by pine needles) producing seemingly better photos. We also made a topographical survey with the theodolite in order to georeference the interior of the shelter with the exterior (already surveyed in 2013).

After the results of the first calculations with the first pictures, we came back one week later to take new pictures of the engraved walls because the results were good but not very precise as there was a lot of noise. This time we used two LED lamps with a neutral tone and we placed the camera on a tripod for stability.

6.2 Calculations

We obtained a global model of 20 million points, a model of the shelter with 15millions points and a model of the engraved table of 7 million points.

The two larger models were hard to manipulate on our computers, so we decimated them on MeshLab with 'Quadric edge decimation' to obtain two models of 5 million points each.

The first batch of models of the table had a lot of noise: the one with the compact camera pictures were unusable and

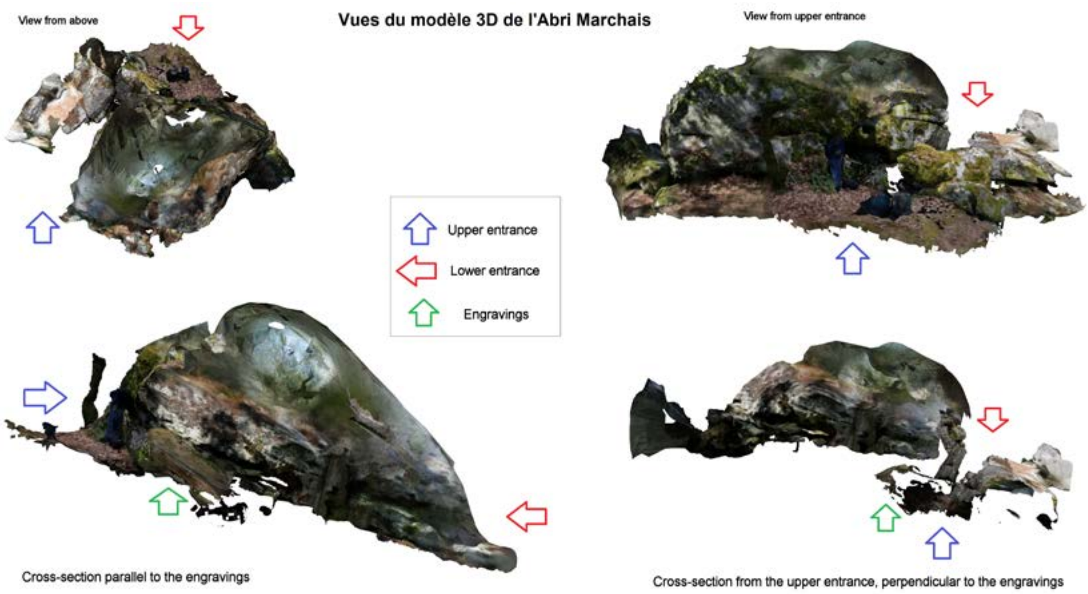


FIGURE 9: VARIOUS VIEWS OF THE PAINTING'S CAVE 3D MODEL.

the other with the bridge pictures would have required too much cleaning. This particular portion of wall is very flat, almost of the same colour everywhere, and engraved with very thin, almost subtle, engravings. We tried a calculation on PhotoScan on the personal computer but it was too noisy to be used. The second take of pictures with the LED lamp gave us much better results with Visual SfM, so we kept this point cloud.

6.3 Meshing and texturing

Meshing was done with the 'Poisson Reconstruction' in MeshLab. The general model needed a lot of cleaning and lacks surface information on some places because of the trees, leaves and various shadows of the forest. The per-vertex colours and texturing aren't very good because of the shadows and light incoherence of the photos: this made impossible for the software to create a uniform texture. Anyway, this model was done only to have a general context of the shelter.

The shelter's 3D model had fewer problems, due to its simpler shape and the presence of many natural geometric features, which helped a lot the camera matching and dense reconstruction. The table model was very thoroughly cleaned and, after the meshing, the result came out really good. Texturing was done in PhotoScan.

6.4 Integrating the different models

We integrated the table model in the shelter's model and then both in the global model thanks to the points taken with the theodolite. As the models were georeferenced in local coordinates, they all come together nicely, but the woods did not allowed us to take precise enough reference points with a GPS, so the shelter isn't correctly orientated and positioned on a global reference space (Figure 9).

6.5 Model presentation

The whole model was decimated again to have a 500.000 faces 3D model in order to create a 3D .PDF with multiple views and two cross sections (as shown on the illustration above).

7. Georeferencing

Georeferencing was done using SfM Georef, an open-source software. We used natural and anthropological ground control points chosen on the pictures and the points surveyed with the theodolite.

This software allows the users to work directly on the pictures to choose the GCP, which is easier than working only on the point cloud. It indicates for every point the residual error x1000 in the general calculation, giving a very accurate feedback on the precision of the work. SfM Georef creates a modified .OUT file, which can be used in PhotoScan to apply the textures.

8. The painting

8.1 The model

We used 50 pictures taken with the bridge Olympus SP-590 UZ using a flash, without tripod or lamp. The idea was to have a much bigger piece of the cave than the painting itself, in order to give it context and surroundings (Figure 10). We needed this model very quickly, so we used the more powerful personal computer and PhotoScan in batch mode. We obtained a 15 million point cloud. We decimated and meshed it in MeshLab in order to obtain a 4 million polygons mesh (Figure 11).

Unfortunately, we could not georeferenciate and place it in the general model, as we did not survey it with the theodolite.

8.2 Dstretch filters

We choose the best picture of the painting and tried several colour filters from the Dstretch plug-in of the ImageJ software. ImageJ comes from medical imaging, where it has been used for almost a decade, now. The plug-in

Figure 10: The painting giving its name to the cave.

Figure 11: The 3D model of the painting and its surroundings.

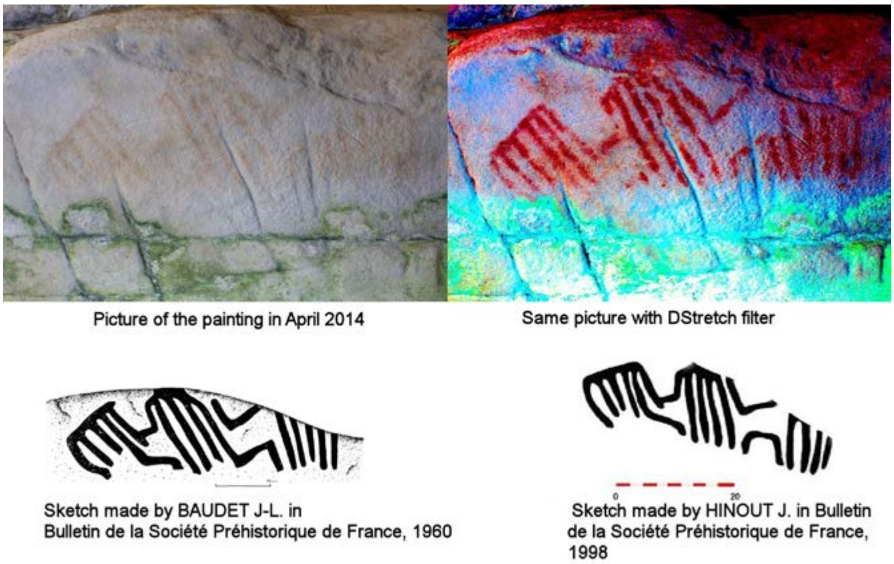


Figure 12: Comparisons between the various representations of the painting.

Dstretch manipulates the colour space in order maximize contrast in the image, bringing out details invisible to the naked eye. This technique is widely used for the study of cave painting by scientists (Figure 12).

We filtered the photo with this tool, obtaining much more readable images, and in comparison with the original picture we can see that it shows much more details than with the naked eye. We also compared the picture and the filtered picture with two sketches done in 1960 and another in 1980 during the excavations, we can observe that in the sketches are present some optical deformations. The filter allows us to observe that the painting is slightly different than its two sketches: the upper lines are almost parallels and the almost vertical lines aren't but they tend to go in the same direction. This, as stated in the introduction, is the result of a non-objective interpretation done at the documentation time; the data collected now is objective, and intrinsically more precise from a technical point of view.

The painting itself is on a horizontal wall nearly 1m from ground, close enough to the entrance to benefit from natural light but not receiving enough to observe it closely or damaging it. So there is no comfortable way to see it, observe it, and sketch it. On the other hand, as good news, these sketches show that the painting seems to be very well preserved.

The advantage of these surveys done in different times is to show the good preservation of the painting itself but also of the wall since no piece of the painting seems missing. It is not always the case, for example a painted cave near Fontainebleau also had different surveys of its paintings and the observations are very different. The Grotte du Croc-Marin at Montigny-sur-Loing has Magdalenian paintings of which we have 3 different surveys:[13] 1960, 1975 and 1980 by different persons. The state of deterioration of the paintings is very visible and preservation is becoming a real problem. It would be interesting to apply Dstretch on those remaining pigments to see if the filters can still see some missing pieces (Figure 13).

Dstretch seems a very interesting tool to help the sketch artist complete his survey. It may even be interesting to have a first

[13] POIGNANT J. (1984), Atlas des grottes ornées, p. 306: http://randos-conviviales.over-blog.com/article-les-abris-de-carriers-de-fontainebleau-1ere-partie-100184475.html

sketch done with the naked eye and a second done with the help of Dstretch to observe the difference between the visual impression of the painting and the real pigment traces. Its use could permit in the future to create a very complete database of the cave paintings of the geological formation around the forest, even around Paris, or even in France. This database could be used, as the 3D models, with a research perspective, but also for the preservation of a very fragile Cultural Heritage.

9. High resolution model of an engraved wall in the Painting's cave

9.1 Digitalization

The photos were taken with two reflex cameras: Canon EOS 500D on a tripod, and two LED lamps. GCP were taken with the theodolite to locate the wall in the cave and have the right scale. We used Visual SfM which worked for 2 hours and produced a 10 million point cloud. We kept it intact but we could only manipulate it on the Desktop n°3.

9.2 Meshing and texturing

With the Poisson reconstruction, we obtained a 20millions polygons model that we decimated into 10 million to be able to manipulate it in real time. This model was then decimated radically in order to obtain a 300.000 polygons model to integrate in the general model for the 3D.PDF for the symposium.

9.3 Integrating the various models

With the GCP taken with the theodolite we were able to georeferenciate the wall within the general model to give it its context, and correct orientation and scale.

10 Tests on measuring the engravings

As a first way to use the high-resolution 3D models of the engravings for study purposes, we tried to define a processing method to highlight them and a method able to measure:

- the depth of the engraving on its whole length and not only on selected areas
- the exact shapes of the engravings to obtain a measurable documentation

We used the high-resolution models of the engraved wall in the Painting's cave and of the table in Marchais' shelter.

After a lot of discussions and thinking we tried 3 methods: we used automatic, semi-automatic and manual mesh processing. Even if we tried several methods, we didn't found an effective processing protocol, able to work on the whole area of the engravings, which do not require significant human intervention and case-specific choices. Using the presented methods we were able to obtain metric information on specific areas with satisfactory results, but a general method is still missing.

10.1 First method: comparison using CloudCompare

We obtained some results by placing a procedural flat surface behind the engraved model, and calculating the point-to-point distance using CloudCompare, but even if this result

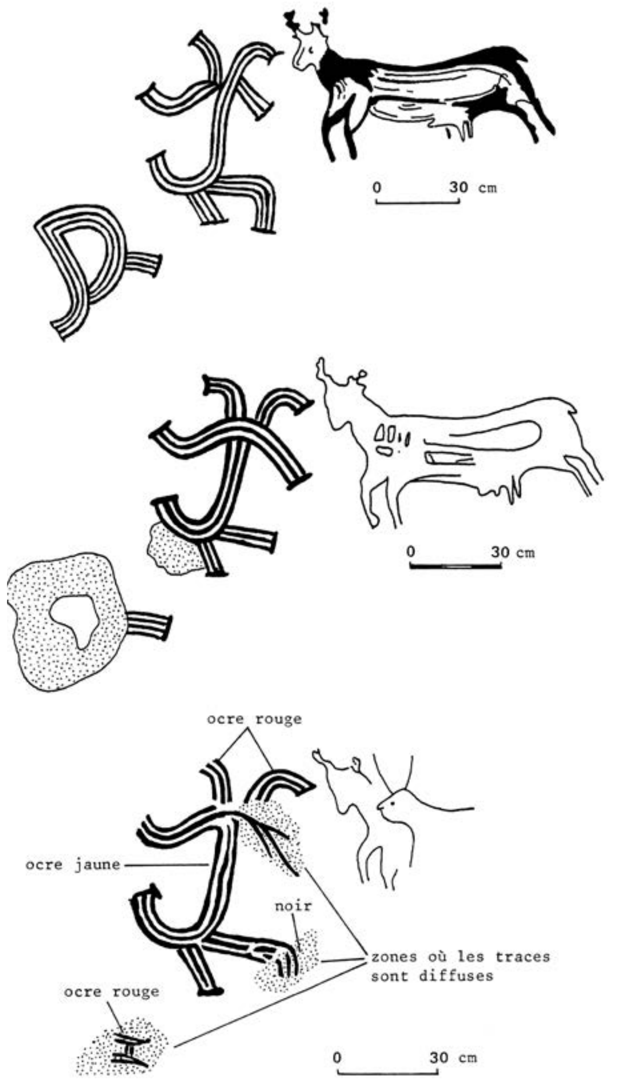


Figure 13: Comparisons between the various sketches of the paintings of the Grotte du Croc-Marin in 1960, 1975 and 1980, where we can clearly see the deterioration of the paintings.

is interesting, the result of the comparison is hard to use, since it is mapped on the 3D surface as a colour-ramp and not directly accessible or rendered as a raster image. Moreover, the approximation of the original surface is quite crude.

10.2 Second method: shifting the normals

10.2.1 Theory

By shifting the normals of the original model of a millimetre for example, we should be able to see which engraving are still crossed by the model, and by doing so, we should be able by shifting the normals little by little to obtain the various depths and come close to obtain the original surface before it was engraved. We would then obtain a topography of the engravings by contour lines.

10.2.2 Application

This shifting keeps the general geometry of the model and the orientation of the concave and convex surfaces. The

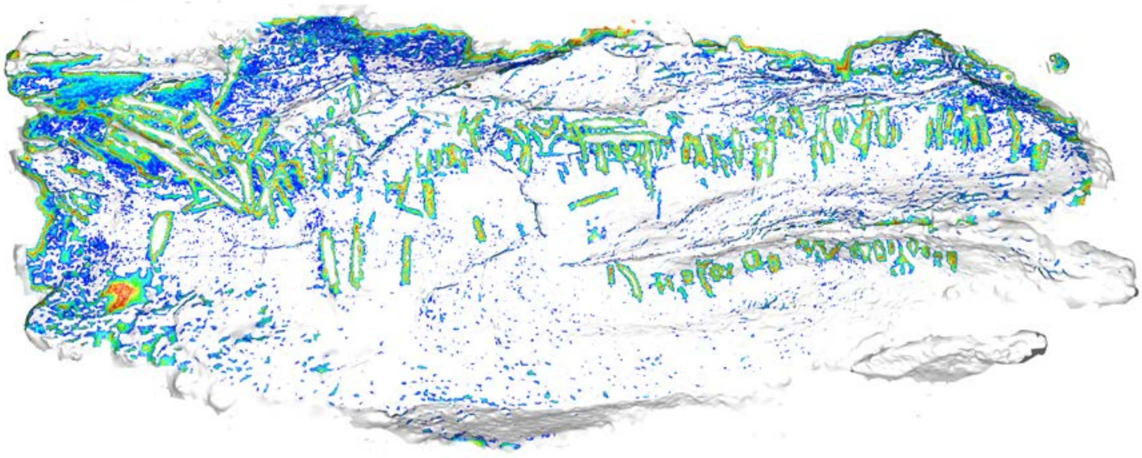


Figure 14: The result of the method 2 with the Hausdorff distance in MeshLab.

new point cloud is then meshed in MeshLab and compared to the original with the 'Hausdorff distance' function. It allows us to visualize the closest points between the two models and enhance the engravings (Figure 14).

11.3 Third method: 3D sculpting

This method was tested using Mudbox software from Autodesk. The idea is to manually 'fill' the engravings on the red mesh, while absolutely not touching the rest of the geometry. In this process, the carving process is revisited in reverse, and the surface offset generated is the geometric 'negative' of the engraving.

We imported in Mudbox two copies of the original mesh: one is locked (in grey) and no more editable, while the other (in red) was edited. The locked grey mesh is acting like a reference to verify that the geometry is altered only in the engraved areas, and that the editing is minimal.

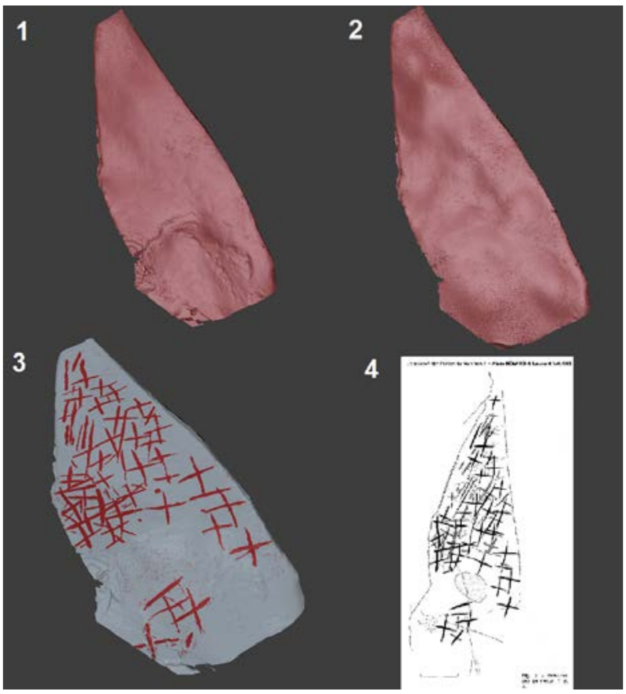


Figure 15: The result of the method 3: 1) hypothesis of the table just before the engravings, 2) hypothesis of the table before the preparation for the engravings, 3) the survey obtained with this method 4) the sketch made by Mr.Benard et Valois in Bulletin n°63 du GERSAR, p.38.

We obtained a 3D model of the Marchais table and the detailed wall of the Painting's cave just before the engravings: once this is done, we can almost think that we are looking at the table just before it was engraved by Humans, nearly 8.000 years ago (Figure 15).

12. Other results using the 3D models

12.1 Roof height and ground level

CloudCompare was able to give us some results while comparing a piece of the shelter and the whole cave with a plane: we have the roof height on both caves and the ground level for the Painting's cave. Having a way to precisely map the height of the ceiling for the whole cave is not a simple task, but with the 3D model it can be automatized (Figure 16).

The illustration with the roof height of the shelter shows the lack of space and its very peculiar shape (Figure 17).

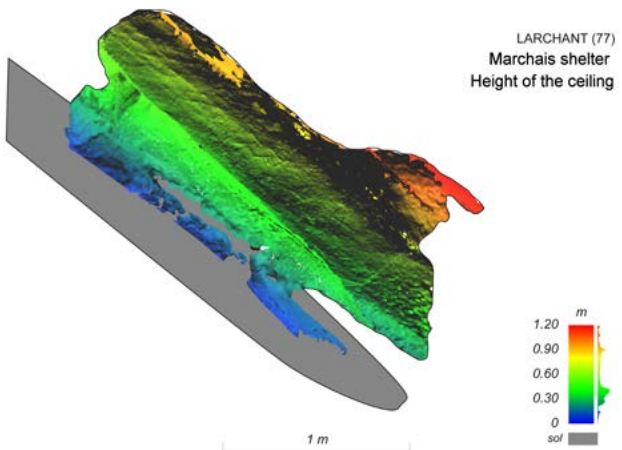


Figure 16: The shelter's roof height.

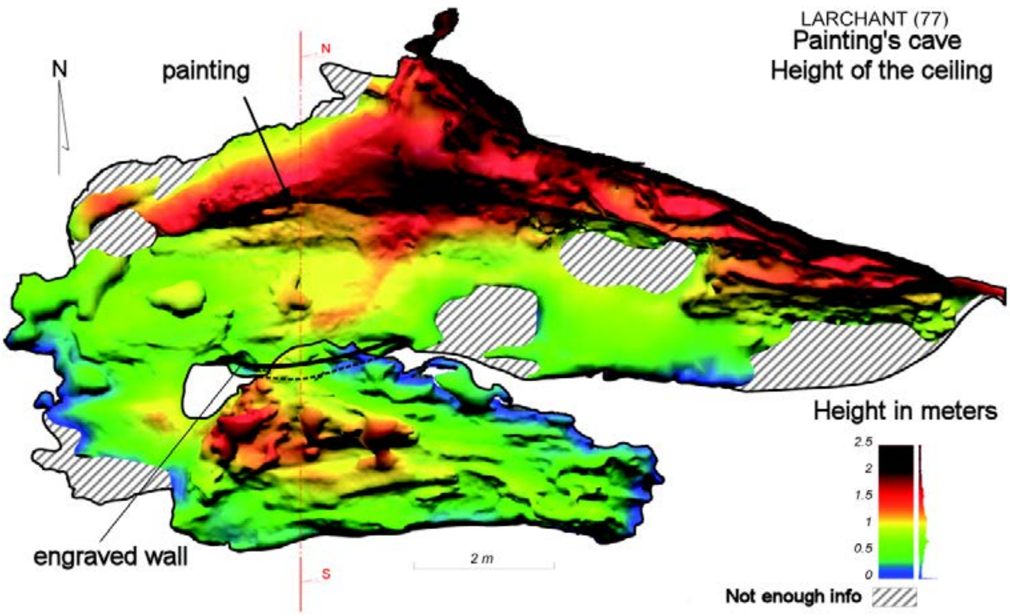


Figure 17: The cave's roof height.

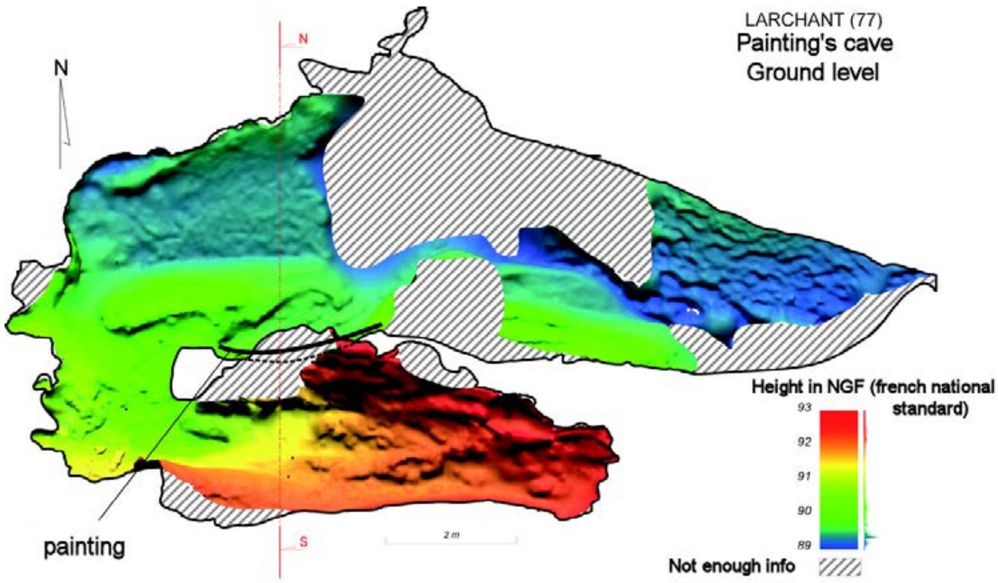


Figure 18: The cave's ground level.

For the cave, we can clearly see the difference of roof heights with an entrance over 1.50m high, the interior of the first chamber at 1m high, the passage between the two chambers at 0.50m high, the entrance of the second chamber at 1.90m high, the middle of the chamber at 1m/1.50m slowly declining towards very low heights.

If future excavations are done, it would be interesting to redo a photogrammetric survey while on the Mesolithic grounds to see if the roof and walls were hard to access and some support was needed in order to reach for the engraved places (Figure 18).

We were also able to obtain the ground levels of the cave, but we still lack some data as the software didn't find enough reference points to elaborate the 3D geometry.

We can however observe that the ground of the second chamber is higher than the first and with a pronounced sloping toward the entrance. It can obviously be seen on field but thanks to the 3D model it can now be quantified, and used to study the accessibility of the different areas of the cave. For example we can observe that the ground level difference over the whole cave is around 4 meters.

12.2 Cross section of the Painting's cave

The 3D models enable us to see the cave and the shelter in a way that would never be possible in reality. This cross section of the Painting's cave is the same as in Section 6.3 but can be easily perceived on a 2D diagram. You can clearly see the difference of size and height of the two chambers, and also that the second one is slightly over the first one (Figure 19).

13. Conclusion

13.1 Concerning the hardware and software

Taking the pictures is the critical moment of the whole operation: one must be very careful because this first step is a condition of good results for all the others after.

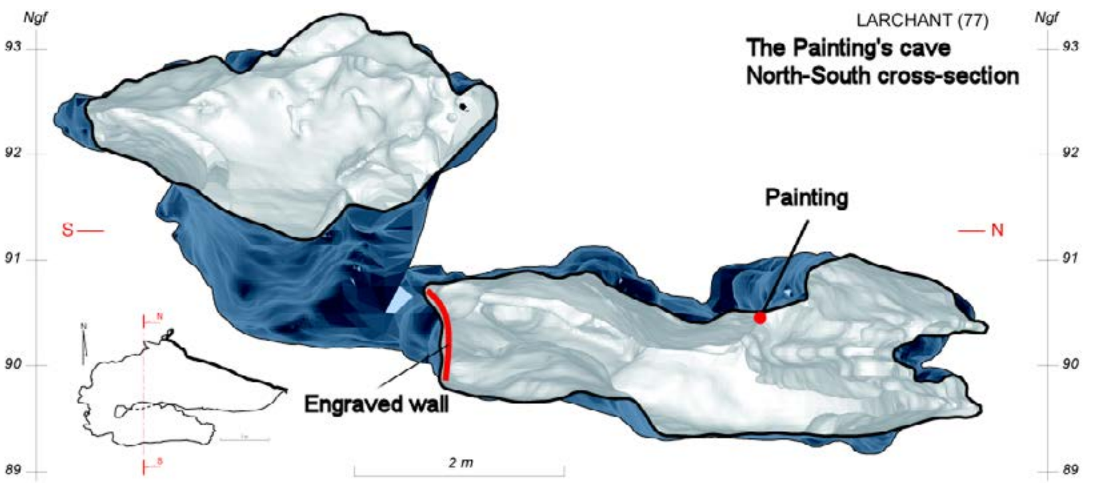


FIGURE 19: CROSS-SECTION OF THE PAINTING'S CAVE.

A very expensive and last generation equipment isn't always a requisite for scientific quality. We showed here that with an old equipment it's possible to obtain scientific 3D data usable for researches.

But it's also clear that an equipment a little more up to date would have allowed us to spend less time in calculations and manipulations, and more time on research and finding ways to get new data out of the 3D models.

13.2 Concerning the presentation method

The visual impact of the 3D .PDF isn't very good as it's very shiny, and it gives away an impression of plastic out of the models. In order to be able to integrate the models into the 3DPDF we had to simplify them a lot, so the data are degraded, which isn't a satisfying solution.

We want to explore new presentations methods, such as the new solution of the VC-Lab, 3DHOP, because unlike popular online methods such as SketchFab, it doesn't require to upload models in external repositories, giving the company access to restricted data, which is generally forbidden in almost every archaeological organisation in France.

We would also like to try out the solution of the QRCode to integrate our models with paperwork, because every French archaeological operation has its paper report.

13.3 Concerning the models themselves

Another interesting aspect of 3D data is that it's always available to further studies, without long explanations needed, and still with an objective aspect.

We hope that our models will be used in the future to do other studies than the ones we did ourselves. We think for example at the study of the light in the caves, which could explain some of the engravings' positions on the walls. We can also think of a metric calculation of the volume of the caves, which could be used for classification.

In any case, we would like to remind that 3D data is not an end, but it must be used for research and confronted with hand surveys as nothing can replace the eye of the expert being the field.

We hope the produced models will be used to further the studies of the caves, by ourselves, with the help of other researchers, or by others. In order to do so, we are thinking about a way of sharing them online with other researchers, and possibly with the public.

14. Future developments

- having all the detailed panels of the engravings in high resolution on the middle resolution shape of the whole Painting's cave
- integration of the 3D models into the mobile app of the ONF
- find a method to measure the engravings using a standardized procedure
- trying to establish a chronology between the engravings
- trying out our method on other caves and engravings in Fontainebleau
- trying out our method on other caves and engravings in France, particularly engravings made with other techniques
- creation of a GIS of the forest and integration of the 3D models
- digitization of the engraving tools and with the help of an expert trying to find comparison points between the tool and the engraving in order to try and match some of them, and to obtain information about the Human behind the engraving (left or right-handed, position during the work, etc.).

References

ABRÀMOFF, M. D., MAGALHÃES, P. J., and RAM, S. J. (2004). Image Processing with ImageJ, *Biophotonics international* 11(7), pp. 36-42.

BAUDET J-L. (1960). Les peintures préhistoriques de l'Île-de-France, *Bulletin de la Société préhistorique de France* 57(3-4). pp. 210-213.

BENARD A., and VALOIS L. (2013). L'abri orné des pentes du Marchais 1 à Larchant (Seine-et-Marne), *Bulletin du GERSAR,* 63, pp.

DELLEPIANE M., CORSINI M., CALLIERI M., and SCOPIGNO R. (2011). *3D models for cultural heritage : beyond plain visualization,* IEEE Computer Society.

DELLEPIANE M., SCOPIGNO R., DI BENEDETTO M., and LARUE, F. (2012). From the digitization of cultural artifacts to the web publishing of digital 3D collections: an automatic pipeline for knowledge sharing, *Journal of multimedia* 7, pp. 132-144.

DE LUCA L., and FLORENZANO M. (2003). Relevé-modélisation-représentation : un regard sur les approches basées sur la vision par ordinateur. In: Vergnieux R. and Delevoie C. (eds.)*. Actes du Colloque Virtual Retrospect 2003, Archéovision 1*, Editions Ausonius, Bordeaux. pp. 28-34. Avalaible from: http://archeovision.cnrs.fr/pdf/vr03_pdf/03_luca.pdf. [Accessed: 28th May 2014].

FORTE M. (2014). 3D archaeology, New perspectives and challenges : the example of Çatalhöyük, *Journal of Eastern Mediterranean archaeology and heritage studies* 2(1). pp. 1-29.

FRANKEN T., DELLEPIANE M., GANOVELLI F. (2005). Minimizing user intervention in registering 2D images to 3D models, *The Visual Computer Journal*. 21 (8-10), pp. 619-628.

FRISCHER B. (2005). New directions for cultural virtual reality : a global strategy for archiving, serving, and exhibiting 3D computer model of cultural heritage sites. In: Vergnieux R. et Delevoie C., (eds.).*Actes du Colloque Virtual Retrospect 2005, Archéovision 2,* Editions Ausonius, Bordeaux, pp. 168-175. Available from: http://archeovision.cnrs.fr/pdf/vr05_pdf/05_frischer.pdf. [Accessed: 28th May 2014].

GILIGNY F. (2011). Informatique et archéologie : une révolution tranquille?*,* In: P. Genet et A. Zorzi (ed.). *Les historiens et l'informatique. Un métier à réinventer*, Rome: Ecole Française de Rome, p. 189-198.

GUNN R. G., OGLEBY C. L., D. Lee and WHEAR R. L. (2010). A method to visually rationalise superimposed pigment motifs, *Rock Art Research* 27, (2). pp. 131-136.

HINOUT J. (1993). La grotte à la peinture à Larchant (Seine-et-Marne), lieu-dit les Dégoûtants à Ratard : le niveau mésolithique, la peinture, les gravures pariétales*, Préhistoire et Protohistoire en Champagne-Ardenne* 17. pp. 25-57

HINOUT J. (1998). Essai de synthèse à propos de l'art schématique mésolithique dans les massifs gréseux du Bassin parisien, *Bulletin de la Société préhistorique française* 95(4). pp. 505-524.

HINOUT J. (1998). Les pétroglyphes mésolithiques des massifs gréseux du Bassin parisien*, Revue archéologique de Picardie* 3-4. pp. 31-52.

JAMES, M. R. and ROBSON, S. (2012). Straightforward reconstruction of 3D surfaces and topography with a camera: Accuracy and geoscience application, *Journal of Geophysical Research* 117. F03017, DOI: 10.1029/2011JF002289 [On line].

ONF website, *Video La mémoire dans la pierre,* ONF, GERSAR et Parc naturel régional du Gâtinais français. [Online] http://www.onf.fr/lire_voir_ecouter/++oid++13d6/@@display_media_video.html [Accessed: 11th April 2014].

PESCARIN S. (2006). Open source in archeologia : nuove prospettive per la ricerca*, Archeologia e calcolatori* 17. pp. 137-155.

POIGNANT J. (1984). *L'art des cavernes : Atlas des grottes ornées paléolithiques françaises,* Ministère de la Culture, Imprimerie Nationale.

REMONDINO F. (2011). *Heritage recording and 3D modeling with photogrammetry and 3D sensing,* Remote Sensing in Natural and Cultural Heritage.

ROCHELEAU M. (2010). *La modélisation 3D comme méthode de recherche en sciences historiques,* 10e Colloque international étudiant du Département d'Histoire de l'Université de Laval, Québec.

ROGER D. *Le grès en forêt de Fontainebleau : expressions et exploitations*, [Online] http://randos-conviviales.over-blog.com/article-les-abris-de-carriers-de-fontainebleau-1ere-partie-100184475.html [Accessed: 28th May 2014].

VERGNIEUX R. (2012). *Partage d'expériences, La 3D au service des SHS : Pratiques, usages, enjeux,* La lettre de l'INSHS.

VERGNIEUX R. (2009). *Sauvegarder les données numériques 3D du patrimoine,* Virtual Retrospect.

WU C. (2013). *Towards linear-time incremental Structure from Motion,* 3DV Symposium.

Forum Romanum: A 3D Model for Self-Service Educational Purposes

Philippe FLEURY, Sophie MADELEINE, Nicolas LEFÈVRE

CIREVE / ERLIS Centre Interdisciplinaire de Réalité Virtuelle, Équipe de Recherche sur les Littératures, les Imaginaires et les Sociétés, Maison de la Recherche en Sciences Humaines de Caen, France
philippe.fleury@unicaen.fr, sophie.madeleine@unicaen.fr, nicolas.lefevre@unicaen.fr

Abstract

The CIREVE (Centre Interdisciplinaire de Réalité Virtuelle) of the University of Caen, Basse-Normandie France has developed an experimental self-service model for educational purposes. It is an interactive 3D model of the forum romanum built from the reconstruction of the entire city of ancient Rome and intended for museums, schools, and universities with educational documentation about each building.

This paper will explain

1. The context of the experiment: the reconstruction of the entire city of Rome as it was in the 4th century A.D for scientific and educational purposes. The interactive model is the work of a research team following scientific methods. Behind the virtual model, there is permanent access to a database with ancient texts, didactic videos, 2D pictures and a bibliography. The software used is 3DSMax for modelling and CryEngine for the interactivity.

2. The aim of the experiment: to provide educational actors with an interactive tool to discover different aspects of classical Antiquity. The interactive model must be used in self-service, i.e. without a scientific mediator. Although a scientific mediator can be essential for knowledge transmission, this interactive model can work alone transmitting knowledge through interactive play.

3. Constraints of an interactive model in self-service: an interactive model, at the service of the public at large, is different from an interactive model used by a researcher for scientific purposes or a model used by a scientific mediator. Firstly, it must be fitted with easily accessible documentation (which replaces dialogue with the scientific mediator). In the experimental model of the CIREVE, there are two types of documentation: automatic documentation which appears when the user is approaching a building and documentation 'on request' (similar to that in a guidebook). Secondly, it must be conceived so that the user cannot get lost in the streets, in public buildings, in dwellings. Virtual gates must be placed in the virtual model like barriers in a real site. In 3DStudio, these are 'collisions'. The software must be legally protected to avoid manipulation and alteration in line with intellectual property rights.

Keywords*:* Ancient Rome, Virtual Reality, Interactivity

Introduction

Being able to use virtual reality to make a virtual visit of an archaeological site – be it a reconstruction of an ancient monument or whole town – is something we consider commonplace nowadays because of a misuse of terminology. The fact is that most of the virtual or interactive visits we read about on the net, for example, do not use virtual reality at all. At the University of Caen we have been using true virtual reality since 2006 to invite the public into a digitally reconstructed 4th century Rome. The Forum Romanum is an experiment in opening up the experience to free autonomous visiting. The context and objectives of this experiment as well as its constraints will be the subject of the following article.

1. The experiment and its context

The Forum Romanum application is taken from the global restitution of Ancient Rome carried out by the 'Plan de Rome' group in the 'Literature, Imagination and Society Research Centre' at the University of Caen (ERLIS, N°4254) with the technical back-up of the Interdisciplinary Centre for Virtual Reality or 'Centre Interdisciplinaire de Réalité Virtuelle de l'Université de Caen' (CIREVE). To our knowledge, the virtual interactive model and the way it has been and will be used which is the fruit of this work within the ERLIS centre is unique (Fleury and Madeleine, 2008). It has made possible visits of the quasi-entirety of 4th century Rome AD for up to 200 people at a time. It offers a truly virtual interactive visit through the streets of the city guided by one of our researchers in a monoscopic or stereoscopic mode according to the numbers of visitors and allowing exploration of the facades and interiors of Rome.

The expressions 'virtual visit' or 'interactive visit' and even 'virtual interactive visit' are used in many contexts these days but rarely correspond to the full meaning of the concept of 'virtual reality'. The concept was formalized in the early 1990's but was little known; it is in fact the interaction with an artificial digitally created world which can be the reproduction of a real world, one which has disappeared or even an imaginary one. It implies two other concepts: 'immersion' which means passing 'through the looking-glass' and entering the image and also 'interaction' which means interacting in real time with the image to manipulate and transform it. Most applications available today on line and which carry the names 'virtual visit' or 'interactive visit' are in fact simply multimedia applications that provide interaction with pre-calculated animations. There is no real-time interaction with the digital model itself.

Contrary to popular belief, virtual reality has hardly penetrated the world of archaeology and we see little evidence of its use in research centres. The modelling of archaeological objects is common in archiving and the presentation of case studies but real-time interaction with these objects is seldom found, as indeed is public access to that interaction. Yet virtual reality is a technique of representation that can be termed 'ecological' to use the vocabulary of neuropsychology, that is, that it offers conditions of experimentation close to real life situations (Klinger, 2006). It goes beyond simple virtual restitution which is already accepted in terms of its representational potential: ability to multiply viewpoints infinitely by moving the 3D camera as well as multiplying hypotheses thanks to the dematerialisation of the medium, and making updates easier. Virtual reality offers, on top of all this, the pertinence of positioning both in time and space.

The virtual model of Ancient Rome created in Caen is therefore one that can be visited as part of an immersive experience in which the conditions of physical viewing and movement through space can be reproduced. Shortcuts allow time saving and enable passage from one building to another or between sections of large public buildings. Without these shortcuts, visiting the virtual city takes the same time as it would in real life. The presence of a human-sized avatar allows the visitor to get an immediate sense of the scale of the environment. The mediating academic guide can conduct the visit through the eyes of the avatar (in his or her position so from a subjective view-point) or follow the avatar a few metres behind and thus have a third person view point. The academic guide can also show the sources that allowed the restitution of virtual Rome by using links to multimedia databases.

Today, the virtual model can be used on two levels. The first is for academic research purposes in that

the virtual model can be used to visualise the city and test theories – two essential aspects of research. The permanent link to the corpus of textual, archaeological and iconographic sources allows users to measure the accuracy of every element of the cityscape. Situational elements as well as the real-life scale helps to test the validity of hypotheses concerning topographical and architectural restitution. It can also be used to test the efficiency of mechanical systems such as the sprinkling of perfumed water in entertainment buildings or the functioning of the curtain in theatres. The second level is a pedagogical one: guided visits of virtual Rome have the same teaching value in terms of the transmission of knowledge as do real visits of the modern city and both forms complement each other. The real visit, when it is possible, gives an idea of the present conservation of the city and enables the visitor to absorb its atmosphere and try to imagine its past. The virtual visit often makes the reading of the lost city easier: the ruins are no longer encrypted by the passing of nearly two thousand years but immediately visible to the eye as they once were.

The Forum Romanum project is an open visit that adapts itself perfectly to these two uses. It is designed for both researchers working in their offices or in libraries on particular areas of the city and for a broad general public as well as schools and universities (sources: Carettoni, 1960; Coarelli, 1994; Guidobaldi, 1998). There are two main advantages of being able to isolate and extract a particular sector of the model of Ancient Rome: the extract has all the benefits of the academic contextualisation of the global model and the reduced space allows information to be concentrated on a particular subject and material constraints are thus limited.

2. Objectives of the experiment

In all archaeological sites and museums there are generally two visit options: a guided visit and a free visit. Each has its advantages and disadvantages. The team Plan de Rome has been practising guided visits in virtual reality for over eight years now in two places: we have rooms equipped for thirty visitors next to a large 70 m2 plaster model of Rome of the same time period as that of the virtual model, and secondly a lecture hall for 200 people. In both locations visitors are given 3D glasses and the academic guide moves through the virtual model in real time thanks to a video-game controller. Attendance figures have been high in both cases and the places are booked months in advance for school parties; we average 1500 entries a year and these could be higher if we had more academic guides. Visits for the general public are available once a month and these average 160 visitors per month. The virtual visit is an enormous advantage in terms of pupils gaining knowledge of Roman civilisation: teachers who come with their classes report greater interest and better knowledge acquisition among pupils.

Our present aim in Forum Romanum is to see if the concept might work as an open access visit without an academic guide. We would target a broader public than the one currently coming to the guided visits which currently necessitate the use of a large screen in a fixed location. We hope to extend our field to include ever larger publics from very different walks of life. We have noticed an upsurge in interest in classical antiquity in the past few years in France and Europe but also as part of a global renewal of interest in the period and its ancient writings (the growth and nature of digital libraries bear witness to this). Paradoxically, this is happening at the very moment when knowledge of the humanities and the mastering of Latin and Greek in schools and universities are on the wane. But if Classics as a subject is less visible in the academic landscape, it is increasingly available today via alternative learning platforms (interactive museums, libraries and websites and smartphone applications to name but a few) which are the new spaces of engagement with the humanities.

Our objective is therefore to develop within an academic framework tools for interactive discovery and interpretation of ancient sources for use by a broad and heterogeneous public. We are studying the conditions and form in which these new spaces of urban experience and reading of ancient texts can usefully take their place alongside the academic tradition and co-exist without replacing or eclipsing it.

The Forum Romanum application offers the possibility of visiting 4th century Rome alone without an academic guide. It is a high level academic historic restitution of Rome led by an established University research group with a board of advisors and readers (Madeleine and Fleury, 2011). The mode of navigation is the same as that used in the global model with a video game controller or a keyboard with movement in real time. The user can run but can't fly or go through walls (a limitation which is also found in the more sophisticated video games which take place in historic cities like Venice, London or Paris). What is new in comparison with the guided visit is that the internaut will receive the information normally given by the guide by means of textual pop-ups. So on approaching a building, its name, its date of construction will be visible at the bottom of the screen as part of a trigger mechanism. Information panels or notice boards can be triggered by button A of the video game controller and the letter I key on touch screens and smartphones. They are the equivalent of labels or captions on objects on archaeological sites or in a museum but made available to the internaut in this case at the touch of a button or screen. If he or she wishes to go further he can touch on a table of contents and gain access to all the textual, visual and video documentation needed on a particular monument including history, functioning of the building and its civic role, academic sources and the bibliography used. The images are maps, plans, photographs of the monument in its present state today and other source material in visual form from museums, librairies. The vignettes appearing at the bottom of the screen can be made larger by the touch of a finger or movement of a mouse so that they can be worked with more easily. The videos treat the functioning of the more complex mechanical systems such as pistons and the stretching of velum.

A virtual visit must be adapted to the device available to the user. The personal computer is the easiest to configure at the moment and the one we are presently working with. We have tested it at the Vieux-la-Romaine Museum near Caen and on the stands of events like the Science Festival (Fête de la Science) with specially configured computers with controllers and mouse provided. We are also currently testing an application called 'Pantheon' which works on Android smartphones and pads. The model of Pantheon was realized in collaboration with the team of Bernard Frischer Rome Reborn. The application can be downloaded free on the 'Plan de Rome' site and a printable plan of the Pantheon is available. Two different technologies are used by the application: augmented reality allows vision of the outside of the building by moving the plan or just turning the device around physically like a camera-eye moving over the surface of the building. Virtual reality is also used to move through the interiors of buildings with the aide of the gyroscope on the device and using a virtual joystick. The target plan is not used in this case. Documents can be read (the font size is adjustable) or listened to. All images can be viewed in full screen.

The next step for us will be to finish the modelling of the forum not yet complete along with the progressive spatial extension of the application. On PCs it is possible to extend the visit to the whole city just as in the guided visit in Caen. For handheld device the application will progress sector by sector. The final stage of this 'self-service' visit will be the uploading of these applications onto the internet. But it is at this point that the technical constraints are the strongest as we will now see.

3. Constraints

The first constraints are methodological. Certain are inherent to any free visit, be it physical or virtual while others are particular to virtual visits. Just like the visit of a real site one has to determine what can be visited and what can't. In Forum Romanum and in the application with the academic guide it is possible to visit the whole area of Rome. When the 'self-service' application will be extended to the whole city such freedom will lose its pertinence since certain sectors still remain 'hypothetical'. This absolute freedom to wander anywhere would encumber the system and confuse and disorientate users unfamiliar with the topography of Rome.

Certain streets will thus be closed off virtually as well as on a physical site where certain streets are closed to the public. For the same reason not all buildings will be open to visit: only the most important monuments will open their doors as well as a few characteristic private buildings such as collective dwellings, private houses and shops. Just as with a real site or museum the position and nature of information must be determined and it is here that we see that the virtual visit makes things a lot easier since particular types of public can be targeted with special information. From the most specialised bibliographies destined for researchers to explanations of the simplest vocabulary for school children the project provides an impressive versatility. The level of the information they receive can be chosen by users at the moment they open the application.

Then come the constraints in terms of software. Screenplay in not developed in the global restitution of Ancient Rome because of our keenness to maintain reality. But for a free visit this will have to be upgraded since a minimal navigation interface must be used, different from that used by an academic guide. This will impact on the quality of access to documentation and its ergonomic and aesthetic dimensions must be defined. Such interfaces are more complicated to set up on portable devices such as handheld device without controller or mouse. For the navigation, particular attention must be paid to 'physicalization' for collision detection. A digital object is naturally penetrable and if we want a virtual visitor to walk on a floor, negotiate a staircase or avoid a statue or if we want an object to direct him or her to documentation then we need to make all these objects 'physical'. To do that, we need to create a physics proxy (in the technical vocabulary of CryEngine, the physics proxy is the geometry that is used for collision detection) around each object which users can reach. Quickest setup is to declare a virtual object as a physics proxy but checking for intersections on complex geometry is very expensive. Then we use as much as possible primitive geometry around an object, converting a complex visual geometry to a six faces collider for example.

This question is part of the world of video games from the 'serious game' to the 'edugame' which is closest to our concerns. The behaviour engine used for Rome is also a video game engine – Cryengine developed by Crytek. The great difference between a video game and an archeological restitution is that in the latter case it is absolutely necessary to respect historical truth and no shortcuts can be taken to make the production easier. Part of the creation of colliders has been automated but arches and staircases still pose a problem the intricacies of which we cannot explain here. Suffice to say that a great deal of manual work is needed. For a virtual visit no errors can remain which means hours and hours of testing, whereas in the guided visit the academic guide knows how to avoid the areas which are not yet solid.

The material constraints linked to the physical capacities of the engines must also be taken into account. If today video game techniques allow the development of virtual visits of huge areas with relatively good and sometimes exceptional detail, it is also true that the creation of a virtual visit of the entire expanse of Rome including details of architecture such as mouldings, ornaments, cornices, not to mention decorations such as statuary, paintings and mosaics (as well as access to a great number of interiors) remains a daunting task. If the extraction of a part of the model reduces the weight of the data to be uploaded onto the internet for example, the graphic treatment of each scene remains the same because software such as CryEngine gets rid of anything which is not visible to the user during his movement round the city. Since academic and pedagogical considerations require constant use of detail the processor and the graphics are highly solicited. The power of the materials needed for diverse user groups is growing all the time, but the fact remains that we have to work with the materials we presently have to hand.

Conclusion

Despite these constraints, our present experiments and research have given very positive results. The Forum Romanum application now has a new offshoot: an application in collaboration with the team of Robert Vergnieux in Bordeaux called Circus Maximus already in place in the Vieux-la-Romaine Museum and first set up for the occasion of the World Equestrian Games which will be held this year in Lower Normandy. We are also creating another application for our region this year to commemorate the 1944 Normandy landings; the 'Centre Interdisciplinaire de Réalité Virtuelle' at the University of Caen will set up an Android application on its site to visit the old University of Caen before the bombings (www.unicaen.fr/cireve). It will be a 'real' piece of virtual reality. The techniques developed for Ancient Rome have proved to be applicable to archaeological sites of all epochs.

References

CARETTONI, G., (1960). Excavations and Discoveries in the Forum Romanum and on the Palatine during the Last Fifty Years. *Journal of Roman Studies* 50, pp.192–203.

COARELLI, F., (1994). *Guide archéologique de Rome.* Hachette, Paris.

FLEURY, Ph. and MADELEINE, S. (2008). 'Problématique d'une restitution globale de la Rome antique. Une visite interactive avec accès dynamique aux sources anciennes'. In: R.Vergnieux et C. Delevoie (eds.). *Actes du Colloque Virtual Retrospect 2007, Archéovision 3*, Bordeaux: Ausonius. pp. 55-60. Available from: http://archeovision.cnrs.fr/pdf/vr07_pdf/07_Madeleine.pdf

KINGER, E., (2006). *Apports de la réalité virtuelle à la prise en charge de troubles cognitifs et comportementaux,* Thèse soutenue le 7 février 2006 à l'ENST (Télécom Paris Tech). URL: http://pastel.paristech.org/index.php?halsid=3f93vjd93rk50ggdur58nusro5&view_this_doc=tel-00161903&version=1

MADELEINE, S. and FLEURY, Ph. (2011). Le 'Plan de Rome' de Paul Bigot à l'Université de Caen et son double virtuel : de l'objet patrimonial à l'outil scientifique. *In Situ*. Les patrimoines de l'enseignement supérieur 17. URL: http://insitu.revues.org/840 ; DOI : 10.4000/insitu.840

GUIDOBALDI, M.P., (1998). *Le forum romain*. Milano, Electa.

Figure 1 : The forum romanum.

Figure 2 : A visit with an academic guide.

Figure 3 : A self service Visit.

Figure 4 : The application 'Pantheon'.

The Virtual Reconstruction of a Small Medieval Town: The Case of Briviesca (Spain)

Mario ALAGUERO,[1] Andres BUSTILLO,[2] Blanca GUINEA,[1] Lena IGLESIAS[1]

[1] Department of Historic and Geographic Sciences, University of Burgos, Spain
[2] Department of Civil Engineering, University of Burgos, Spain
malaguero@ubu.es, abustillo@ubu.es,
blanca.guinea@universidaddeburgos.es, liglesia@ubu.es

Abstract
A variety of methods are employed for the virtual reconstruction of many cultural heritage items, from small objects to entire cities. However, the time required for the development of these methods and the associated human resources are often limited. This study develops a virtual reconstruction method that optimizes the use of both resources, which in this case is applied to a small medieval town. The method is based on an estimation of modeling time that depends on two variables: final visual quality and the ultimate purpose of the reconstruction. In accordance with those variables, different 3D meshes and image textures sizes are considered and the results are presented to the final user to obtain feedback prior to the 3D modeling stage. It is validated in the virtual reconstruction of the town (villa) of Briviesca (Spain) as it was in the 15th c., with the final use of off-line rendering for the preparation of a TV documentary.

Keywords*:* 3D Reconstruction, Middle-Age Cities, Design Method, Project Management

Introduction

Central to the selection of the best technique to create any virtual reconstruction is the intended use to which the 3D reconstruction will be put. The most frequently mentioned final uses described in the literature refer to digital exhibitions of historic items for the public in museums (Gobbetti, Marton, 2004; Fabio *et al.*, 2010; Callieri *et al.*,2011; De Paolis, 2013) or through the Internet (Chow, Chan 2009; Wang *et al.*, 2008) as well as historical or archaeological research (Martin, 2010; Cheng, Jin, 2006; Lambers *et al.*, 2007; Lucet, 2009). Some, but not all of these final uses require real-time rendering of 3D models. For example, 3D caves or web engines of 3D environments will require real-time rendering unlike when rendering images and videos. Moreover, certain research-related uses will require very accurate 3D models, while others use low-quality 3D models, for instance, in the context of teaching. This broad range of uses and suitable qualities for 3D models makes it difficult to identify any single way of creating a virtual reconstruction of an historical or archaeological item.

If it is a question of the virtual reconstruction of a small medieval town, then CAD tools represent the only solution at present, because the vestiges of the streets and squares are all that now remain. However, a virtual reconstruction of this sort encounters important challenges: first the estimated requirements, in terms of development time and human resources, are often very limited; second the virtual reconstruction should be optimized in terms of both accuracy and visual quality for its final use. Although one immediate solution might be the use of existing software tools for large-scale automated modeling, such as those proposed for the virtual reconstruction of major urban centres (Dore, 2013), such a solution would imply excessive use of resources in the case that concerns us here of a medium sized Medieval village, due to its limited extension, complex orography, organic town planning, and the heterogeneity of almost all of its buildings. As an alternative, the present research proposes a virtual reconstruction method applicable to small towns that optimizes the modeling time depending on the two main variables that can affect it: final visual quality and the ultimate purpose of the reconstruction. Up until the present, research into CAD methodologies for virtual reconstruction has focused mainly on the relation between model size and computer capabilities for real-time rendering (Guidi *et al.*, 2005; Fabio *et al*., 2010). Different 3D meshes and image texture sizes are considered in relation to those two indicators. The results are then presented to the final user to obtain feedback prior to the start of the 3D modeling stage. Different 3D modeling and shading strategies are considered, depending on the singularity of each building and its structural elements. A further novelty of this CAD method is its emphasis on the generation of 3D models and image textures, which may be re-used in future virtual reconstructions with similar characteristics.

The structure of this work is as follows: Section 2 describes the CAD method for the virtual reconstruction of a medium-sized medieval town; Section 3 shows its application to a case study: the virtual reconstruction of the town of Briviesca (Spain) in the 15th c. and the main performance indicators of this reconstruction; Section 4 contains the conclusions of this work and identifies future lines of work.

2. Method

The first step of this method was to define the main objectives of the virtual heritage reconstruction project (VHRP), on the basis of three factors:

- Quality level selection: the quality level of the final project has a direct influence on the human and technical resources that are expended on the VHRP. Thus, in order to take decisions and control the project framework throughout the modeling process, a resource cost estimation has to be completed at the start of the project. The resource costs will increase exponentially in accordance with the desired quality level of the model.
- Display mode selection: this will influence the project resource costs in accordance with the format (image, video, interactive virtual reality...) and the type of device (personal computers, interactive booths, large-format projections...). Specific requirements will then be defined to develop the project.
- Creating the script: the script will focus attention on the highest quality models, to which more resources will be dedicated in the modeling process. It will help save time on the modeling of structures that will only appear briefly in the background or that will not appear at all in the final display mode.

Apart from these influential there is a further external one: the availability of accurate sources of historical and archaeological documentation.

The depth of historical and archaeological research is very important in the following steps of the project. Insufficient and inaccurate historical information are a few of the many reasons why the modeling process will require constant supervision by historians and heritage experts, who will respond to doubts and take informed decisions whenever historical information is doubtful or missing. Accordingly, this factor has a lot of influence on resource costs, which should be considered and quantified if possible.

The next step of the method is to draw up a forecast of human, technical and time resources that will be dedicated to the project in accordance with its main objectives. This forecast will be done by means of experimentation to establish the capabilities of the modelers (as each forecast will greatly depend on their working experience), who will complete different modeling tests with 3D models to establish a timeframe. Subsequently, these temporal results will be extrapolated to calculate the resources cost forecast for the whole project. Usually, this is a provisional forecast that may be modified in subsequent phases, so a best case and worst case estimate should be prepared to adjust the calculation of the resource costs to the project.

Once the resource cost forecast is defined, the next step is to compare it with the capabilities of the team and the objectives of the project. An appraisal of all these factors provides a clear picture of project feasibility.

The following step is to begin the modeling and texturing tasks. The design team should take photographs of the original textures of the walls and roofs of the buildings and other similar structures and textures. These photographs serve two ends: as a visual reference for the modeling phase and to be processed as textures for use in the texturing phase. At this point, the hues of all the textures used in the project are harmonized and tiled textures are generated, for use in large-scale models. The textures are then classified with meta tags, scaled to similar file sizes and sorted on a database to facilitate the texturing process and for use in future projects. The modeling team will decide whether sufficient photographs have been taken for the whole project and the historians and heritage experts should also review all of the images.

In the modeling phase, the first step is to draw artistic sketches of the model considering previous research, the data collected from an archaeological investigation if any, and the photographs taken from both the original structure and other reference buildings.

This artistic sketch has to be done carefully, to avoid any inaccuracies or biased views that the modeling team may have. It should be thoroughly reviewed by the historians and heritage experts to provide interpretive solutions in cases where the historical documentation is insufficient. If there is little information on something and the sketch has a high level of guess, historians experts and modelers must choose a way to show what is actually according to the documentation and what is notional (Mehta, 2001).

Once accepted and reviewed, the sketches may be used to begin the modeling phase. The first step will be the completion of a small-scale 3D prototype of the model that the historians and heritage experts will then review. Their approval is needed and it should be verified whether the quality level satisfies the objectives defined in the project description and whether sufficient human, technical and time resource costs have been forecasted. Hence, the need at this stage to measure the resource costs, so as to extrapolate them to the whole project.

Once the results of this assessment process are acceptable, the whole modeling phase can begin. Developed for a virtual reconstruction project of a small town, this method divides the modeling structure into the following four sections:

– The modeling of singular buildings: this phase cannot be systematized. In each particular case, there will be singular buildings with so many different characteristics that specific modeling will be needed in each case.
– Modeling of ordinary buildings: some ordinary buildings will be modeled to represent the urban

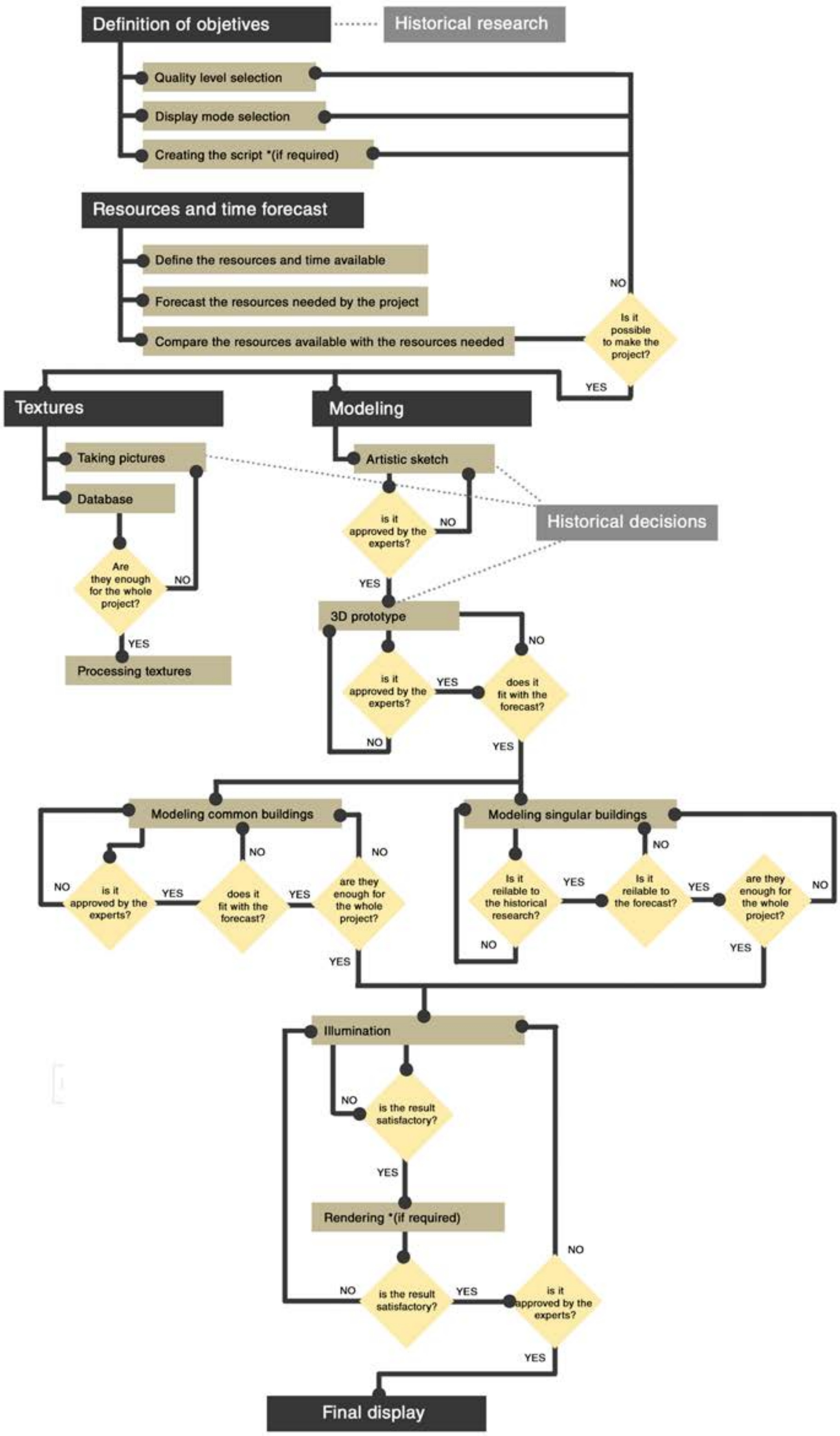


FIGURE 1: METHOD FLOWCHART.

layout of the town. These models will be of the buildings that are most frequently found in the original town. In this phase, several different buildings, with certain modifications, will be placed around the town, with little chance of the viewer perceiving duplicate buildings. The number of public buildings that are required for each case is determined by the characteristics of the urban centre and by defining the following objectives.

- Modeling of urban accessories: this provides realism and contributes to the creation of the identity of the buildings. These accessories may in certain cases be used to highlight relevant areas and to singularize certain buildings.

- Joining models and environmental modeling: this process cannot be systematized and will vary in each particular case. The resources invested in environmental modeling depend on the script and the project objectives; so the importance of this task will in each case vary.

If the historians and heritage experts approve the results of this modeling phase, the project will move on to the lighting and rendering phase.

More technical resources are needed in this phase than in any other to improve the quality level. Better lighting and better quality rendering require high-performance computers and more time from the modeling team. In addition, this phase depends directly on the objectives that are defined, especially the display mode and the quality level that is selected. Based on these factors and the resource costs that are forecast for the whole project this phase can be estimated, but it can not be systematized. Total rendering time, for example, can be estimated, by adjusting the values provided by the rendering engine and by completing timed tests, to determine how much computing time would be required to process the final results (Figure 1).

3. Case study: the city of Briviesca in the 15th century

The method described in the previous section was applied to a case study, to establish whether it actually optimizes the modeling process: the virtual reconstruction of the town (villa) of Briviesca (Spain), in the 15th c. This decision was taken, because the Municipal Council of Briviesca was prepared to participate in the project. It was willing to assume a role in decision-making, it agreed to review the quality level of the results and would provide assistance with historical research. Furthermore, this case was a very interesting experiment with which to test our method, because its orthogonal urban layout is both unique and rare.

Briviesca is located in northern Spain in an area that was formerly the old Kingdom of Castile. For several centuries preceding the demise of the Muslim civilization in the Iberian Peninsula, the various Christian kingdoms in this area lived through an age of political and social upheaval as a consequence of constant rivalry and warfare between each other and with the Moorish kingdoms to the south. Many small towns and cities built walls in an effort to protect their inhabitants against attack. Walled towns also offered protection to the inhabitants of the numerous small villages scattered across the north of Spain that lived off agriculture and livestock. These improvements in the defenses of many towns meant that the inhabitants of outlying villages could seek refuge from raids. A small population could flee its village very quickly at the sound of its church bells that would also warn other villages of the impending threat. The region of Briviesca was no stranger to this situation: documentary evidence exists of several small hamlets little more than a few hundred meters apart the main town of which was Virovesca, located on a hill near the modern-day town of Briviesca. Virovesca had been populated since pre-Roman times. From around the 13th c., the dominance of the Christian kingdoms in the Kingdom of Castile returned the region to a period of stability and prosperity during which time the bourgeoisie and the nobility began to exercise greater influence. In the early 14th c., Doña Blanca de Portugal purchased an inheritance of Juana Gómez de Manzanedo, consisting of the cluster of hamlets around Virovesca. Doña Blanca is recognized as the founder of Briviesca and was actively involved in planning a new villa on a plain, bringing together farmers from other villages in the region to build a larger and more influential settlement. She also decided to layout the town in a very organized manner, imitating the classic urban Roman layout of a grid (a very unusual example for Medieval urbanism in the Iberian Peninsula) (Figure 2) and ordered the construction of defensive buildings and other infrastructures to attract new inhabitants.

In this study, we performed 3D modeling of the urban environment of the new Briviesca around a century after its foundation, showing the evolution of the town from its foundation up until its consolidation in the 16th c.

In line with our method, we contacted history and heritage experts to draw up a descriptive profile of the town in the 15th c., its specific characteristics and its singular buildings, before beginning to sketch and taking pictures. The worst problem of this phase was the dearth of documentation. Nowadays, only the urban layout and some walls of the two temples in the village are preserved in Briviesca as they were in the 15th c. Thus, in this particular case, a constant feed of advice from the historians and heritage experts on each of the virtual reconstruction stages was necessary.

We had to select the desired quality level in the model to define the objectives of the project, the display mode and the narrative structure. We decided that in view of the potential impact and importance of testing our method with a complex case, it would be appropriate to develop a good quality level rather than a high-resolution model. We planned a scale to select the level of quality from 1 to 5, where 1 is a very low quality model (volumetric, 12-60 polygons/building) and 5 a high-resolution model (10K-80K polygons/building). In this case, we selected a quality level of 4/5 (1000-3000 polygons/building).

Our team consisted of two modelers with experience of virtual heritage reconstruction projects and two computing systems (Intel(R) Core(TM) i7-2630QM CPU 2.00 GHz, 8GB RAM). We tested the modeling and rendering speeds to establish a forecast for the project duration and resource costs. As planned in the definition of the method, a 3D model was performed at two different quality levels to quantify the three variables (Figure 3). Modeling time, the number of polygons, the weight of the images and the model level of quality were all measured. The 3D low-detailed-model was completed in under 16 minutes. It had

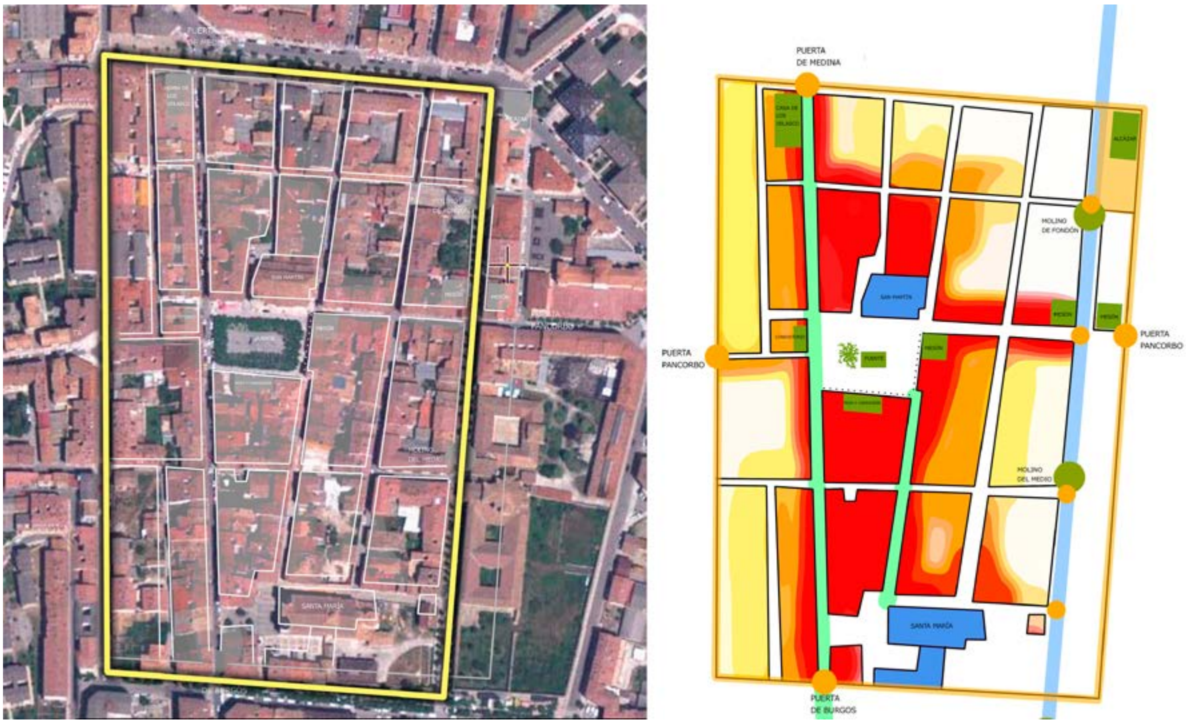


Figure 2: Briviesca nowadays and its urban layout on the 15th c. (Ibarra, 1998).

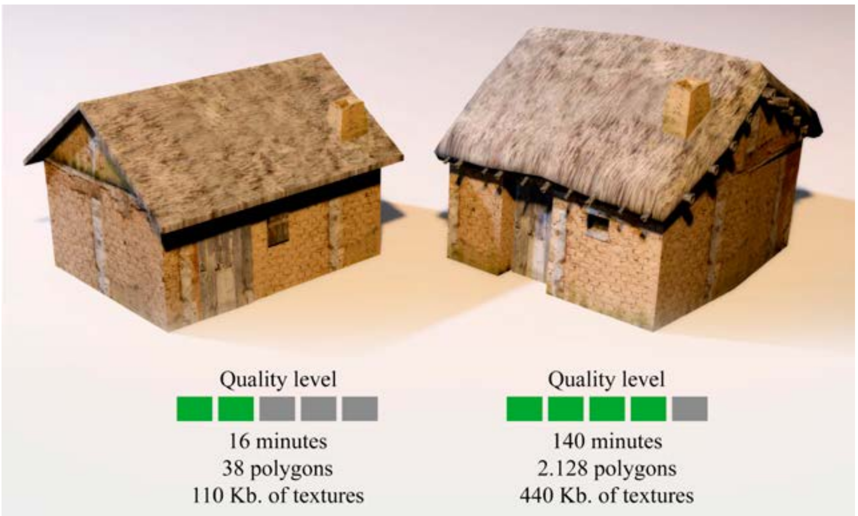


Figure 3: Modeling and rendering test with two differents quality levels.

38 polygons and 110 Kb of textures and scored 2 on a scale of 5 for visual appeal. The version with most visual appeal was completed in 2 hours and 20 minutes. It had 2,128 polygons and 440 Kb of textures and scored 4 on the same scale of visual appeal. A lower level of quality would not meet the objective expectations and a higher one would be extremely heavy (the 3D model of a building would have as many polygons as the entire village modeled at a level 2 on the scale of visual appeal).

Historical-archaeological research cite records of around 300 ordinary buildings in Briviesca at the turn of the 15th c. (including houses, stables, sheep pens and sheds) and over 20 singular buildings.

There were also some urban accessories that had to be modeled (markers, fences, plants, wells ...). These additional items led us to estimate that the total number of polygons in the whole project would be between 1,000 and 2,000 times that of a single house at a degree of visual appeal of 2/5. Compared with the modeling experiment, the degree of visual appeal of 2/5 would mean that the whole village would have between 38,000 and 76,000 polygons and a visual appeal of 4/5 would mean from 3 to 6.5 million polygons. The forecast pointed out that the two modelers of our work-group would dedicate from 300 to 500 hours to develop the project and would need several additional computers to render the entire video in the time frame requested by the customer.

The next step was taking pictures to create the database to apply textures to the 3D model and to provide a source of inspiration for the modeling process. We visited several small villages with similar materials and construction styles which may have been found in the Briviesca of the 15th c. (Figure 4).

The pictures had to be properly prepared for their use. All the textures were edited with an image editing software to eliminate flaws and anachronisms (modern materials, fixtures, broken tiles ...) and to ensure uniform sizes and hues (Figure 5). This task is hard but facilitates the subsequent process of modeling and rendering and helps to correct differences between textures in the whole project. Afterwards, the textures were categorized in a database to facilitate their easy retrieval and future use. This systematic process helped the modeling process and the generation of buildings with different appearances, simply by changing the textures in accordance with the model stored on the database.

The sketching and picture-taking phases are usually very useful in cases where there is little information. The historians and heritage experts used them to take specific decisions and guesses rather than elaborating a theoretical description all by themselves. For example, doubts arose over the morphology of the houses, the distribution of residential areas, the form of singular buildings and other interpretations. The advice of the historians and heritage experts contributed to an informed and accurate sketch of the town that was very useful in the modeling phase.

The next step in our method was the creation of a prototype model. This phase served to prevent future errors and to check whether the resource cost forecast was correct. A building was modeled, textured, lit and rendered, with information from historical research and from the photographs. The resource costs test gave positive results and adapted to our expectations. One of the modelers dedicated 92 minutes to model the building with 1,149 polygons. In addition, the prototype was approved by the historians and heritage experts except for some minor changes in the structure that were taken into account for modeling the rest of the town. The model was moreover presented to the Municipal Council of Briviesca, which considered the visual quality satisfactory. However, it should be stressed that this small-scale test can not guarantee the reliability of the overall project.

Both the modeling and texturing tasks depended on each other. Singular buildings, common buildings and accessory buildings were modeled step by step to fill the city layout, before completing the environment modeling. Forty different houses were modeled (Figure 6) with several variations (2-5), to complete the modeling of the ordinary buildings. So, we had 160 different kinds of buildings, changing only a few models and textural details (Figure 6). This helped us to achieve heterogeneity in the urban landscape, a realistic global view of which is important. So, modeling such a large number of houses was not excessively time consuming, at least in comparison with the modeling of singular buildings.

The modeling of singular buildings can not be systematized at all. The distinguishing features and characteristics of each case mean that the models are so different that it is hardly possible to carry out any joint action (Figure 7). Greater efficiency in this phase may only be achieved by constantly adjusting the modeling time to the forecast, because the modeling times of complex structures are very often longer than expected, which can change the forecast timing. In addition, this phase presented a further problem: only rarely were textures found that could be likened to those of Briviesca in the 15th c. Numerous existing monuments were repaired with modern techniques or have been subjected to intensive cleaning processes that were non-existent in the 15th c. We resorted to image editing software to add damage, stains, mould and dampness to the original textures taken from real buildings.

These phases were constantly supervised by the historians and heritage experts to pick out anachronisms, errors in the architectural design and the historical bias of modelers.

Modeling elements and the details of accessories was an intricate part of the process that required complex models and had a large number of polygons in the final model. One of the most representative cases is the orchards that were very common to every village, which required a lengthy modeling time and a lot of polygons. This was a problem that had to be solved, so an experiment was performed to check the number of polygons needed for each solution. Several types of plants with various levels of detail were modeled (Figure 8). A single plant had a low number of polygons (10 to 40), but this method became untenable when multiplied by the 300 or so plants in an orchard. A solution was proposed: to create fallow crops in the biggest parcels, locate orchards in small fields and use particle distribution in large format crops, which reduced the overall processing time.

Once modeled, all the buildings and accessories were grouped along the main roads of the town. The main

Figure 4: Picture of a house from Yugueros (Spain) a sketch made about its structure and the final 3D model with the original textures applied.

FIGURE 5: ORIGINAL PICTURE, PROCESSED TEXTURE AND APPLIED TEXTURE TO A 3D MODEL.

FIGURE 6: SEVERAL COMMON BUILDINGS MODELS AND THREE VARIATIONS OF THE SAME BUILDING.

buildings (when their approximate location was known) were placed in position and the other buildings were then joined to create the households. After this, we began to draw the blocks of the town, consisting of courtyard houses, vacant lots (the historians and heritage experts stressed that not all the plots would be built up a century later) and crops. The area of the town in which each building was situated influenced its characteristics. The houses were bigger and built with better materials on the main roads, the central square and around the churches. In the back streets and around the periphery, the houses tended to be far poorer, as they were built of cheaper materials in larger blocks (because there was less land demand in the poorest areas of the town). The larger crops were placed on the margins of an artificial canal running through the town from north to south on the eastern side.

The historians and heritage experts and representatives from the Municipality of Briviesca were invited to review the whole model of the town, before beginning the lighting and the rendering stages. There were several issues to solve, following which a number of lighting tests were performed (Figure 9) before the rendering process began.

Once rendered, all the images and videos underwent a final editing and post-production phase, which brought the project to a conclusion (Figure 10).

Conclusions

This study has proposed a method for the virtual reconstruction of small towns. It optimizes two requirements: development time and the human resources required for the task. The method is

specially thought through for research groups with limited human resources and limited time, because it generates ready-to-use materials (3D Models, textures…) for subsequent projects. It is based on the estimation of modeling time, depending on the two main variables that can affect it: the final visual quality (evaluated on a scale from 1 to 5) and the ultimate purpose of the reconstruction (from off-line rendering applications of high visual quality to real-time rendering applications with limited render hardware). In accordance with those variables, different 3D meshes and image textures sizes were considered and the results were presented to the final user to obtain feedback prior to the 3D modeling stage. Two different 3D modeling techniques were proposed depending on the singularity of each building and its structural elements. The problems that relate to the complexity of a large 3D project (final rendering time, etc) are surmounted by estimating final 3D mesh sizes at the start of the modeling stage.

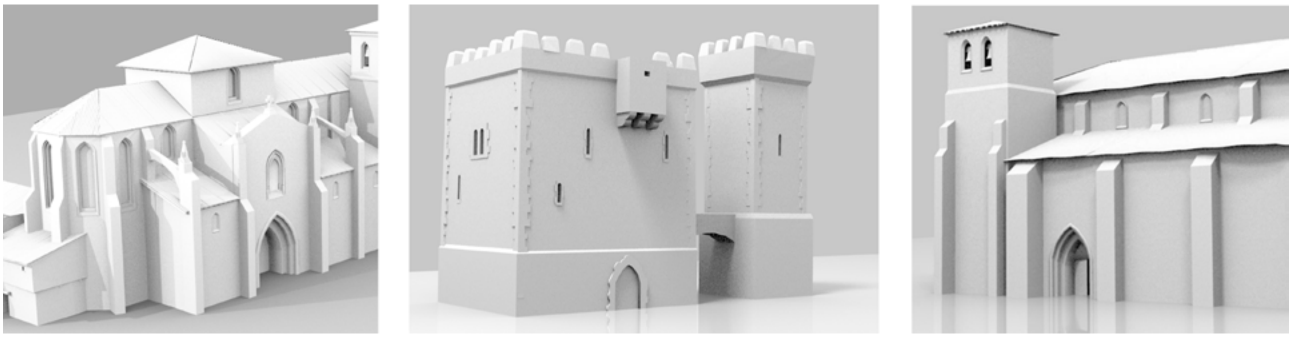


FIGURE 7: EXAMPLES OF SINGULAR BUILDINGS: CHURCH OF SANTA MARÍA, FORTIFICATION AND CHURCH OF SAN MARTÍN.

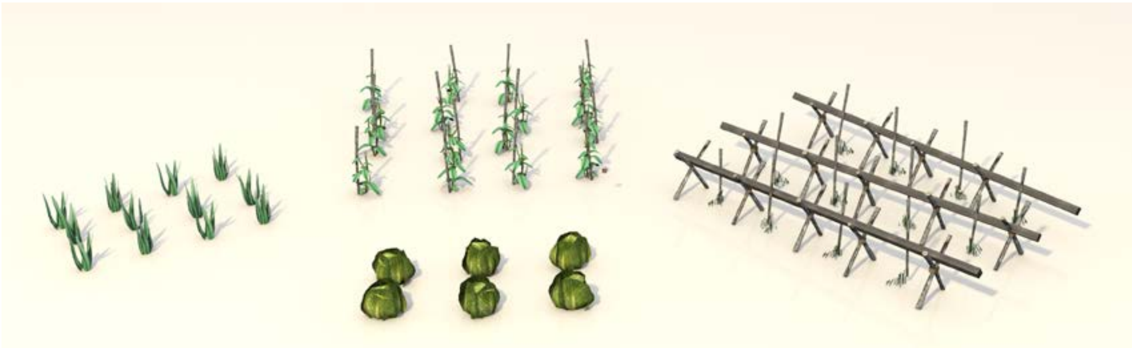


FIGURE 8: SEVERAL TYPES OF PLANTS.

The method has been validated through a virtual reconstruction of the villa of Briviesca (Spain) as it was in the 15th c., with the final use of off-line rendering for the construction of a TV documentary. The continuous support of experts in medieval architecture and history in Spain was requested to define such details as the urban layout of the city and the structure of ordinary housing and singular buildings, such as places of worship and stately homes that are still easily recognizable in the city. This medieval village includes around 300 standard houses and 20 singular buildings. A level of 4 on the pre-defined scale of 5 was selected for this virtual reconstruction, because no real time rendering was included in the final product. 3D models of standard houses were done at two quality levels; level 2 required 16 minutes 3D model work including 38 faces and the image textures had a size of 110 Kb, while level 4 required 9 times more 3D model work, including 80 times more faces and image textures that were 4 times larger. Including secondary elements, the city was estimated to include between 38,000 and 76,000 polygons at a quality level of 2, and between 3 and 6.5 million with a quality level of 4. Therefore, different quality levels were considered for singular buildings and standard housing depending on their presence in the story telling and the final video, allowing us to conclude the virtual reconstruction on time with a limited use of human resources and rendering capabilities.

Further research is planned on a combination of this method and the automated generation of housing to reduce the human resources needed for such projects, but retaining the final visual quality and scientific rigor of the virtual reconstruction. Therefore, large scale automated modeling techniques will be tested and, depending on the end purpose of the reconstruction and its required visual quality, their suitability will be determined for integration in the proposed methodology.

Acknowledgments

This research was partially supported by the Spanish Program "Impulso de la Industria de Contenidos Digitales desde las Universidades" of the Spanish Ministry of Industry, Tourism and Commerce and the definition of the "Centro de Creación Digital (DigiUBU)" of the University of Burgos and the Municipality of Briviesca (Spain). The authors would especially like to thank María Jesus Olivares for her kind help in providing historical information on the city of Briviesca.

References

CALLIERI, M., CHICA, A., DELLEPIAE, M., BESORA, I., CORSINI, M., MOYÉ, S.J., RANZUGLIA, G., SCOPIGNO, R., BRUNET, P. (2011). Multiscale acquisition and presentation of very large artifacts: The case of portalada. *Journal of Computing and Cultural Heritage* 3 (4). Article number 14. doi>10.1145/1957825.1957827

CHENG, X.J., JIN, W. (2006). Study on reverse engineering of historical architecture based on 3D laser scanner. *Journal of Physics: Conference Series* 48. pp. 843–849.

CHOW, K., CHAN, K. L. (2009). Reconstruction of photorealistic 3D model of ceramic artefacts for

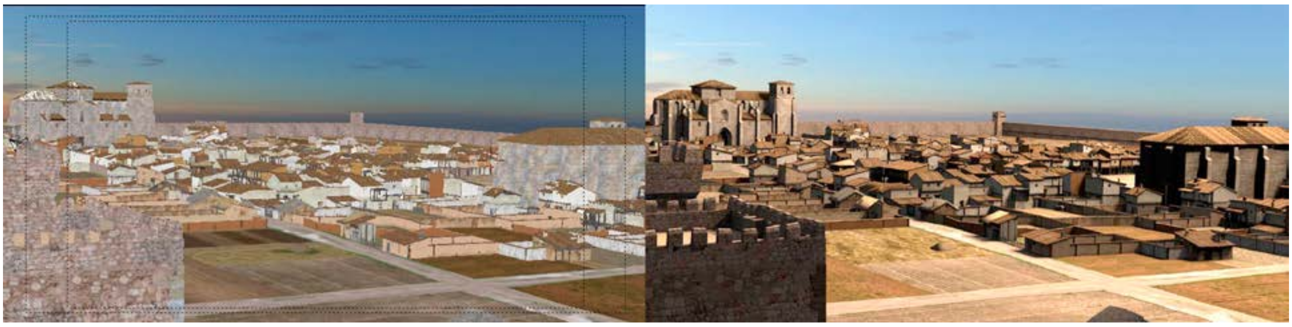


Figure 9: Test of lighting of the whole model.

Figure 10: Post-production effects applied to final render.

interactive virtual exhibition. *Journal of Cultural Heritage* 10 (2). pp. 161–173.

DE PAOLIS, L.T. (2013). Walking in a virtual town to understand and learning about the life in the middle ages. Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics). 7971 LNCS (PART 1). pp. 632-645.

DORE, C. MURPHY, M. (2013). Semi-automatic techniques for as-built BIM façade modeling of historic buildings, In: *Proceedings of the Digital Heritage 2013 - Federating the 19th Int'l VSMM, 10th Eurographics GCH. Marseille (France); 28 October-1 November 2013*, Volume 1, Article number 6743786, pp. 473-480.

FABIO, B., BRUNO, S., DE SENSI, G., LUCHI, M., MANCUSO, S., MUZZUPAPPA, M. (2010). From 3D reconstruction to virtual reality: A complete method for digital archaeological exhibition. *Journal of Cultural Heritage* 11(1). pp. 42-49.

GOBETTI, E., MARTON, F. (2004). Layered point clouds: a simple and efficient multiresolution structure for distributing and rendering gigantic point-sampled models. *Computer and Graphics* 28. pp. 815–826.

GUIDI, G., FRISCHER, B., DE SIMONE, M. (2005). Virtualizing Ancient Rome: 3D acquisition and modelling of a large plaster-of-Paris model of imperial Rome. Videometrics VIII. 5665. pp. 119-133.

IBARRA, J.L., ORTEGA, A.I. (1998). La villa de Briviesca en la Baja Edad Media: datos y reflexiones para su estudio. *Boletín de la Institución Fernán González* 217. pp. 321-352.

LAMBERS, K., EISENBEISS, H., SAUERBIER, M., et al. (2007). Combining photogrammetry and laser scanning for the recording and modelling of the late intermediate period site of Pinchango Alto, Palpa, Peru. *Journal of Archaelogical Science* 34(10). pp. 1702–1712.

LUCET, G. (2009). Virtual Reality: A Knowledge Tool for Cultural Heritage. *Computer vision and computer graphics: theory and applications* 24. pp. 1-10.

MARTIN, P., LLAMAS, J., MELERO, A., GOMEZ-GARCIA, J. and ZALAMA, E., (2010). A practical approach to making accurate 3D layouts of interesting cultural heritage sites through digital models. *Journal of Cultural Heritage* 11(1). pp. 1-9.

MEHTA, M. (2001). Virtual Reality applications in the field of architectural reconstructions. In: *Proceedings of the Seventh International Conference on Virtual Systems and Multimedia. Vienna (Austria) 9-12 September 2009,* IEEE Computer Society. pp. 183-190.

WANG, J., TIAN, F., SEAH, H.S. (2008). Sketch-up in the Virtual World. In: *Proceedings of the 2008 International Conference On Cyberworlds.* Hangzhou (China), 22-24 September 2008, IEEE Computer Society. pp. 109-116.

25 Years of Experience in Virtual Reconstructions - Research Projects, Status Quo of Current Research and Visions for the Future

Mieke PFARR-HARFST

Technische Universität Darmstadt, Germany
pfarr@ika.tu-darmstadt.de

Abstract

More than 20 years of research in the area of virtual reconstruction at the Department of Information and Communication Technology in Architecture (IKA) at the TU Darmstadt have produced virtual reconstructions in a wide variety of no longer existent buildings and settlement structures. In all research projects the emphasis is on the scholarly validity of the ultimate results. This precondition influences the operational methodology as well as the insightful dealing with available source materials. A constant discourse with leading scholars plays a major role in the reconstruction process. Thus virtual reconstructions are always designed as international and interdisciplinary research projects; their sustainability is guaranteed through publication at various platforms. This paper is a retrospective of previous projects, their conceptual content, methodology and results as well as a critical examination of future challenges in this area. The status quo of virtual reconstructions for imparting knowledge and research as such will be delineated.

Keywords*:* Virtual reconstruction, Cultural Heritage, Research Tool, Challenges, Visions

1. Introduction

The intention of this paper is to give a general, theoretical survey of virtual reconstructions. Background for this overview are the research activities of Department IKA and its twenty years of experience in the area of scientifically based three-dimensional reconstructions.

Therefore this paper is organised into the following parts:

- Background, guidelines and principles of the research
- Survey of past research projects
- Two projects presented and discussed in detail
- Potentials of virtual three-dimensional models for the transfer of knowledge and research in and of itself
- Challenges and opportunities for future research landscapes
- Summary completing

2. Background, guidelines and principles of the research

The Department IKA is a part of the faculty of Architecture at the TU Darmstadt and was founded by Prof. Manfred Koob 25 years ago. During these years of research experience following guidelines or rather principles concerning the methods of research have been developed.

- Definition of model structures and levels at the beginning of a reconstruction process and their classification into macrostructures and microstructures
- The binding determination of nomenclature for the 3-D model as well as for the scholarly sources
- Classification and structuring of sources and the creation of source catalogues for all elements
- Creation of reference models within the defined model structures to enable editing.
- Archiving of the given status of all important models
- Documentation of the reconstruction process in an uniform system
- Creation of a definitive filing system suitable for the project.

These principles are the basis for all of the research projects of the Department IKA.

3. Project Cross-Section

The research projects carried out at the Department IKA can be divided into different categories which reflect the diversity of the research topics. Many reconstructed buildings or estates are of significance to constructional and architectural history and therefore can be assigned to cultural heritage. All virtual reconstructions created at the Department IKA are essentially scientifically founded and therefore contribute towards the investigation and decryption of the constructed cultural heritage. In the following, the individual categories with concrete project examples are listed.

- category 1: churches:
  Religious buildings and their construction and design history are the focus of this category, e.g. the virtual reconstruction of the St. Peter's Basilica in Rome (Figure1).
- category 2: palaces

The virtual reconstruction of palaces and castles, their interior fittings and constructional historic development are closely linked and viewed in terms of their political context and cultural history. This category includes buildings such as the Berlin Palace or the Dresden Palace (Figure 2).

- category 3: 'Cultures of the World'
  This category focuses on the mediation of knowledge concerning other cultures and is therefore linked with an educational mandate for a university environment and for the general public. The reconstructions of the imperial tombs of Xian (Kunst-und Ausstellungshalle Der Bundesrepublik Deutschland, 2006), Angkor Vat or the Moscow Kremlin (Figure 3) provide examples.

- *category 4: 'Immaterial Memory und New Forms of Memory',*
  The virtual reconstruction of destroyed synagogues in Germany and Europe (Grellert, 2007) and the handling of new memory forms associated with this provide the focus of this category. The new media and their potentials are integrated into the memory process and therefore create new forms of memory and knowledge mediation. (Figure 4).

- *category 5: 'Urban Structures and their Development'*
  Urban settlement structures and their investigation are always an important topic in architecture and discussions on building cultural heritage. In projects such as the reconstruction of the history of Venice (Figure 5), the surroundings, constructional structures, economic context and geographical features were investigated.

- category 6: monasteries
  The monasteries of the middle ages are mostly no longer intact, and yet nevertheless still have great influence on the architecture and settlement structures of today and on cultural development. By means of three-dimensionality, the inner correlations of the monasteries and also the interaction with their surroundings were examined. This includes projects such as the reconstruction of the Lorsch Monastery (Kloster Lorsch) or the Plan of St. Gall (Figures 6-8).

- category 7: 'Icons in Architecture'
  In the history of architecture there are several key buildings which have sustainably characterized construction, building constructions or the concept of architecture. The focus of the projects in this category is therefore dependent on the significance of these buildings. Therefore, in the case of the reconstruction project for the Crystal Palace (Figure 9), the future-orientated steel-glass construction was transferred into three dimensions and the impressive spatial effect of the building was visualized.

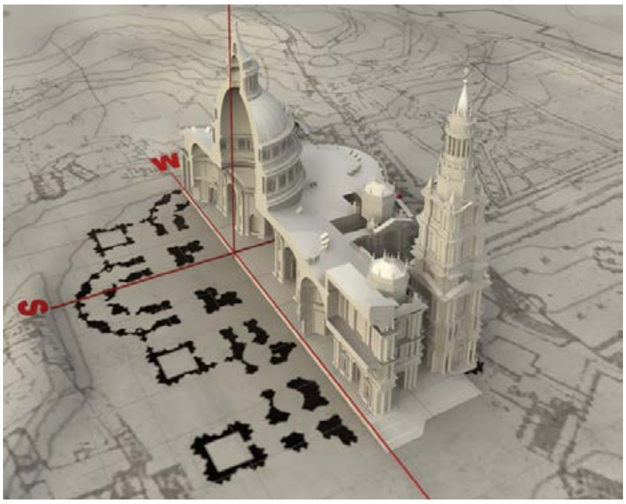


Figure 1: Digital Reconstruction of St. Peter in Rome. © Dept. IKA, TU Darmstadt, 2006.

Figure 2: Digital Reconstruction of Dresden Castle in Dresden – Giants Hall.
© Dept. IKA, TU Darmstadt, 2011.

Figure 3: Digital Reconstruction of Kremlin.
© Dept. IKA, TU Darmstadt, 2004.

- category 8: royal and imperial palaces, castles and fortresses
  This category encompasses the visualization of royal palaces and fortresses of the past to facilitate the investigation of the power centers and space-defining structures. The Pfalzanlage (palace complex) in Frankfurt can be named as an example.

Figure 4: Synagogues.
© Dept. IKA, TU Darmstadt, 2006.

Figure 5: Digital Reconstruction of Venice
© Dept. IKA, TU Darmstadt, 2006.

Figure 6: Benedictine Monastery Plan - Basilica.
© Dept. IKA, TU Darmstadt, 2006.

Figure 7: Benedictine Monastery Plan - Enclosure.
© Dept. IKA, TU Darmstadt, 2006.

Figure 8: Benedictine Monastery Plan - Guesthouse.
© Dept. IKA, TU Darmstadt, 2006.

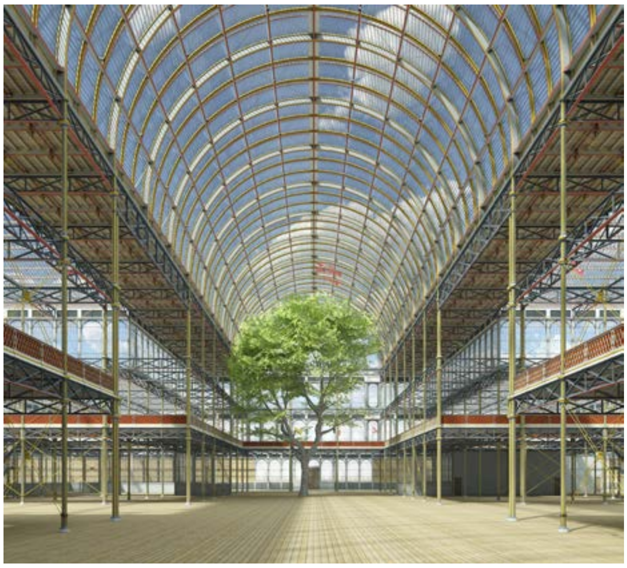


Figure 9: Digital Reconstruction of Crystal Palace London.
© Dept. IKA, TU Darmstadt, 2013.

4. Detail project descriptions

Following this overview of the project categories and the complexity of the research projects, which already provide an initial impression regarding the potential of such three-dimensional models for the field of research, two projects are then described in detail. In the course of this, the methodology, the sources and the related obtaining of knowledge are addressed in particular.

4.1. Project: Benedictine Monastery Plan

The goal of the EU project 'Benedictine Monastery Plan' was an examination of two hypotheses regarding the interpretation of the 9th century plan using a virtual three-dimensional model. The project partners were Benedictine monasteries throughout Europe and the Technical University of Vienna. Dr. Baraba Schedl, University of Vienna, was the expert, who confirmed the knowledge about the monastery in the 9th century and the existing scientific publications about the plan (HECHT 1983, JACOBSEN

1992, REINHARDT 1952). In particular the two hypotheses were as following:

- Hypothesis 1: The St. Gall Monastery plan can be understood as an ideal and organisational chart that demonstrates the organisational structures, spatial relationships, building hierarchies and furnishings of an ideal monastery.
- Hypothesis 2: The drawn red lines can be understood as clear space, an architectural concept describing actual space between two spatially limiting construction elements such as walls.

If one combines both hypotheses with each other and considers the plan to be an organisational diagram and the red lines as clear space independent of the thickness of the walls, the following conclusion results:

The monastery plan of St. Gall is a sort of construction manual for Benedictine monasteries all over Europe. The distribution of the buildings, their construction and the appearance of the entire monastery complex are thus dependent on the financial resources of the patron, the topographical situation and the climatic conditions as well as on the available construction materials. These assumptions were to be examined three-dimensionally using the potentials of digital media on the basis of three fictional monasteries.

The parameters for these three monasteries in culturally and climatically differing regions of Europe, northern, central and southern Europe were defined as follows:

Northern European monastery complexes :

- steep roof pitches due to snow load and rainwater
- roofing: thatch or straw
- walls: natural stone, thickness approx. 80 cm

Central European monastery complexes

- walls: wood, thickness approx. 20 cm
- walls of buildings positioned higher in the hierarchy of stone, thickness approx. 50 cm
- walls plastered with clay, lime or brick dust
- medium roof pitch
- roofing: straw or wooden shingles for important buildings

Southern European monastery complexes

- flat roof pitches as are common in Italy or Spain
- roofing: tiles
- walls: bricks, thickness approx. 30 cm

With the hypothesis described here all areas of the monastery plan were first incorporated into a three-dimensional outline using the red lines. For this the interior spatial measurements for each building were determined in advance. The indications here were the furnishings shown in the plan, the number of users as well as the basis of the construction history and design with respect to traffic areas and required space. The basic measurements of the furnishings were reconstructed on the basis of comparable objects from previous research. After the structures and sizes of the three monasteries had been roughly determined the following buildings of the northern model were more closely examined.

- Basilica (Fig. 07)
- enclosure with refectory, dormitory, calefactory as well as cloister (Fig. 08)
- abbot's residence
- guesthouse (Fig. 09)
- school
- mill
- bakery

The model structures and the model levels were defined according to the described guidelines:

- monastery complexes (northern, central and southern Europe): macrostructure II (scale 1:500)
- detailed individual buildings (exterior): macrostructure I (scale 1:200)
- detailed individual buildings (interior): microstructure I (scale 1:100)
- facilities and furnishings: microstructure II

The individual buildings were linked to the models of the monastery complexes since changes, for example in the construction, affected the structural composition of the complex as a whole.

The guesthouse serves as an example for clarification of the methodology presented above.

It can be discerned in the plan that the guesthouse has a main room surrounded by four small chambers. The furnishings of the rooms are also shown. In the main room benches and tables are grouped around an open fireplace and the chambers are furnished with a bed, a chest and a stove. A stable for the guests' animals is located at the rear of the building.

With the given height and size of the building the construction of the wooden roof support structure and the associated laws of span and construction had to be taken into consideration. In addition to the dimensions this was the key.

The result of the this research project is, that the monastery plan of St. Gall can indeed be read and understood as a generalised plan and that the interpretation of the lines offers flexible adaptation to local conditions and cultural as well as construction alternatives.

The research methodology was based on the detailed analysis and the consolidation of the research results, up to date at the time, concerning the plan and the monastery buildings as well as construction and materials in the Middle Ages. The transfer into three-dimensionality occurred step-by-step. Regular verification phases were carried out by the participating researcher, Dr. Barbara Schedl.

In addition an extensive library of furnishings for mediaeval monasteries was created which was then archived and used for further projects.

The EU project clearly demonstrates the potentials of the 3D models for research, the verification of hypotheses in three-dimensionality and the fusion and generation of new insights concerning construction, furnishings and spatial relationships.

4.2. Project: Construction History of the Florentine Cathedral

The most recent project 'Construction History of the Florentine Cathedral' will be used to discuss further potentials and methodologies. The two main points of emphasis of the project were:

- the construction history of the cathedral and its precursor buildings
- the planning history of the cathedral with plans for the façade, the planning of the dome construction and the cathedral as a whole.

In this project as well different model levels at varying scales were created:

- the city of Florence: macrostructure III
- the cathedral in its urban context: macrostructure II
- the cathedral and its precursor buildings (exterior): macrostructure I
- the cathedral and its precursor buildings (interior): microstructure I
- façade plans: microstructure I
- construction of the dome, construction machines: microstructure II

The sources and basis of this project included all of the previously published scholarly works, plans, surveying reports (Belli, 1994; Braundfels, 1964; Fanelli, 2004; Gurrieri, 2010; Rocchi, 1996; Rocchi, 2006). Besides this, the excavation results as well as hitherto unpublished information collected in the archives of the participating projectpartners were the scientific base, too. The projectpartners were: Art and Exhibition Hall of the Federal Republic of Germany, Bonn; Kunsthistorisches Institut in Florenz, Florence; Museo dell'Opera del Duomo; Florence; Università degli Studi di Firenze, Florence.

With respect to the precursor structure, Santa Reparata, research is still in its inception and its actual appearance is still unclear. In this case first the secured excavation results became the basis of the 3D model. This was then compared with the predominant theories concerning Santa Reparata. An evaluation together with the experts is still pending for this part and for that reason we chose a somewhat vague manner of presentation (Figure 10).

Furthermore, the original design of Arnolfo was transposed onto the present-day dome in order to visualise the spatial-structural and proportional relationship between the structures (Figure 11).

Particular emphasis was placed on the planning theory and the construction of the dome. The newest research results were applied three-dimensionally as hypotheses. It was possible to verify these hypotheses by means of simulations and exact structural application (Figure 12).

Figure 10: Cathedral of Florence – St. Reparata. © Dept. IKA, TU Darmstadt, 2013.

Figure 11: Cathedral of Florence – Concept of Arnolfo. © Dept. IKA, TU Darmstadt, 2013.

In addition, the research model served as a fusion of all sources that have been available up to now concerning the construction history of the cathedral. The history of the façade plans is an example (Opera Di Santa Maria Del Fiore Di Firenze, 1987). The original plans of the facades collected in the archives of the Museo dell'Opera del Duomo, Florence, were transferred into three-dimensionality and then correlated directly with the construction of the cathedral. A three-dimensional archive of the designs resulted that can be made available to research (Fig. 13a-13d). The same applies to the entire building model that exists in various levels.

For the first time a deliberately cautious type of presentation during the preparation of the model for imparting knowledge were chosen in order to distinguish clearly the elements of uncertainty. The result of this interdisciplinary and international research has been a part of the exhibition 'Florenz!' at the Art and Exhibition Hall of the Federal Republic of Germany in Bonn. Within the context of the exhibition on the city of Florence a 3D stereo projection showed a broad public the cathedral as an important historical monument. The digital data set is available for further research.

5. Potentials as Virtual Scientific Models

The following hypothesis arose from the projects described in the overview: virtual reconstructions offer potentials that are predestined as a three-dimensional tool for the investigation of our cultural heritage (Koob, 1995; Hermon, 2007; Hermon, 2012). In the following these

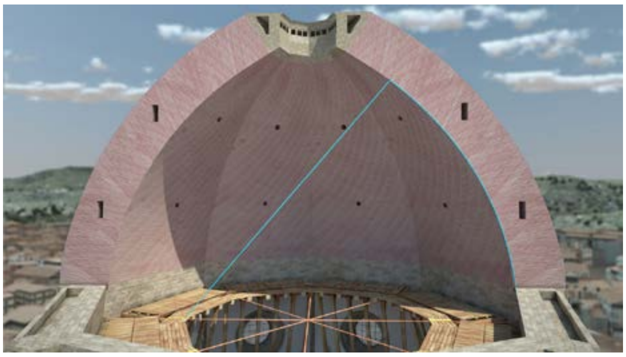


Figure 12: Cathedral of Florence – Dome. © Dept. IKA, TU Darmstadt, 2013.

potentials will be described in detail and the relationship to our current research at Department IKA will be shown.

A clear and binding definition of terms for the use of virtual three-dimensional models or reconstructions as research tools does not exist up until now, and most terms have become inadequate to express the underlying potential of three-dimensional models (Münster, 2010). Because the potentials and the application possibilities associated with them are many and diverse, the term 'Virtual Scientific Models' is introduced in this paper and is defined as follows:

Virtual Scientific Models, as a three-dimensional medium, offer a promising opportunity for processing and passing on knowledge.

In the broadest sense, these Virtual Scientific Models are to be assigned to the virtual research environment of E-Humanities.

The potentials are generated from the characteristics of virtual three-dimensional building models such as:

- digitality
- three-dimensionality
- visual language
- virtuality

The following potentials arise:

- diversity of output forms
- visualisation of complex interrelationships of content and space
- presentation of alternative forms
- consolidation, generation, verification and imparting of knowledge.

By means of these potentials it is possible to apply these models in all phases of a research project involving cultural heritage (Fig. 14).

5.1. Potential 'Diversity of Output Forms':

Here the basis is primarily the characteristic of digitality that enables the production of a data set in every technically possible way. Among these are renderings, simulations, film sequences, web-based applications or haptic models. All of these output formats have been used to impart knowledge for many years.

The following are various opportunities for research:

- open research model, a virtual research space
- scientific model, a virtual three-dimensional archive system

Figure 13a: Cathedral of Florence – Façade of Antonelli. © Dept. IKA, TU Darmstadt, 2013.

Figure 13b: Cathedral of Florence – Façade of DeFabris. © Dept. IKA, TU Darmstadt, 2013.

Figure 13c: Cathedral of Florence – Façade of Graziani. © Dept. IKA, TU Darmstadt, 2013.

Figure 13d: Cathedral of Florence – Façade of Saslow. © Dept. IKA, TU Darmstadt, 2013.

- adaptable to various scientific problems, for example, simulation of events or stills as cross-fades.

5.2 Potential 'Clarification of Complex Interrelationships of Content and/or Space',

This potential of Virtual Scientific Models is based in three-dimensionality, virtuality and visual language. Hence the following possibilities arise for the scholarly context:

- understanding of structures that are not visible, determining spatial context
- localisation of finds
- verification of joining principles, structural details
- interactive stage connected to the databank, similar to an open research model
- representation of various construction phases.

5.3. Potential 'Presentation of Alternative Forms'.

The focus of research is always on knowledge through digitality and virtuality and the resulting possibility of working in immaterial space. Applied to the scholarly process this potential means the following:

- verification and refutation of scientific hypotheses
- juxtaposition of various approaches to solutions as three-dimensional hypotheses or sketches.
- basis for scholarly discourse
- new insights into partial questions
- detection of contradictions.

5.4. Potential 'Consolidation, Generation, Verification and Imparting of Knowledge'

Scientifically well-grounded virtual reconstructions rely on a basis of knowledge that is generated from sources of differing type, origin and author. This results in probably most important group of potentials that focus on knowledge in virtual reconstructions. They concern the consolidation, fusion, verification and generation of knowledge. For the research process that means:

- hitherto existing research results are assembled
- a virtual scientific model emerges as a method of scholarly research
- a basis for further research is created
- new knowledge can be generated and passed on.

One could say that the bottom line of this list of potentials is that digital cultural heritage is created through the digitalisation of cultural heritage in three-dimensional space.

6. Challenges and Research Topics

Exactly at this point challenges arise that on the one hand are opportunities, but on the other offer risks as well. It is indispensible that they be dealt with by research. At the moment there are no uniform structures, standards and rules nor 'best practice' principles. Indeed individual networks, both national and international, have been set up but their interlinking is still in the beginning stages.

The Department IKA is active in research and networking with respect to the establishment of a new research methodology. The basis for research can be summarised with the keywords: fusion, generation, verification, securing and imparting of knowledge (PFARR 2011*)*. The focus of research is always on knowledge in and about virtual reconstructions and the accompanying scholarship.

The current research questions will be listed in the following part:

- research topic 1: Virtual Sketches
  Understandig virtual building models as three-dimensional hypotheses it is necessary to find an adequate language for their presentation. This means photographic reality as opposed to a sketch. What does a sketch-like representation mean in the computer?

- research topic 2: Presentation Form
  This topic includes the question of presentation form with respect to the knowledge basis. How do I present verified sources as opposed to hypotheses?

Research topic 1 and 2 can be assigned to the area of generation and verification of knowledge.

- research topic 3: Open Research Model
  The use of three-dimensional building models as virtual research environments or open research models. Here the primary question arises as to what extent these can be practicably integrated into the scholarly process.

- research topic 4: Standard for a virtual research environment
  At this point the vision of virtual research space goes a step further. The questions in this connection are:
  - technical requirements and challenges for such a principle
  - practical feasibility (simulator sickness)
  - networking of the locations
  - localising information and the knowledge base
  - additional benefits of such a system in the scholarly process.

Research topic 3 and 4 can be considered to be in the area of fusion and generation with respect to a future scientific method.

- research topic 5: Best practise
  Verification of scholarship, the introduction of binding standards and the assurance of 'best practise', for example, a three-dimensional archive with documentation. This topic has to do with the verification and sustainability of knowledge.

- research topic 6: VR and AR Applications
  This is concerned with imparting of knowledge. The focus is primarily on the question of how expedient the use of VR and AR applications is.

All of these research topics regard virtual reconstructions as digital cultural heritage, as a bearer of knowledge. We must deal circumspectly with this heritage and define strategies for working with it and we must guarantee its sustainability.

Bibliography

BAUMEIER, S. (2010). *Entwicklung eines internetbasierten Werkzeuges zur Unterstützung von Forschungs- und Kommunikationsprozessen in der Siedlungsforschung. Untersuchung der Potentiale digitaler 3D-Visualisierungen für die Rekonstruktion von baulich-räumlichen Siedlungsstrukturen*. Available from: http://hss.ulb.uni-bonn.de/2010/2075/2075.htm. [01.07.2014]

BELLI, F., BENVENUTI PAPI, A., DALLA NEGRA, R.; FABBRI; P., GURRIERI; F. (1994). *La Cattetrale die Santa Maria del Fiore. Cassa di Risparino di Firenze. (Vol. I, II).* Florenz, Giunti Ed.

BRAUNFELS, W. (1964). Der Dom von Florenz. Olten: Urs Graf Verlag.

FANELLI, G., FANELLI, M. (2004). *Die Kuppel Brunelleschis. Geschichte und Zukunft eines großen Bauwerks.* Florenz: Mandragora S.R.L.

FRINGS, M. (2001). *Der Modelle Tugend. CAD und die neuen Räume der Kunstgeschichte*. Weimar: Verlag und Datenbank für Geisteswissenschaften.

GRELLERT, M. (2007). *Immaterielle Zeugnisse. Synagogen in Deutschland. Potentiale digitaler Technologien für das Erinnern zerstörter Architektur*. Bielefeld: transcript Verlag.

GUIDI, G. (2008). *3D Modeling of Large and Complex Site Using Multi sensor Integration and Mulit-resolution Data.* Eurographics Digital Library. Available from: http://diglib.eg.org. [14.05.2014]

GURRIERI, F. (2010). *La Cupola del Brunelleschi Firenze*. Firenze: Accademiadell´Iris.

HECHT, K. (1983). *Der St. Galler Klosterplan*. Sigmaringen: Thorbecke-Verlag.

HEINE, A., HENZE, K. and RIEDEL, F. (eds) (2006) .*Modellieren, Strukturieren, Präsentieren. Von Handaufmass bis High Tech II. Informationssysteme in der historischen Bauforschung.* Mainz: Verlag Philipp von Zabern.

HERMON, S. (2007). *3D Modelling as a Scientific Research Tool in Archaeology*. Academia.edu. Available from: http://www.academia.edu/1226903/3D_Modelling_as_a_Scientific_Research_Tool_in_Archaeology. [01.06.2014]

HERMON, S. (2012) *Scientific Method, Chaîne Opératoire and Visualization - 3D Modeling as a Research Tool in Archaeology*. Available from: https://www.academia.edu/1227042/Scientific_Method_Chaine_Operatoire_and_Visualization_-_3D_Modelling_as_a_Research_Tool_in_Archaeology [01.06.2014]

JACOBSEN, W. (1992). *Der Klosterplan von St. Gallen und die karolingische Architektur: Entwicklung und Wandel von Form und Bedeutung im fränkischen Kirchenbau zwischen 751 und 840*. Berlin: Dt. Verlag für Kunstwissenschaften.

KOOB, M. (1995). *Architectura Virtualis. Projekt Weltkulturerbe*. Lecture notes, Technische Universität Darmstadt.

KUNST-UND AUSSTELLUNGSHALLE DER BUNDESREPUBLIK DEUTSCHLAND (eds) (2006). *Xi´an. Kaiserliche Macht im Jenseits*. Mainz: Verlag Philipp von Zabern.

MÜNSTER, S. (2010). *Entstehungs- und Verwendungskontexte von 3D-CAD.Modellen in den Geschichtswissenschaften.* In MEISSNER, K. *Virtual Enterprises, Communities & Social Networks*. Dresden: TUDpress,

OPERA DI SANTA MARIA DEL FIORE Di FIRENZE (eds.) (1987). *Due Granduchi tre Re e una facciata. Opera di Santa Maria del Fiore di Firenze*. Firenze: Alinea Ed.

PFARR, M. (2010). *Dokumentationssystem für digitale Rekonstruktionen am Beispiel der Grabanlage Zhaoling, Provinz Shaanxi, China. Darmstadt*. Technische Universität Darmstadt: tuprints. Available from: http://tuprints.ulb.tu-darmstadt.de/2302/ [01.07.2014]

PFARR, M. (2011*) Documentation system for digital reconstructions. Reference to the Mausoleum of the Tang-Dynasty at Zhaoling, in Shaanxi Province, China*. In: BÖRNER,W., UHLIRZ, S. (2011) *Proceedings of the 16th International Conference on Cultural Heritage and New Technologies (CHNT 16, 2011)* Vienn; Museen der Stadt Wien. Available from: http://www.chnt.at/proceedings-of-the-16th-international-conference-on-cultural-heritage-and-new-technologies/ [14.01.2013]

REINHARDT, H. (1952). *Der St. Galler Klosterplan: [Facsimile-Wiedergabe in acht Farben]; mit einem Ausschnitt aus dem Klosterplan in sechs Farben und zehn Bildzeichnungen von Hans Bühler*. St. Gallen: Fehr.

RIEDEL, A., HEINE, K. and HENZE, F. (eds) (2011). *Modellieren, Strukturieren, Präsentieren. Von Handaufmass bis High Tech III. 3D in der historischen Bauforschung*. Mainz: Verlag Philipp von Zabern.

ROCCHI, G. (1996). *S. Maria del Fiore: rilievi, documenti, indagini strumentali, interpretazioni ; piazza, battistero, campanile*. Firenze

ROCCHI, G. (2006). *S. Maria del Fiore: teorie e storie dell'archeologia e del restauro nella città delle fabbriche arnolfiane*. Firenze: Alinea.

The Archaeological Excavation of the Early Neolithic Site of Portonovo as a Case Study for Testing a 3D Documentation Pipeline

Cecilia Conati BARBARO,[1] Eva Savina MALINVERNI,[2] Anna Nora TASSETTI,[2] Chiara LA MARCA[1]

[1]Dipartimento di Scienze dell'Antichità, Sapienza Università di Roma (Italy)
[2]DICEA, Università Politecnica delle Marche, Ancona (Italy)
cecilia.conati@uniroma1.it, chiara.lamarca@uniroma1.it, e.s.malinverni@univpm.it,n.tassetti@univpm.it

Abstract

An archaeological investigation carried out by INRAP (Institut national de recherches archéologiques préventives) rThe paper outlines how different state-of-the-art survey workflows can be applied to map the Early Neolithic site of Portonovo - Fosso Fontanaccia (Ancona, Italy), establishing a straightforward, fast and often low-cost workflow for excavation recording. Different survey experiences are carried out to map the five domed ovens (VI millennium BC) excavated during the 2013 field campaign, ranging from digital photogrammetric to terrestrial laser scanner data acquisition and from open source to commercial processing. The option of quick, well-tested and often low cost/open source survey pipelines makes the research experience a case-study highlighting new approaches that can be integrated in the general excavation methodology and additional interesting features such as model/data reusability. The produced photorealistic 3D models together with all the other digital data are integrated inside a GIS environment satisfying the need to manage on situ the documentation of on going excavations.

Keywords : archaeological excavation, Portonovo, 3D documentation, GIS

1. Introduction

Stratigraphic archaeological excavations demand high-resolution documentation techniques for 3D single-surface documentation because only within a destructive process each single deposit can be uncovered, identified, documented and interpreted. Documentation must be systematic, objective and independent from the given interpretation and should permit the re-examination of the work even after a long span of time. This 3D recording is typically accomplished using total stations but other techniques such as photogrammetry and 3D scanning are nowadays standing out because of their performance and affordable deliverables. These 3D methodologies for Cultural Heritage 3D techniques should be taken into consideration because overcoming previous constraints in terms of processing time and cost (Cowley, 2011). Moreover it is documented how they can be really affordable establishing a straightforward, fast and often low-cost workflow for excavation recording and Cultural Heritage in general (d'Annibale, 2014). More important they allow the generation of very realistic 3D results (in terms of geometric and radiometric accuracy) that can be used for many other purposes: digital preservation and conservation, cross-comparisons, monitoring of shape and colours, simulation of aging and deterioration, virtual reality/computer graphics applications (Bruno *et al.*, 2010), 3D repositories and catalogues (Fangi *et al.*, 2013), web-based geographic systems (Manferdini *et al.*, 2010), computer-aided restoration, multimedia museum exhibitions and so on (Barcelo *et al.*, 2000, Cowley, 2011).

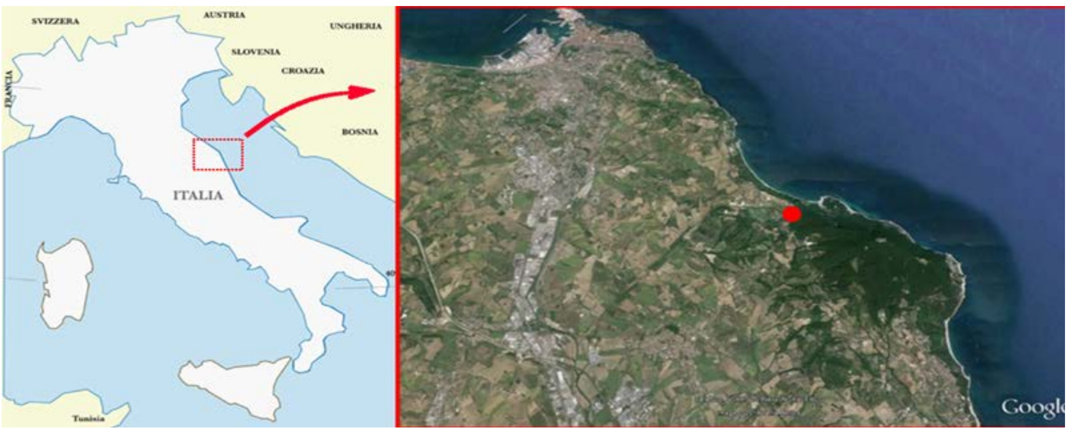


FIGURE 1: THE POSITION OF THE ARCHAEOLOGICAL AREA IN PORTONOVO.

Figure 2: The archaeological area in Portonovo.

2. Excavated Area

The site of Portonovo dates back to the ancient Neolithic of the middle-Adriatic Italian peninsula (Fig.1), and the structures that have been found represent a unique evidence in Italy and in the Mediterranean area. The archaeological site is located on a south-facing slope (Fig. 2), along the right bank of the river Fontanaccia, at an altitude of about 120 m asl.

Soundings were conducted in 2006 by the Soprintendenza per i Beni Archeologici delle Marche (a peripheral organ of the Ministry of Culture with the institutional task of protecting, conserving and valorising the architectural and landscape heritage in the Marche Region). Since September 2011 systematic excavation campaigns has been undertaken by the Sapienza University of Rome (2011, 2012 and 2013) over an area of about 300 m2 (Conati Barbaro 2013).

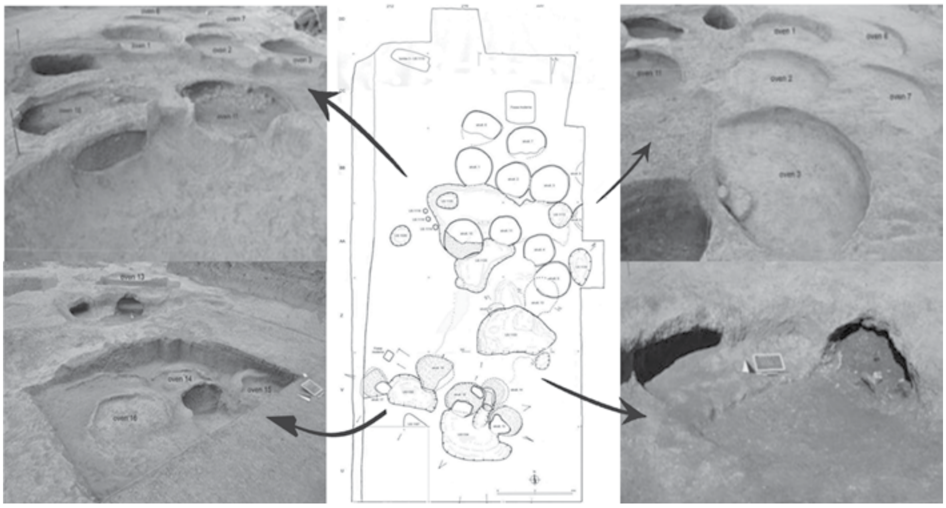


Figure 3: The excavated area and the domed ovens.

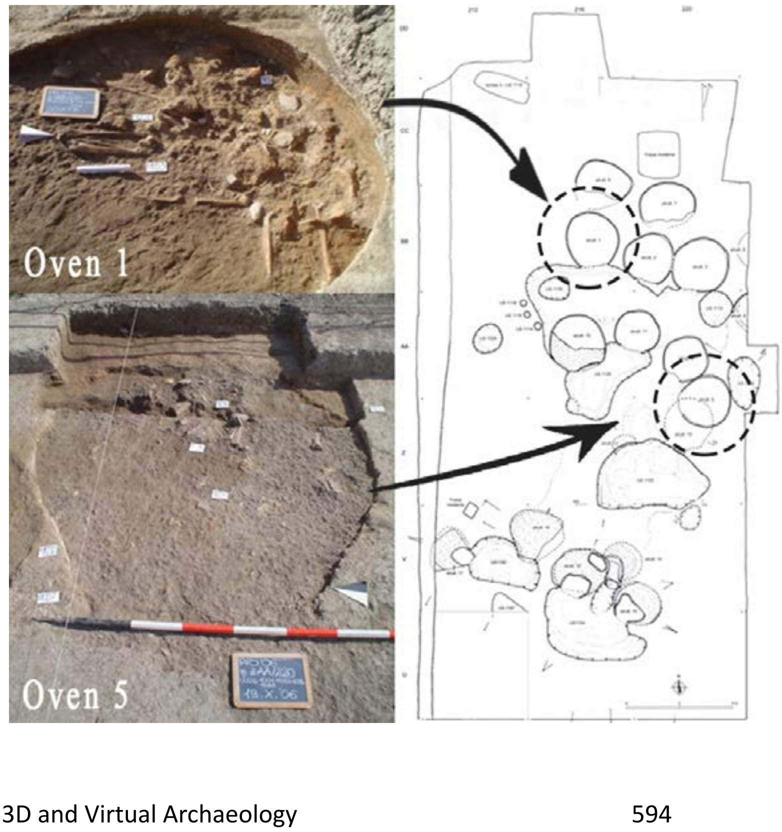


Figure 4: 3 badly preserved burials were found within two ovens.

Figure 5: Charred barley caryopses confirming the hypothesis of use of the ovens for roasting cereals.

These investigations revealed 18 domed ovens dug in the ground along the hill slope and dated at the half of the VI millennium BC cal. The ovens were aligned at different heights along the hillside overlooking shallow and irregular pits, which were dug to facilitate the excavation of the ovens and provide their access for use (Fig.3).

The downslope ovens are well preserved because they were protected by a thicker surface deposit, while the upper slope ovens are badly damaged by erosion and ploughing.

Six ovens are entirely preserved, while the others were cut by erosion and ploughing: these appeared as circular structures lined with clay. The ovens have circular and smoothed bases with diameters ranging from 1,80 to 2,00 m and an average height of 0,50 m. The inner lining is partly made by firing of the natural sediment and partly by addition of a clay paste.

Three badly preserved burials were found within two ovens: oven 1 contains the remains of two individuals while oven 5 contains one adult male of more than 55 years.

The PXRD analysis (X-ray powder diffraction) of hardened sediment samples of the inner walls of the ovens showed that the sediments were affected by exposure to very low temperatures not exceeding 500°C (Muntoni, Ruggiero 2013). This low value does rule out the possibility of using the ovens for firing pottery, but is compatible with other uses such as cooking and food processing.

Moreover, during the 2013 campaign dozens of charred barley caryopses were found inside three well-preserved ovens, thus confirming the hypothesis of use of the ovens for roasting cereals. Three radiocarbon dates were obtained so far:

- 6555 ± 45 BP - 5620-5460 BC cal σ 2, on a barley caryopses, found at the entrance of oven 14;
- 6500 ± 50 BP - 5560-5350 BC cal σ 2, on a charcoal from the base of oven 5, under the male burial;
- 6418 ± 50 BP - 5480-5310 BC cal σ 2, on a bone of the male burial in oven 5.

2.1 The archaeological materials

Archaeological materials are quite scanty suggesting that this area was devoted to ‘specialized’ activities related to food processing, which were performed ouside or near to a village, although any evidence of houses or domestic structures so far. were found.

Typical central Adriatic Impressa pottery, groundstones (mainly fragmented grinding slabs and grinders) and chipped stone artifacts were found in the pits in front of the ovens (Fig.6).

Flint cores, blades and blade lets, some of which showing thermal alteration, were found both outside and inside the ovens. These data allows us to hypothesize that ovens

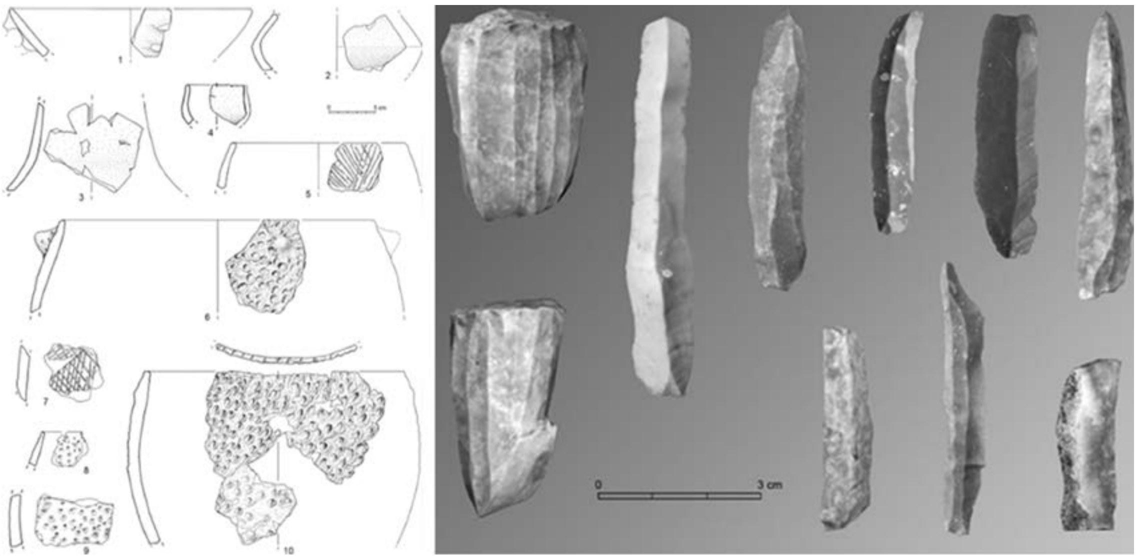


Figure 6: The archeological materials.

could also have been used to heat flint in order to facilitate the knapping operations.

According to the radiometric dates, the site of Portonovo-Fosso Fontanaccia represents a good evidence of the process of Neolithization of the middle-Adriatic Italian peninsula, standing out for the pottery and lithic production. This particular characteristics is widely spread from Abruzzo to Romagna regions.

3. Previous Documentation

2013 fieldwork recording strategies have so far been distinguished from the rest of the excavation documentation, which has largely remained on paper

During 2006 -2013 field campaigns covered around 300 square meters and recorded the archaeological excavations manly by hand drawn documents. The main working tool was the daily field report, used to register all the relevant observations and data interpretation during the workday.

Mapping draws were done manually using optical levels to record the height of each SU and the total stations to georeference grid points in a topographic reference system.

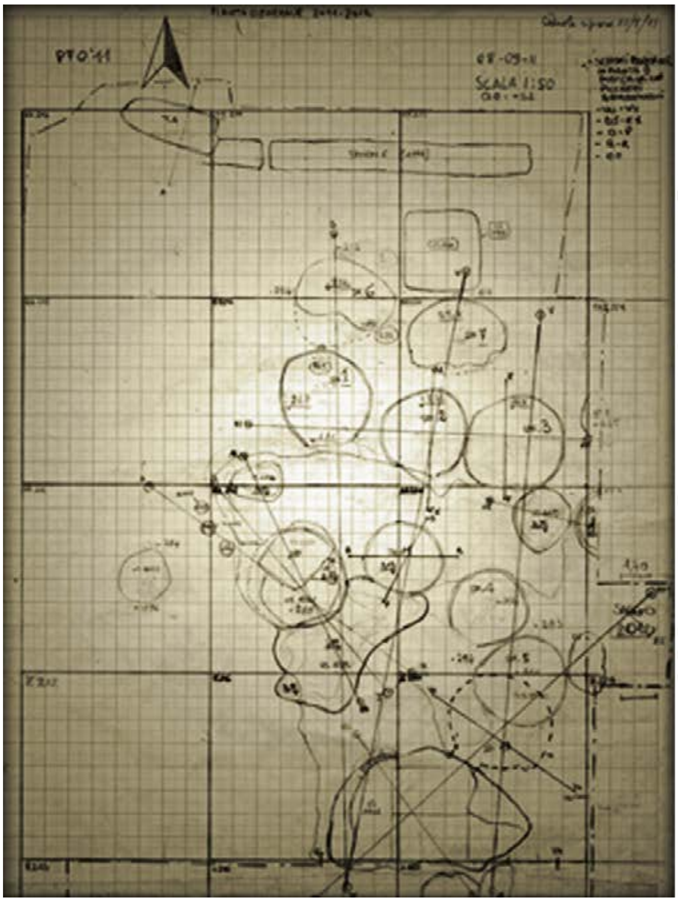
A grid system of 4mx4m was set up in 2006 , using the X axis as main reference system. This grid gradually enlarged over the years, according to the expansion of the excavation area. A Archaeological materials (lithics, pottery, faunal and botanical remains) were recorded together with their position, depth, Stratigraphic Unit.

Artefacts were removed and stored in numbered collections, marking them in detailed plans (scale 1:10), in order to help the identification of possible associations and refittings during post-dig study

Each structure was recorded by field finds labels, detailed plans (1:10) and sections (scale 1:20) to study morphology and depth articulation (Fig.7, left). General sections were also drawn in a smaller scale (1:50), in order to relate the structures along the hillside. All the recorded structures were placed in a general plan (scale 1:50). Each SU was documented by taking photographs at various stages of excavation: at the beginning, when the SU is recognized, while being excavated and finally when it is completely dug up. Moreover the wide shot and aerial photos allowed to record the excavation area as a whole.

SU Sheets were filled during the fieldwork, according to a widely shared form (Parise Bodoni, Ruggeri Giove 1984), storing all the SU information together with all the observed stratigraphic links, in order to recreate the stratigraphic sequence of the entire site.

After the excavation the work went on creating computerized database for the stratigraphic units, digitalizing the italian context sheets (SU reports) and rebuilding stratigraphic sequences (Fig.7, right). Part of the post-excavation study



Informazioni generali
ID: 49 — US: 1171 — Tagli:
Località: Portonovo Fosso Fontanaccia — Anno: 2013 — Data scavo:
Area: — Settore: V215-216 — Saggio: — Ambiente:
Quote: -105, -110 — Piante:
Schizzi: — Sezioni:
Prospetti: — Foto:

Descrizione US
Definizione: riempimento della struttura 18
Criteri distinzione: colore, consistenza — Natura (modo formaz.): naturale
Compon. organiche: numerose componenti organiche in disfacimento (carboni) — Compon. inorganiche: rari frammenti ceramici
Misure: — Stato conservaz: buono
Consistenza: friabile — Colore: bruno-nerastro
Campioni: — Flottazione:
Setacciatura:
Descrizione: strato di terra di colore bruno-nerastro a matrice argillosa, consistenza poco compatta, ricco di materiale organico dissoltosi
Osservazioni:

Rapporti stratigrafici
uguale a: — si lega a: — anteriore a: 1150
gli si appoggia: — si appoggia:
coperto da: 1150 — copre: 1170 — posteriore a: 1169
tagliato da: — taglia:
riempito da: — riempie: 1169
Interpretaz: riempimento del forno 18
Elementi datanti: ceramica
Datazione: Neolitico Antico — Fase:
Resposabile: La Marca

Figure 7: Previous documentation: Hand drawn plan (left) and SU reports (right).

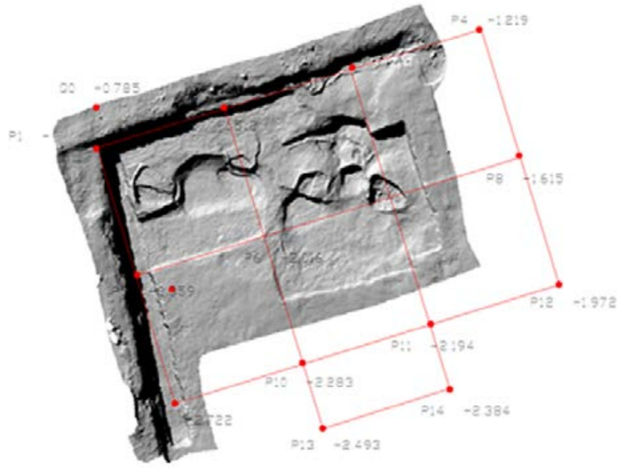


FIGURE 8: 3D MODEL AND DIG WORKING GRID.

consisted in the digital reworking of hand drawn plans and sections and in the analysis of the archaeological materials.

The SU form uses as a starting base the Italian Archaeological Find Sheet (Fresina 2006), with some modifications such as new categories and types.

A filemaker database was developed for pottery study, focusing on qualitative and quantitative analysis, technological and typological aspects.

Finally, graphic and photographic reproduction of artefacts is important both for documentation and scientific publication and allows to compare archaeological remains with those of other contemporary sites.

Previous fieldwork seasons revealed detailed stratigraphic sequences requiring a complex documentation, highlighting the need for managing data in a GIS platform, able to collect all the stratigraphic, topographic and cartographic data.

For this reason in 2013 first, new recording techniques (Photogrammetry and Laser-scanner) were tested. Second, a dedicated GIS has been developed and closely linked to the digital recording and data processing of archaeological excavation context.

4. Research Issues

The research approach tries to test a 3D documentation and GIS management pipeline for the archaeological site of Portonovo bringing in a systematic way 3D digital field recording procedures and combining them in a GIS environment with previous integrated data storage mechanisms.

In this context, different data are collected during last excavation (September and October 2013) by means of terrestrial laser scanner and photogrammetry and different deliverables are produced (photorealist 3D surface models, orthophotos, 3D sections etc.). 2D and 3D representations of all excavated features are then linked together with other digital data taking full advantage of recent open source GIS functionality in data management and supporting the archaeological knowledge production process.

Moreover the creation of photorealistic 3D models giving realistic representation of excavation features in digital 3D space will help to make visible and accessible this unique site in the panorama of Italian prehistory. This result gains in importance because the ovens of Portonovo are very fragile and prone to rapid deterioration. Therefore, the idea of an open-air musealization of these structures is not feasible.

5. 3D documentation

Different survey experiences are carried out to map the five ovens excavated during the 2013 field campaign, ranging from close-range Photogrammetric to Terrestrial Laser Scanner data acquisition and from open source to commercial processing. In particular, three surveys are carried out to map different excavation stages and the overall archaeological site:

- At excavation beginning (September 19), Photogrammetric survey (PHMY)
- During the excavation (September 27), Photogrammetric and Terrestrial Laser Scanning survey (PHMY and TLS)
- At excavation closing (October 10), Photogrammetric and Terrestrial Laser Scanning survey (PHMY and TLS)

FIGURE 9: GEOREFERECING THE TLS RANGE DATA BY GCPS AND EXCAVATION GRID.

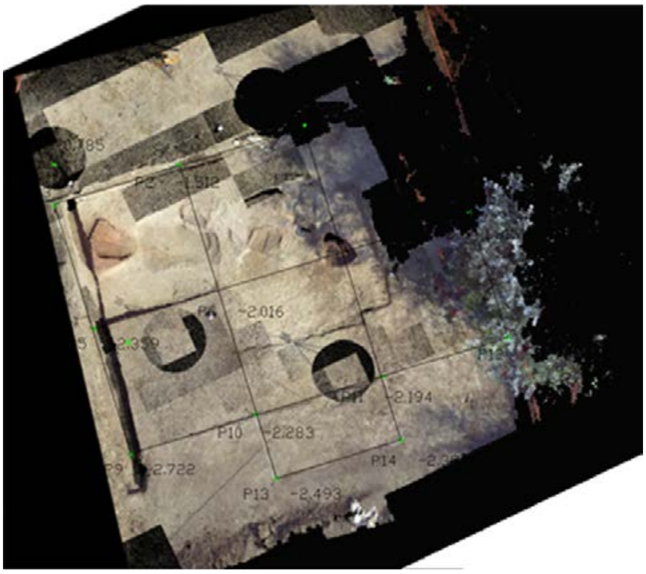
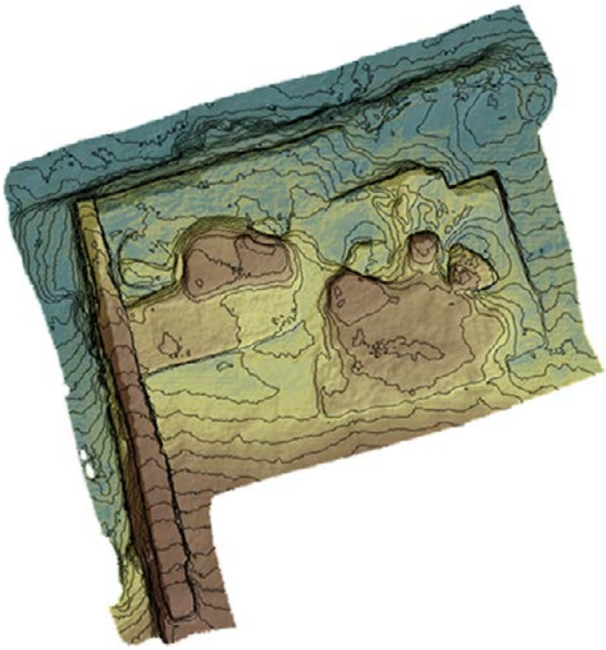


FIGURE 10: 2D ORTHOIMAGE.



FIGURES 11: DEM (GSD OF 2CM) AND 5 CM COUNTOUR LINES.

Ground control points are acquired with a total station to geo-reference the 2013 surveys, bring the two models (PHMY and TLS) into a common reference system set them in the existing dig working grid and link in with archival data from previous excavations (Fig.8).

Several recent publications compare the two mentioned technologies based on factors such as accuracy and resolution. Even if both technologies are nowadays capable of providing similar accuracy and resolution when supported by a well-designed digitization plan, their combination and integration remains the ideal solution to survey large and complex sites and improve the extraction of features with different geometric level of details. In particular laser scanning data is used as reference for a comparison with photogrammetric data and as documentation for the creation of high precision DSM and 2D metric deliverables (archaeological sections and orthoimage of the site). On the other hand, textured photogrammetric 3D models are used for visualization purposes of the entire site and assist archaeologists in interpretations of past uses of space. The collected images and range data are processed independently as described below.

The 3D data gathered so far (point clouds, triangulated meshes, orthophotos etc.) document with high quality the excavated surfaces but need to be archaeologically interpreted. These data formats have the potential to replace the traditional analogue excavation maps and drawings and to increase the scientific value and extend the application of the recorded archaeological data. During the process of interpretation, these collected documents have to be combined with all the other excavation data, especially with finds and samples. This is best done using GIS.

5.1 Laser Scanner survey

Acquiring large number of precise 3D data points, Laser scanners are effective tools to measure the topography of

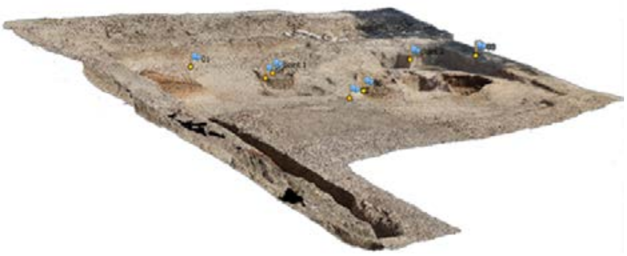


FIGURE 12: GEOREFERENCED TEXURED 3D MODEL.

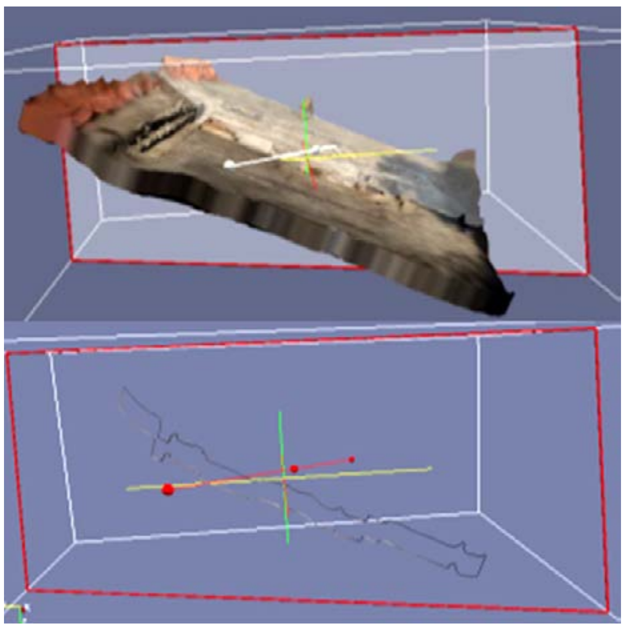


FIGURE 13: SECTIONS BY PHOTOGRAMMETRIC SURVEYING AND OPEN SOURCE PARAVIEW.

FIGURE 14: ORTOPHOTO BY PHOTOGRAMMETRIC AND LASER SCANNER SURVEYING.

the site as well as of the surface single archaeological units. Additionally they become invaluable when upstanding features like walls or cross sections make the recording process by total station time consuming and complicated or where, as happens to our case-study, the excavated surface is too vulnerable to be walked on (e.g. waterlogged environment, mosaics, organic material) (Doneus et al., 2005).

In total eight scans are performed by a CAM2 Focus3D laser scanner in two surveys carried out to map different campaign steps (September 27 and October 10). collecting data with a good average sampling distance of 3.1 mm and resulting in a dataset of ca. 45 million points acquired. Acquisition parameters are chosen as an acceptable compromise between level of detail of the final 3D model and computing resources needed for data processing. The number of the range acquisitions and stations for each temple depends on the dimensions and on the complexity of the monument. The positions of the different acquisitions have been organized to cover the entire volume of the monument, taking account of shadows, obstacles and undercut.

The laser scanner data are used as metric reference to scale the image data, for a geometric comparison with the photogrammetric data and for the creation of archaeological sections and plans of the site.

Range data are cleaned and aligned with the specific brand Faro SCENE software, then meshed, georeferenced (Fig.9) and edited using different software tools (Cyclone, Pointools, ParaView) as each of them has limitations in some processing functions.

Other bottlenecks, typical of the laser scanning processing are:

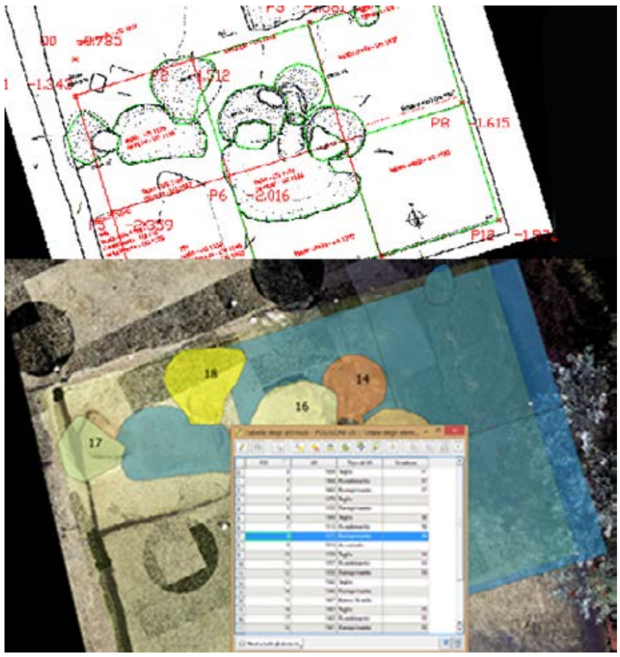


FIGURE 15: SU DIGITALIZATION.

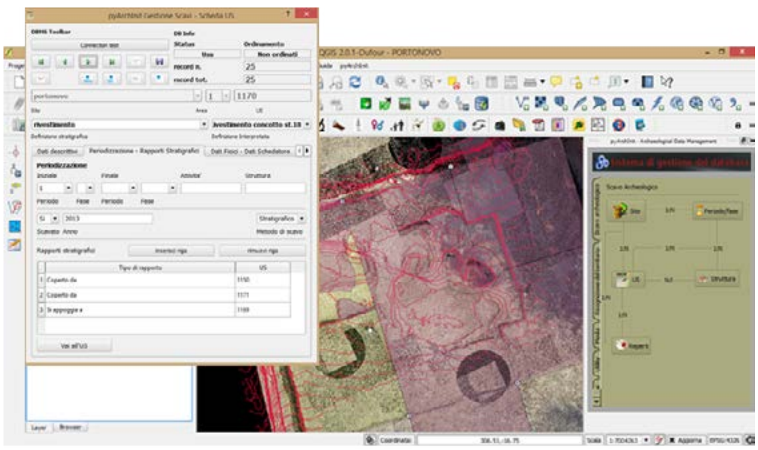


FIGURE 16: FILLING US REPORTS AND LINKING TO GIS FEATURES VIA US_TABLE.

- long editing / working time to generate a suitable polygonal model
- heavy 3D models not suitable to manage by a computer/laptop with medium characteristics (even after a decimation of the triangles)
- good texture normally missing and so requiring texture mapping procedures to map high resolution images onto the range-based 3D geometry

Nevertheless a 3D data from laser scanner data are easy to be acquired, accurate and different analyses and studies can be performed. Figures 10 and 11 show some processing steps and examples of final delirevables : 2D orthoimage, DEM (GSD of 2cm) and 5 cm countour lines.

This documentation helps in analysing the differences in shape and height of the different structures and in creating layout digitisations in Autocad and/or QGis, useful later to identify US and populate the PostgreSQL implemented database.

5.2 Close Range Photogrammetry survey

Over the last years, close-range photogrammetry methodology is showing promising characteristics for 3D spatial-data acquisition. Several photogrammetric approaches have already been used in archaeology and often dedicated to field archaeology, thanks to 3D restitutions of excavation units meeting archaeological accuracy requirements (Daniel *et al.*, 2008).

Approximately 300 images are acquired in three different surveys (September 19 and 27 and October 10) keeping the camera at the minimum focal length while the image resolution is set at the highest level in order to acquire good quality textures. Two different camera and settings are simultaneously used in order to speed up the survey: a Canon PowerShot SX30 IS (12mm, 4320 x 3240 pixels) and a SONY SLT-A65V (18 mm, 6000 x 4000 pixels). The distance to which the images are taken is quite constant due to the limited articulation of the archaeological site. The images are taken both convergent and nadir with a good overlap.

Terrestrial photogrammetric data are collected for the entire site because ideally-suited for 3D modelling purposes (confined, standing with no presence of grass etc.).

Both image alignment and 3D model reconstruction are fully automated using Agisoft Photoscan and some small user interaction allowed. Total station's GCPs and laser scanner data are used to drive the orientation and bundle adjustment and control the right scaling of the mesh. The results are quite dense, complete, with high quality texture and georeferenced (figure 12 & 13).

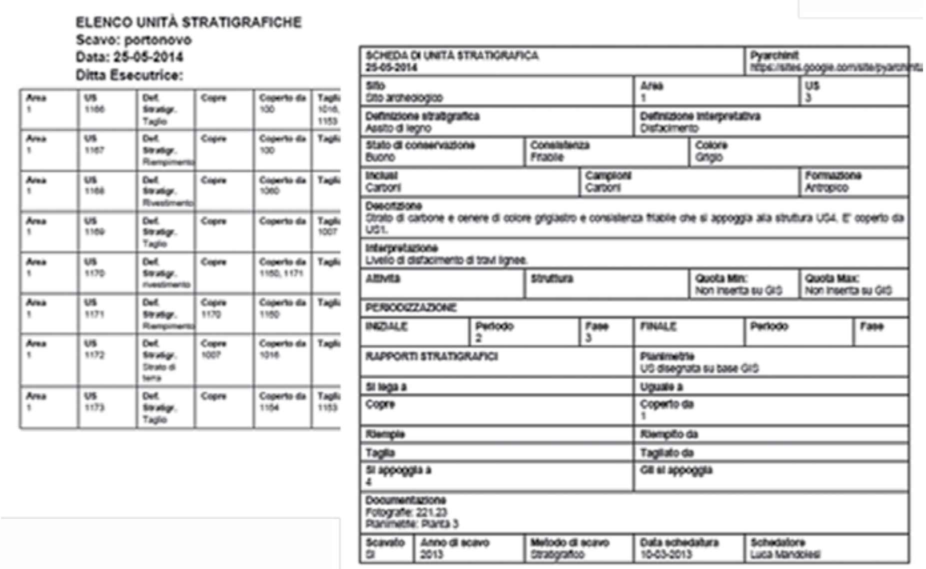
To improve the final result Photoscan is use to wrap the Photogrammetric textures over the TLS mesh and extract orthoimages (Fig.14)

6. Open Source GIS management

The digital recording solution introduced in this last excavation campaign resulted in a data production booming allowing: a systematic collection of excavation unit and artefact documentaion, a fast and detailed production of digital excavation plans and sections and an increase in the photographic documentation of the excavation

It was argued that in order to achieve a true integration of excavation data recording, management and representation, GIS technology had to be in the core of such attempt. Implementing a Geographical Information System was in fact possible to handle alphanumerical table, vector geometries, topographical and multimedia data within a single solution, keep as much as possible the integrity of the raw data and provide a very fast and robust management tool for daily life of archaeological fieldworks and archaeological site comprehension.

With the idea to test a really affordable working pipeline, it was decided to adopt the open source plug-in pyArchInit, well meeting the requirements of a user-friendly



ELENCO UNITÀ STRATIGRAFICHE
Scavo: portonovo
Data: 25-05-2014
Ditta Esecutrice:

Area 1	US 1166	Def. Stratigr. Taglio	Copre	Coperto da 100	Tagli 1016, 1153
Area 1	US 1167	Def. Stratigr. Riempimento	Copre	Coperto da 100	Tagli
Area 1	US 1168	Def. Stratigr. Rivestimento	Copre	Coperto da 1060	Tagli
Area 1	US 1169	Def. Stratigr. Taglio	Copre	Coperto da	Tagli 1007
Area 1	US 1170	Def. Stratigr. rivestimento	Copre	Coperto da 1150, 1171	Tagli
Area 1	US 1171	Def. Stratigr. Riempimento	Copre 1170	Coperto da 1150	Tagli
Area 1	US 1172	Def. Stratigr. Strato di terra	Copre 1007	Coperto da 1016	Tagli
Area 1	US 1173	Def. Stratigr. Taglio	Copre	Coperto da 1104	Tagli 1153

SCHEDA DI UNITÀ STRATIGRAFICA
25-05-2014
Pyarchinit
https://sites.google.com/site/pyarchinit

Sito: Sito archeologico | Area: 1 | US: 3
Definizione stratigrafica: Assito di legno | Definizione interpretativa: Disfacimento
Stato di conservazione: Buono | Consistenza: Friabile | Colore: Grigio
Inclusi: Carboni | Campioni: Carboni | Formazione: Antropico
Descrizione: Strato di carbone e cenere di colore grigiastro e consistenza friabile che si appoggia alla struttura US4. E' coperto da US1.
Interpretazione: Livello di disfacimento di travi lignee.
Attività | Struttura | Quota Min: Non inserita su GIS | Quota Max: Non inserita su GIS
PERIODIZZAZIONE
INIZIALE | Periodo: 2 | Fase: 3 | FINALE | Periodo | Fase
RAPPORTI STRATIGRAFICI | Planimetrie: US disegnata su base GIS
Si lega a | Uguale a
Copre | Coperto da: 1
Riempie | Riempito da
Taglia | Tagliato da
Si appoggia a: 4 | Gli si appoggia
Documentazione: Fotografie: 221.23; Planimetrie: Pianta 3
Scavato: Si | Anno di scavo: 2013 | Metodo di scavo: Stratigrafico | Data schedatura: 10-03-2013 | Schedatore: Luca Mandolesi

Figure 17: PDF outputs by PyArchInit plug-in.

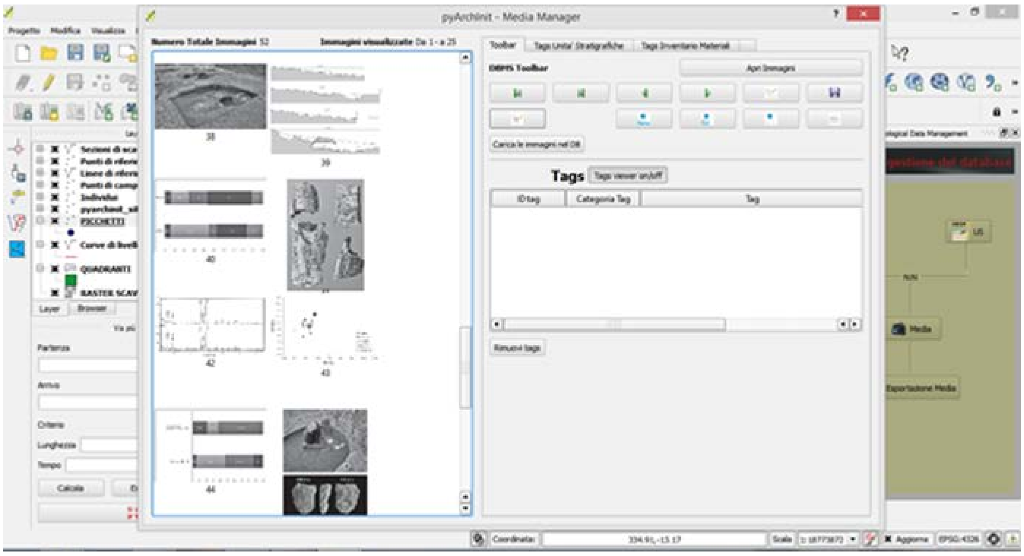


Figure 18: Media management by PyArchInit plug-in.

archaeological FOSS (Free and open-source software). PyArchInit is created by Mandolesi in 2009 (Mandolesi, 2009 and Mandolesi and Cocca 2012) and integrates itself inside the QGIS environment but it can stand alone too. It is pre-prepared, already successfully tested in several context (Gugnali et al., 2011) and perfectly in line with the needs which emerged during the work. Being able to manage excavation data (compatible with Italian ICCD standard record sheets), PyArchInit can support the archaeologist for his daily work in rescue excavations.

The first step of the GIS project was to georeference old and new excavation plans in the same reference system tied to the surveyed GCPs and reference excavation grid. Then all the stratigraphic units (pyarchnit_US) were created in order to manage on situ the documentation of on going excavations.

After having digitized the stratigraphic units, it was possible to fill US reports and link to GIS features via US_table (Mandolesi, 2009). Fig. 15 and Fig.16 shows some working steps related to Oven number18, from US digitalization to US_table filling-in by US reports and GUI data entry.

Alphanumeric, topographic and multimedia data coming from the above mentioned sources were stored in PostreSQL thanks to the different GUIs supporting the data entry and the database management. The geoDB allows to manage all the topographical objects: sections, sites, projects, stratigraphic units, small finds,etc. In particular seven management GUIs (Stratigraphic units, Site, Chronology, Infrastructures, Taphonomy, Record Archaeological, Multimedia) comes with the plug-in and allowed to handle and display in a single analysis environment different data (georeferenced plans, orthoimages) simultaneously. After data entry it was possible first to query the database and transform these queries into GIS visualization, then look for US and visualize linked US reports (Fig.17), images and excavation plans to work with.

Fig.18 shows how was possible, by means of the Multi Media manager, to tag images and link to US/artifacts (N:N) too.

A particular issue dealt with old excavation data and databases management. pyArchInit lacks of specific tools to import vector data created by other GIS software or stored in other databases. To use all the produced documentation it was necessary to import it in QGIS and then modify the relative attributes in order to match the pyArchInit PostGIS layer, to which all the vector data were further added. A similar matter was the data migration from Filemaker and access databases (used for US and pottery) to pyArchInit.

7. Conclusions

The entire methodology supports a complete digital excavation recording workflow and facilitates interpretation in a data-led synthetic manner. The produced photorealistic 3D models can be used to extract maps, plans, cross-sections, orthoimages for a technical public or for conservation issues and together with all the other digital data can be integrated inside a GIS environment satisfying the need to manage on situ the documentation of on going excavations.

The option of quick, well-tested and often low cost/open source documentation makes the research experience a case-study highlighting new approaches that can be integrated in the general excavation methodology and additional interesting features such as model/data reusability

Moreover, the photorealistic 3D models can be refined and drive other cultural heritage experiences such as multimedia enjoyment and virtual museums, enhancing the knowledge of the site or helping education and promotion. This communication feature becomes especially important when the archaeological site (as happens to our case-study) must be covered/closed in default of other protecting / museum solution.

On-going research focuses in augmented reality to visualize GIS layers and 3D models and allow the exploration of the archaeological area of Portonovo on situ even post-excavation.

Acknowledgements

Special thanks go to L. Mandolesi for his released PyarchInit plug-in and his essential on-line support (PyarchInit mailing list, blog etc) and to ing. C. A. Bozzi for his help during the 3D survey and processing.

References

BARCELO, A.; FORTE, M.; SANDERS, D.H. (2000). *Virtual Reality in Archaeology.* BAR International Series 843, Oxford: Archeopress.

BRUNO, F.; BRUNO, S.; DE SENSI, G.; LUCHI, M.L.; MANCUSO, S.; MUZZUPAPPA, M. (2010). From 3D reconstruction to virtual reality: A complete methodology for digital archaeological exhibition. *Journal of Cultural Heritage*, 11, pp.42-49.

CONATI BARBARO, C. (2013). Cooking, working and burying in ancient Neolithic: the ovens of Portonovo (Marche, Italy), *Origini* XXXV, pp. 31-51.

COWLEY, D.C. (2011) *Remote Sensing for Archaeological Heritage Management. EAC Occasional Paper No. 5; Occasional Publication of the Aerial Archaeology Research Group No. 3*; Europae Archaeologiae Consilium: Budapest, Hungary; p. 307.

d' ANNIBALE, E., TASSETTI, A.N., MALINVERNI, E.S. (2014). Finalizing a low-cost photogrammetric workflow: from panoramic photos to Heritage 3D documentation and visualization, International J*ournal of Heritage in the Digital Era* 3(1), pp.33-50, doi 10.1260/2047-4970.3.1.33.

DANIEL, S., DUGUAY, J., POULIOT, J. (2008). Close range photogrammetry experiments for 3D modeling of archaeological excavation site, *Geomatica* 62(3), pp.411-417.

DONEUS, M.; NEUBAUER, W. (2005). 3D Laser Scanners on Archaeological Excavations. In: *Proceedings of CIPA 2005 XX International Symposium, Torino, Italy, 26 September–1 October 2005:* Volume 34(5/C34/1), pp. 226-231.

FANGI, G., MALINVERNI, E.S., TASSETTI, A.N. (2013). The metric documentation of Cham Towers in Vietnam by Spherical Photogrammetry. In: *ISPRS Annals of the Photogrammetry, Remote Sensing and Spatial Information Sciences, Volume II-5/W1, 121-126, XXIV International CIPA Symposium, 2-6 September 2013, Strasbourg, France*, doi:10.5194/isprsannals-II-5-W1-121-2013,.

FRESINA A. (2006). *Norme per la catalogazione di reperti archeologici e numismatica: Scheda RA-N: beni archeologici* / Centro regionale per l'inventario, la catalogazione e la documentazione grafica, fotografica, aerofotografica, audiovisiva e Filmoteca regionale siciliana, Palermo: CRICD.

GUGNALI, S., MANDOLESI, L., DRUDI, V., MIULLI, A., MAIOLI, M.G., FRELAT, M.A., GRUPPIONI, G.F. (2011). Design and Implementation of an Open Source G.I.S Platform For Management of Anthropological Data, *Journal of Biological Researches-Thessaloniki*. n.1, Vol. LXXXIV, Rubettino-Soveria Mannelli, 84, pp. 350-353.

MANDOLESI L. (2010). PyArchInit – Qgis, e PostgreSQL per la gestione dei dati di scavo. In: Cignoni, P., Palombini, A., Pescarin, S. (eds.). *ARCHEOFOSS. Open Source, Free Software e Open Format nei processi di ricerca archeologica. Atti del IV Workshop (Roma 2009), Archeologia e Calcolatori, Supplemento 2*, Firenze: All'Insegna del Giglio, pp. 209-222.

MANDOLESI L., COCCA, E. (2013). PyArchInit: gli sviluppi dopo Archeofoss 2009. In: Serlorenzi, M., (eds.). *ARCHEOFOSS, Free, Libre and Open Source Software e Open Format nei processi di ricerca archeologica. Atti del VII Workshop (Roma 2012), Archeologia e Calcolatori, Supplemento 4*, Firenze: All'Insegna del Giglio, pp. 128-138.

MANFERDINI, A.M.; REMONDINO, F. (2010) Reality-Based 3D Modelling, Segmentation and Web-Based Visualization. In: *Proceedings of EuroMed'10 the Third International Conference on Digital Heritage, Lemessos, Cyprus, 8–13 November 2010*, pp. 110-124.

MUNTONI I.M., RUGGIERO G. (2013). Estimating the Firing temperatures of pyrotechnological processes in Neolithic site of Portonovo, *Origini* XXXV, pp.52-56.

PARISE BADONI F., RUGGERI GIOVE M (eds), (1984). *Norme per la redazione della scheda del saggio stratigrafico*, Roma; Multigrafica

PYARCHINIT WEBSITE (2014). PyarchInit - free open source software per l'archeologia documentation. [Online] https://sites.google.com/site/pyarchinit/ [Accessed: 14th July 2014].

Chapter 9. Multi-Agent Systems and Complex System Modelling

A Density-Based Simulation Approach for Evaluating Prehistoric Population Fluctuations in Finland

Tarja SUNDELL,[1] Juhana KAMMONEN[2]

[1]Department of History, Philosophy, Culture and Art Studies, University of Helsinki, Finland
[2]Institute of Biotechnology, University of Helsinki, Finland
tarja.sundell@helsinki.fi

Abstract
Finland's isolated geographical location in North Eastern Europe has greatly affected its settlement history. In this paper we expand upon our prior studies on prehistoric population events and fluctuations in the area. We apply a forward in time population genetics simulation environment simuPOP which allows continuous development of the simulation model. Here, we adjusted the individual mobility model to follow a widely acknowledged method known as density-based clustering. The results of our simulations demonstrate that density-based clustering manages to capture the subpopulation fine-structure and produces population genetic diversity measures similar to those attained in our previous research. The revised model also allows integration of additional archaeological data into the simulation. This work contributes to the description of prehistoric population processes in Finland and elsewhere.

Keywords: Prehistoric Population, Population Simulation, SimuPOP, Density Based Clustering, Population Bottleneck, Stone Age

1. Introduction

The prehistoric population dynamics and fluctuations in Finland have been studied by a number of researchers during the last years (e.g. Oinonen *et al.,* 2010; Tallavaara *et al.,* 2010; Hertell & Tallavaara, 2011; Rankama & Kankaanpää, 2011; Kammonen *et al.,* 2012; Tallavaara & Seppä, 2012; Sundell *et al.,* 2010, 2013, 2014; Sundell 2014; Manninen & Knutsson, 2014). Furthermore, prehistoric population processes inaccessible by other means have been studied by archaeological and genetic population simulations (Sundell *et al.,* 2010, 2013). In these simulations populations can be simulated either forward in time or backward in time into the coalescent (the most recent common ancestor) i.e. coalescent simulations.

For this work we chose the forward simulation approach as it allows population genetic processes, such as mutation, migration and genetic drift to be included in the process from the beginning of the simulation. With forward in time simulations virtual populations are created and simulated through their entire histories. The simulated populations are then sampled and analyzed and the results compared with real data from present day genetic diversity. SimuPOP (Peng & Kimmel 2005), a Python-operated forward time population genetics simulation environment, was used in this study (http://www.python.org). We chose simuPOP as it is the most versatile forward time population simulation software available. SimuPOP consists of a number of components from which users assemble a suitable simulator. The simulations were run in the Ukko supercomputer of the Department of Computer Science at the University of Helsinki.

2. Materials and Methods

We performed four different population genetic simulation scenarios moving forward in time to evaluate possible past demographic events in prehistoric Finland (see Table 1). The simulations start with small pioneer populations, go through two archaeologically justified prehistoric population bottlenecks and finally the population grows exponentially, known from historical sources. We used the same demographic model as in our previous work (Sundell *et al.,* 2013) to be able to make comparisons and draw conclusions between different parameters. An overview of the general simulation model is illustrated below (Figure 1).

Whereas the previous simulation model included only relative geographical location information, the simulated individuals now have separate information fields for geographical longitude and latitude (x and y coordinates). In order to simulate fine-structured subpopulations, we now employ a density-based clustering method called DBSCAN (Density Based Spatial Clustering of Applications with Noise) (Ester *et al.,* 1996) to divide individuals into subpopulations based on geographical distance. The method promises to find clusters of arbitrary shape and have good efficiency on large databases of thousands of objects. Evidently, a simulated population of individuals with spatial location serves as a large database. DBSCAN is widely used in geospatial applications (e.g. Wang *et al.,* 1997, 2000; Handl *et al.,* 2005).

The key structural elements in our simulation scenarios are:

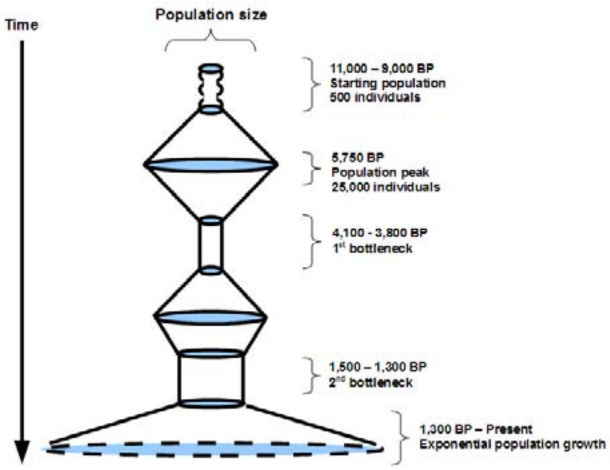


Figure 1: The general model of prehistoric demography in Finland used in our simulations. The width of the cone represents the relative population size at the time before present (BP).

Genetic components

- Both mitochondrial DNA and Y chromosomes are simulated including mutations
- Reproductive ages are set at 20-60 years for males and 20-40 years for females
- The maximum lifespan is 60 years
- The simulated population is age-structured, in other words, generations overlap
- The natural mortality rate is 15% per ten years
- The mating of individuals depends on the density-based clustering of groups
- The mortality rate is fixed and is based on random sampling
- Each mating produces 1 to N offspring according to Poisson distribution. Offspring is produced for each pair until it, together with the surviving part of the population, reaches the population size at the next simulation step.

Starting point

- Pioneer population of 250 females and 250 males that is divided into two subpopulations
- Fluctuating start (serial founder effect) vs. stable start
- The simulation moves forward in ten-year-steps.

Population bottlenecks

- The population size is set to grow exponentially up until 5750 BP, after which the population begins to gradually decline over 1600 years towards the bottleneck
- The models include two bottlenecks, first at 4100-3800 BP with population census sizes with 1000 and 200 individuals and a second, less severe bottleneck at 1500-1300 BP with population census size with 10,000 individuals.

Migration and gene flow

- Immigration involved in the appearance of Typical Comb Ware (TCW) and Corded Ware (CW)
- Internal gender-specific mobility: the mobility rate of females is ten times larger than that of males.
- Constant gene flow from neighbouring populations: archaic European, archaic Scandinavian and Saami.

Final population

- Timespan simulated 11,000 years, from early prehistory to this day
- The population is simulated with true census numbers i.e. the whole population
- Four separate simulation scenarios selected for comparison from previously simulated 24 scenarios (Sundell *et al.,* 2013, see Table 1)
- Nine separate simulation checkpoints selected to evaluate the indicators of genetic diversity compared with previous simulations (Sundell *et al.,* 2013, see Table 2)
- Each simulation scenario is run 50 times to obtain enough replicates for evaluation of random variation to the result.

The simulations begin at 11,000 BP when the first postglacial pioneers settled the region following the retreat of the continental ice sheet. The population growth is simulated according to the strength of the archaeological signal (Tallavaara *et al.,* 2010). We employ two archaeologically justified bottlenecks: one at 4100-3800 BP and another at 1500-1300 BP. The second, Iron Age bottleneck, has been less severe. The simuPOP package accepts absolute BP years only. Therefore we use BP as the unit of time in all the simulations, contrary to the other datings in this research.

The simulation starts with a small initial pioneer population of 250 females and 250 males that is divided into two subpopulations: Saami and Other Finland. We employ a scenario with fluctuating population size in the initial phase, reminiscent of serial founder effects. Serial founder effects often take place in real world founding populations. Here, the population size fluctuates between 240-630 individuals. Both subpopulations have the same size change rate, with population minima reached c. every 200 years. With Saami, here, we refer specifically to genetic Saami whose origins remain partially open. The total population size is set to grow exponentially from 9000 BP onward until the first bottleneck occurs.

The genetic effects were measured with two basic indicators of genetic diversity: the number of haplotypes (A) and haplotype diversity ($\hat{H}$) in a sample. The first is simply a direct count of different haplotypes (differing in at least one nucleotide position or microsatellite locus in mtDNA and Y chromosomes, respectively). $\hat{H}$ (Nei, 1987) is based on population haplotype frequencies and measures the probability of observing two different haplotypes when sampling two random chromosomes or, as in this case, haploid individuals, from a population. Haplotype diversity is calculated with the formula $\hat{H} = n(1-\Sigma xi2)/(n-1)$ where N is the number of individuals and xi the frequency of the ith haplotype. These definitions of A and $\hat{H}$ are the same as in our previous simulations (Sundell *et al.,* 2013) to allow comparisons between different simulation instances during the development of the methodology e.g. the effects of density-based clustering. Mitochondrial DNA (mtDNA) is maternally transmitted and simulated as a 631 bp DNA sequence, corresponding to HVS-I and HVS-II of the mtDNA control region. Hypervariable segments (HVS) of mtDNA are the segments that typically are sequenced when mtDNA lineages are studied. Any nucleotide at any position can mutate creating new variation. The mutation rates used in the simulations are based on values published in Heyer et al. 1997, 2001; Kayser & Sajantila 2001 and Sigurdardottir et al. 2000. In order to make the model even more realistic, the chromosomes were initialized with actual Finnish genetic data including distinct mutations (Hedman et al. 2007). In addition, a binary approximation for the mtDNA nucleotides was used in order to save random access memory during the simulation (Kammonen, 2013).

Background populations are used to model external migration waves and minor gene flow into Finland: Archaic European, Archaic Scandinavian and Saami. The composition of the background populations is the same as used previously (Sundell *et al.,* 2013). The background populations evolve separately for 12,000 years and a snapshot is saved every 2000 years. The snapshot populations are then used as source pools for immigration. The Archaic European background population includes 50,000 individuals, the Archaic Scandinavian 25,000 individuals and the Background Saami 5000 individuals.

We include two migration waves from neighbouring populations, the TCW migration wave at around 6000 BP and the CW migration wave at around 5200 BP. In the migration waves, the Typical Comb Ware migration wave replaces c. 2% of the north-eastern subpopulation with Archaic European and the Corded Ware migration wave replaces c. 0.8% of the south-western subpopulation with Archaic European. In addition to the above-mentioned specific migration waves, moderate constant gene flow replaces 0.01% of the population with Archaic European every 10 years during the entire simulation. Additionally, 0.005% of the area populated by Saami and the north-eastern area is replaced with gene flow from Saami background population every 10 years and after 3500 BP 0.05% of the south-western area is replaced with migration from Archaic Scandinavian background population. The migration rates were chosen based on our previous work (Sundell *et al.,* 2010, 2013). In our previous study we used higher gene flow rates and came to the conclusion that lower gene flow rates should be explored since the high rates forced the simulated populations' diversity to the background populations' which is not true even for the current Finnish population.

Total population is set to reach a maximum of 25,000 individuals shortly before 5750 BP to correspond to the accumulating archaeological evidence for a Stone Age population peak in Finland (Tallavaara *et al.,* 2010; Tallavaara & Seppä, 2012; Sundell *et al.,* 2014). The exponential growth slows down before this and the

Scenario	Population size first 2000 years	Internal mobility between subpopulations	Migration waves (TCW and CW)	Constant gene flow	Bottleneck census size at 4100-3800 BP	Bottleneck census size at 1500-1300 BP
1	fluctuating	yes	-	-	1000	10 000
2	fluctuating	yes	-	-	200	10 000
3	fluctuating	yes	small	moderate	1000	10 000
4	fluctuating	yes	small	moderate	200	10 000

Table 1: Four different simulation scenarios. The highlighted scenarios with temperate constant gene flow (3-4) are those where the strongest similarity to present day genetic diversity was observed in the previous simulation study (Sundell et al. 2013).

	Checkpoint	Time BP	Simulation steps
1	Starting point	11,000	0
2	Population starts to grow	9,000	200
3	Sampling before the population split	7,000	400
4	Population peak	5,750	525
5	Before the first bottleneck	4,650	635
6	Immediately after the population minimum	4,090	691
7	Second population peak	2,000	900
8	In the middle of the second bottleneck	1,400	960
9	Present time, final generation	0	1099

Table 2: Simulation checkpoints. In every checkpoint, a random sample of 1000 individuals is extracted from the population extant at the time. If total population size at this point is smaller than 1000, all individuals in the population are included in the sample. Checkpoints are used to verify the location of individuals in the study area and to measure different genetic effects in the simulations.

population size remains approximately the same for a brief period after this time point. After 5750 BP the population gradually begins to decline, which continues over 1600 years towards the Stone Age bottleneck. After the bottleneck the population slowly recovers, followed by a less severe second bottleneck at 1500-1300 BP. During the last 1300 years the population size is set to grow to 1,000,000 individuals, the final population size of the simulations.

3. Results

We carried out four different simulations (1-4). The genetic effects were measured with two basic indicators of genetic diversity: the number of haplotypes (A) and haplotype diversity (Ĥ) in a sample. The results of our simulations demonstrate that density-based clustering manages to capture the subpopulation fine-structure and, in addition, produces diversity measures similar to those in our previous research which validates the model. All simulation instances in all four scenarios produce diversity measures that are in the close vicinity of real-world data (Figures 2-4). Furthermore, the number of Y-chromosomal haplotypes (Figure 2) is even closer to the real-world reference line than the values observed in our previous simulations (Sundell *et al.,* 2013).

Our results are best illustrated in the three boxplots below (Figures 2-4). The simulation result datasets were visualized with IBM SPSS's PASW Statistics 18, boxplot utility (SPSS Inc. 2009). The boxes contain 50% of the observed values. The vertical lines that end in a horizontal stroke, whiskers, give information about the spread of the data: approximately 95% of the data lie between the inner fences. The stroke in the middle implies the median of the observations.

We tested different mobility patterns by allowing individuals to move at different probabilities. Our results indicate that the density-based clustering of simulated individuals seems to be a more defining parameter than a high versus low probability for mobility among the individuals, in other words, density-based clustering seems to obscure the effect of different mobility probabilities as a less important factor. Clustering of individuals mimics the natural processes in which people form families and larger groups.

4. Discussion

Prehistoric population fluctuations and their consequence to the present day genetic diversity can be studied by population simulations where evolutionary forces i.e. mutation, migration and genetic drift are incorporated. We performed four different simulation scenarios including permutations of the following parameters: fluctuating population growth during the first 2000 years, internal mobility between subpopulations, varying levels of migration as well as constant gene flow and two prehistoric population bottlenecks (Table 1). In order to model population movement more accurately, we introduced a discrete measure of geographical distance between the simulated individuals and altered the individual mobility model to follow a widely acknowledged method known as density-based clustering. This method allows individuals to be divided into subpopulations and form clusters based on geographical distance. Whereas the previous simulation models included only relative geographical location information, the simulated individuals now have separate information fields for geographical longitude and latitude (x and y coordinates). In addition of individuals moving

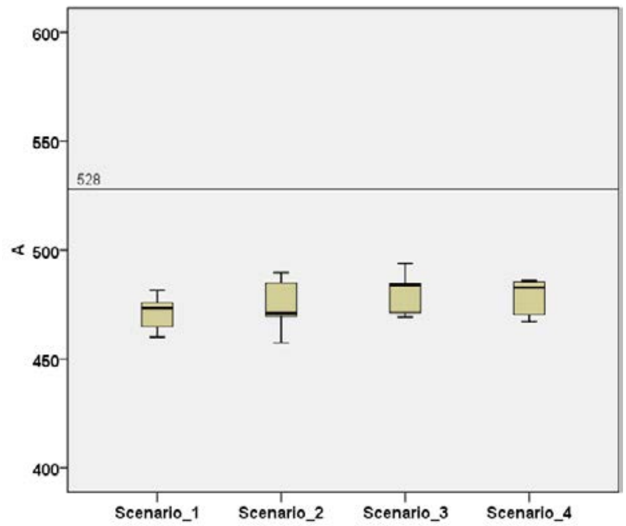


Figure 2: The number of Y-STR haplotypes (A) in final population (0 BP) with scenarios 1-4 with 50 simulation replicates, each with a sample of 907 individuals. The reference line represents the value observed in the present (528 objects) (Palo et al. 2009). The number of individuals in a sample corresponds to the sample size in Palo et al. 2009 enabling direct comparison between the simulations and the real world.

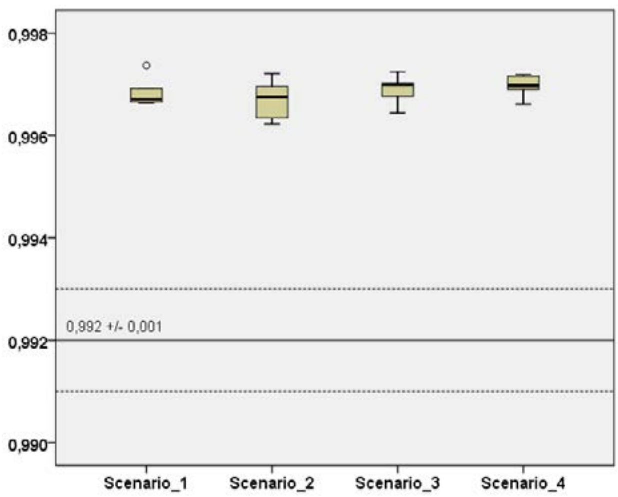


Figure 3: The Y-STR haplotype diversity (Ĥ) in final population (0 BP) with scenarios 1-4 with 50 simulation replicates, each with a sample of 907 individuals. The reference line represents the value observed in the present (0.992 + 0.001) (Palo et al. 2009). The number of individuals in a sample corresponds to the sample size in Palo et al. 2009 enabling direct comparison between the simulations and the real world.

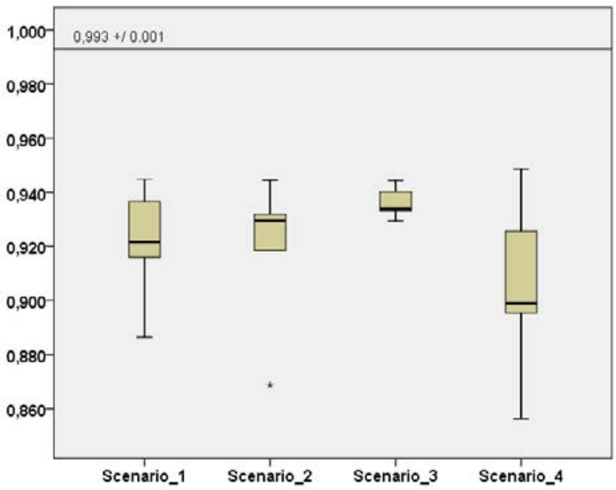


Figure 4: The Mitochondrial haplotype diversity (Ĥ) in the final population (0 BP) with scenarios 1-4 with 50 simulation replicates, each with a sample of 832 individuals. The reference line represents the value observed in the present (0.993 + 0.001) (Palo et al. 2009). The number of individuals in a sample corresponds to the sample size in Palo et al. 2009 enabling direct comparison between the simulations and the real world.

in a distant area, here, they are prone to settle areas and form groups if the conditions are favourable. This real-world phenomenon is better modelled with clustering of individuals.

The number of possible simulation settings is almost limitless. However, simulation models can never be an exact replication of the complex reality, but instead, are used to evaluate which population scenarios are most likely to be true. One of the principal benefits of simulation models is that you can begin with a simple approximation of a process and gradually refine the model as your understanding of the process improves or new data becomes available.

The main advantage of this new revised approach lies in its multi-function usability: it allows for integration of various types of additional data, such as archaeological location data, terrain maps and GIS-datasets, also used in our previous research on the Finnish population prehistory (e.g. Pesonen et al. 2011, Kammonen et al. 2013). Archaeological and genetic research continuously produces new data that can be corporated into the forthcoming work. Due to this new enhanced version of the simulation tool, the software allows real-time simulation of neighbouring populations, in other words, the background gene pool. This allows simulation of multifold background populations which produces more dynamic migration and constant gene flow effects.

Acknowledgements

This work was supported by the Nordenskiöld Foundation and the Jane and Aatos Erkko Foundation.

References

ESTER, M., KRIEGEL, H-P., SANDER, J. & XU, X. (1996). A Density-Based Algorithm for Discovering Clusters in Large Spatial Databases with Noise. Simoudis, E., Han, J., Fayyad, U.M., (eds.). *Proceedings of 2nd international Conference on Knowledge Discovery and Data Mining (KDD-96).* Palo Alto, California: AAAI Press pp. 226-231.

HANDL, J., KNOWLES, J., & KELL, D.B. (2005). Computational cluster validation in post-genomic data analysis. *Bioinformatics* 21, pp.3201–12.

HEDMAN, M., BRANDSTÄTTER, A., PIMENOFF, V., SISTONEN, P., PALO, J. U., PARSON, W. & SAJANTILA, A. (2007). Finnish mitochondrial DNA HVS-I and HVS-II population data. *Forensic Science International* 172, pp. 171-8.

HERTELL, E. & TALLAVAARA, M. (2011). Hunter-Gatherer Mobility and the Organization of Core Technology in Mesolithic North-Eastern Europe. In: T. Rankama (ed.). *Mesolithic Interfaces. Variability in Lithic Technologies in Eastern Fennoscandia.* Monographs of the Archaeological Society of Finland 1, pp. 95-100.

HEYER, E., PUYMIRAT, J., DIELTJES, P., BAKKER, E. & DE KNIJFF, P. (1997). Estimating Y Chromosome Specific Microsatellite Mutation Frequencies Using Deep Rooting Pedigrees. *Human Molecular Genetics*. 6(5), pp. 799-803.

HEYER, E., ZIETKIEWICZ, E., ROCHOWSKI, A., YOTOVA, V., PUYMIRAT, J. & LABUDA, D. (2001). Phylogenetic and Familial Estimates of Mitochondrial Substitution Rates: Study of Control Region Mutations in Deep-Rooting Pedigrees. *American Journal of Human Genetics* 69(5), pp. 1113-26

KAMMONEN, J. (2013). *Improving the throughput of the forward population genetic simulation environment simuPOP.* MSc thesis, Department of Computer Science, University of Helsinki. Helsinki.

KAMMONEN, J., SUNDELL, T., MOLTCHANOVA, E., PESONEN, P., OINONEN, M., HEGER, M., HAIMILA, M. & ONKAMO, P. (2012). Bayesian Spatial Analysis of Archaeological Finds and Radiocarbon Dates: An example from Finland 4000-3500 cal BC. In: Zhou, M., Romanowska, I. Wu, Z., Xu, P. & Verhagen, P. (eds.). *Revive the Past. Computer Applications and Quantitative Methods in Archaeology (CAA). Proceedings of the 39th Annual Conference, Beijing, April 12-16.* Amsterdam: Pallas Publications. pp.318-25.

KAMMONEN, J., SUNDELL, T., PESONEN, P., OINONEN & ONKAMO, P. (2013). Bayesian Spatial Modelling of Radiocarbon Dated Archaeological Artefacts Using R-INLA. In: Earl, G., Sly, T., Chrysanthi, A., Murrieta-Flores, P., Papadopoulos, C., Romanowska, I. & Wheatley, D. (eds.). *Archaeology* *in the Digital Era. Papers from the 40th Annual Conference of Computer Applications and Quantitative Methods in Archaeology, Southampton, 26-29 March 2012.* Amsterdam University Press, pp. 414-419.

KAYSER & SAJANTILA. (2001). Mutations at Y-STR loci: implications for paternity testing and forensic analysis. *Forensic Science International* 118, pp.116-21.

MANNINEN, M. A. & KNUTSSON, K. (2014). Lithic raw material diversification as an adaptive strategy – Technology, mobility, and site structure in Late Mesolithic northernmost Europe. *Journal of Anthropological Archaeology* 33, pp. 84-98. doi:10.1016/j.jaa.2013.12.001.

NEI, M. (1987). *Molecular evolutionary genetics.* Columbia University Press.

OINONEN, M., PESONEN, P. & TALLAVAARA, M. (2010). Archaeological radiocarbon dates for studying the population history in eastern Fennoscandia. *Radiocarbon* 52, 393–407.

PALO, J. U., ULMANEN, I., LUKKA, M., ELLONEN, P. & SAJANTILA, A. (2009). Genetic markers and population history: Finland revisited. *European Journal of Human Genetics* 17, pp. 1336-1346.

PENG B. & KIMMEL M. (2005). SimuPOP: a forward-time population genetics simulation environment. *Bioinformatics* 21, pp. 3686-3687.

PESONEN P, J. KAMMONEN, E. MOLTCHANOVA, M. OINONEN & P. ONKAMO. (2011). Archaeological Radiocarbon Dates and Ancient Shorelines - Resources and Reservoirs. In: Ikonen, A.T.K. & Lipping, T. (eds.). *Proceedings of the seminar "The Sea Level Displacement and Bedrock Uplift", Pori, Finland.* Posiva Working Report 2011-07, pp. 119-29.

Python Programming Language. Accessed 31 July 2014. http://www.python.org.

RANKAMA, T. & KANKAANPÄÄ, J. (2011). First evidence of eastern Preboreal pioneers in arctic Finland and Norway. *Quartär* 58, pp. 183-209.

SIGURDARDOTTIR, S., HELGASON, A., GULCHER, J.R., STEFANSSON, K., & DONNELLY, P. (2000). The Mutation Rate in the Human mtDNA Control Region. A*merican Journal of Human Genetics* 66, pp. 1599-1609.

SPSS INC. (2009). PASW Statistics for Windows, Version 18.0. Chicago.

SUNDELL T., HEGER M., KAMMONEN J. & ONKAMO P. (2010). Modelling a Neolithic Population Bottleneck in Finland: A Genetic Simulation. *Fennoscandia Archaeologica* XXVII, pp. 3-19.

SUNDELL, T., KAMMONEN J., HEGER M., PALO J. & ONKAMO P. (2013). Retracing Prehistoric Population Events in Finland Using Simulation. In Earl, G., Sly, T., Chrysanthi, A., Murrieta-Flores, P., Papadopoulos, C., Romanowska, I. & Wheatley, D. (eds.). *Archaeology in the Digital Era. Papers from the 40th Annual Conference of Computer Applications and Quantitative Methods in Archaeology,* Amsterdam University Press, pp. 414-419.

SUNDELL, T., KAMMONEN, J., Halinen, P., Pesonen, P. & Onkamo, P. (2014). Archaeology, genetics and a population bottleneck in prehistoric Finland. *Antiquity* 88, 342, pp. 1132-47.

SUNDELL, T. (2014). *The past hidden in our genes. Combining archaeological and genetic methodology: Prehistoric population bottlenecks in Finland.* Doctoral Dissertation. Department of Philosophy, History, Culture and Art Studies. University of Helsinki.

TALLAVAARA, M., P. PESONEN & M. OINONEN. (2010). Prehistoric population history in eastern Fennoscandia. *Journal of Archaeological Science* 37,, pp. 251-260.

TALLAVAARA, M. & H. SEPPÄ. (2012). Did the mid-Holocene environmental changes cause the boom and bust of hunter-gatherer population size in eastern Fennoscandia? *The Holocene* 22(2), pp. 215-225.

WANG, W., YANG, J., & MUNTZ, R. (1997). "STING: A Statistical Information Grid Approach to Spatial Data Mining". In: M. Jarke, M. J. Carey, K. R. Dittrich, F. H. Lochovsky, P. Loucopoulos, M. A. Jeusfeld (Eds.). *VLDB'97, Proceedings of 23rd International Conference on Very Large Data Bases, August 25-29, 1997, Athens, Greece.* Morgan Kaufmann, pp. 186-195.

WANG, W., YANG, J., & MUNTZ, R. (2000). "An approach to active spatial data mining based on statistical information. IEEE Transactions on Knowledge and *Data Engineering* 12(5), pp. 715-728.

Multi-Agent Modelling of the Neolithic LBK

Jean-Pierre BOCQUET-APPEL,[1,2] Richard MOUSSA,[1] Jérôme DUBOULOZ[3]

[1]CNRS, UPR 2147, Paris, France
[2]Ecole Pratique des Hautes Etudes (EPHE, Paris) France
[3]CNRS, UMR 8215, Trajectoires Nanterre, France
jean-pierre.bocquet-appel@cnrs.fr, richard.moussa@gmail.com, jerome.dubouloz@mae.cnrs.fr

Abstract

A modelling approach is presented in which archaeological data on the first farmers in Central and Western Europe, called the LBK (5600–4900 cal BC), are cross-analysed with the corresponding environmental data, at a geographical scale of 1.9 million km² (pixel). The purpose of this approach is to simulate the geographical and demographic expansion of the LBK culture and to gain insights into the responsiveness and resilience of its socio-natural system to climatic impacts, and into its possible dissolution. This is done through multi-agent modelling of the LBK socio-natural system, based on intermediate, decontextualised anthropological and ethno-historical partial models and with input from palaeo-environmental and bio-archaeological disciplines, cultural archaeology, social anthropology, palaeodemography and economics. To reduce the effect of stochasticity in the historical model, which is its one intractable weakness, the modelling centres on the infrastructure, which is constrained to the longue durée of geohistory. The partial simulation results presented here show the resilience of this socio-natural system, where demographic growth sporadically slows due to the impacts of historical environmental collapse.

Keywords: LinearBandKeramik LBK, Multi-Agent Modelling, Decontextualised Anthropological Partial Models, Decontextualised Ethno-Historical Partial Models, Stochasticity, Longue Durée, Geohistory.

Introduction

The LBK was the main Neolithic agricultural society in temperate Europe. It formed in Central Europe at around 7500 BP (5550 calBC), on the pioneer front of the Balkan Neolithic cultures, which had become established earlier (Bánffy and Oross, 2010; Pavùk, 2004), and extended longitudinally across Europe from Hungary to Normandy and to Poland, Ukraine and Romania. The first aim of our research was to build up a model of the simplified functional patterns of the socio-natural LBK system, that is to say the system that enabled the agricultural society to generate livelihoods and to ensure its reproduction against a background of strong demo-geographic expansion of that society, until its expansion ceased before the dissolution of its cultural area. Agricultural societies are extremely vulnerable to climate hazards. Famines have regularly hit proto-historical and historical agricultural societies, and the possibility that such crises occurred over the 600 years of the LBK culture is a serious hypothesis. The second aim was to use this model to combine the archaeological data with the corresponding palaeo-environmental data in order to assess the responsiveness and resilience of the LBK to environmental hazards. This can be done by constructing computer models of artificial societies and running simulations according to reconstructed natural contexts.

Agent-based historical simulation models raise a great many questions (see Barton 2013; Kohler, 2012; Lake 2010, 2014, 2003; McGlade, 2014; Premo, 2007, 2010; Clark and Hagemeister 2006, Wobst 2010). The intractable problem of historical multi-agent simulation is stochasticity (e.g. McGlade, 2014), which makes such simulations improbable even though they produce so-called emerging phenomena. Our modelling approach therefore focused on the least stochastic fraction of total History, represented by structural variables over the longue durée of geohistory (Braudel, 1958). Our modelling approach follows the line of experimental archaeology (Axtell *et al*, 7275).

After a brief presentation of the theoretical framework of Obresoc modelling and of the environmental data used, this paper outlines the main modules of the general Obresoc model. From the standpoint of the anthropological approach, the questions needing answers are: What is the sensitivity of the agricultural system and the population to climate variability, in terms of the frequency of food shortages and famines? Are there any detectable climate impacts that may have premised or accompanied the dissolution of the LBK cultural area? From the modelling standpoint, the question raised by Joshua M Epstein (2006: 15) remains: Does the hypothesized microspecification suffice to generate the phenomenon observed? The simulation results presented in this article concern only the variation of simulated aggregated variables over time. Their spatio-temporal variation and the emergence of a geographic pattern will be presented in a volume describing the Obresoc approach as a whole.

1. Reducing the stochasticity of the historical model

History can be seen as the continuous flow of human activity, both collective and individual, since the origins of humanity and across the entire planet. In all societies, with or without written records, these human activities produce traces that, if kept, become the historical archives of humanity. From this raw material, historians produce narratives of history. But the flow of information, possibly increasing in magnitude and heterogeneity with the size of the population, makes it difficult to integrate in narrative terms. In the last century, for reasons to do with ordering information for ease of analysis, the historian Fernand Braudel and the Annales School of ESC (Economies, Societies and Civilizations) found it was necessary to subdivide History into three durations in time: long, medium and short (Braudel 1958). The general flow of information over time, expressing all (georeferenced) human activities, represents the total History. Braudel and the Annales School devised a methodology to obtain kind of spectral decomposition of historical information into qualitative phenomena distributed into three nested classes of shorter and shorter durations, whose sum makes up total history. Phenomena of different durations can interact or not. This model for the spectral analysis of history was extended to ancient archaeology nearly 25 years ago by Bintliff (1991), in analytical and narrative terms (see e.g. Sweitz 2005).

This subdivision is useful in archaeological modelling when it has to resolve the two great mysteries of history: stochasticity and contingency. Stochasticity, as we know, is the random variation of any phenomenon whose law of distribution is known; contingency, in the sphere of human activities, is the unpredictability of their succession. An example of such successive historical sequences is that given by Thucydides in the 5th century BC, in his famous History of the Peloponnesian War (Thucydide, 1936 and 1948). The turbulent flow of the history narrated by Thucydides would be unknowable from the Hellenic archaeological remains alone, however vast and numerous. It should be noted that the effects of stochasticity and contingency, which make any prediction of human activities random, are not the same in all three of Braudel's durations. If, in a Braudelian analysis, we leave aside what pertains to civilizations, world interpretations and mentalities, we see that the long duration of geohistory concerns that which ensures the biological survival of societies, including techniques and social modes of production, which can be grouped together by using the Marxist term of infrastructure. The infrastructure is highly constrained by biological and geo-environmental data. For ethnographic societies (with no written records) we can add social structures (clans, matrimonial regimes) and, perhaps, institutions. Social structures begin to be visible during Prehistory (see Bentley 2007). We see that the long-term trend of History provides the flow of human activities which is least subject to contingency. It is on this fraction that concerns infrastructure, and its consequences in terms of demographic and geographic colonization, that Obresoc modelling is mainly focused. The purpose of this model equates to addressing an issue concerning human expansion in the universe, where the question would be: given a specified economic infrastructure and reconstructed environmental conditions, could humans colonize Mars? Thus, looking beyond the insoluble question of modelling total History, the archaeological modelling tool applied to reproduce the long-term trend continues to demonstrate its unique advantages in terms of non-linear interactions and retroactions between categories of unsubstantiable structural variables, and thus in terms of practical transdisciplinarity (see Epstein 2006). Failing actual experimentation, such practical transdisciplinarity is inconceivable without computer simulations.

Based on structural geo-historical data on the one hand, and on an intermediate socio-natural multi-agent model on the other hand, what was the potential of LBK infrastructure in terms of colonization between 5550 and 4950 cal BC? By structural data, we mean the geological fertility of soils, altitude, the reconstructed original vegetation cover, recurrent climate simulations and estimates of the corresponding soil fertility in the expansion area. The socio-natural model incorporates sub-models representing an agro-pastoral production system, a structured household and population dynamics, as well as reference archaeological data on agro-pastoralism. The set of sub-models is inferred from ethno-archaeology. On the variable geo-historical structural background, the goal of Obresoc modelling is to cause the 'household' agents to act through mechanical interactions between the different sequences of sub-models and to observe their traces, i.e. the archaeological effects, in terms of spatial and temporal distributions of households and hamlets whether living or dead. Without computer simulations, it would be hard to imagine how overall effects could be discerned. The fit between the Obresoc model and the LBK data was measured by visual inspection, via the degree of coincidence of simulated expansion areas with the observed expansion areas at six consecutive archaeological points in time. The area located east of longitude 23.5 ° E was excluded for reasons of lack of access to comprehensive documentation.

2. The basic environmental data

Fertility of Neolithic soils

We followed the hypothesis that if the LBK soils cannot be exactly reconstructed, the nature and the distribution of the main pedological domains did not undergo a notable evolution during the Holocene. Indeed, at the spatial scale of temperate Europe, these pedological domains are driven by a series of stable physiographic factors, the most primary of which are i) the lithology and ii) the climatic nuances. The modern European soil map was thus simplified in 13 classes and then considered from a 'fertility' point of view through different criteria : (i) soil thickness and granulometry, with variable available soil resources, (ii) average chemical fertility (acidity, potential richness in cations), (iii) soil saturation level, with variable aeration and plant asphyxia conditions. A map of relative fertility was produced including 5 main classes ranging from the most to the least fertile soils (see : Schwartz, Ertlen, Davtian, Dubouloz, Gauvry, Vyslouzilova, Berger, in press).

Climatic variations

To remedy the lack of high-resolution hydrological and palaeo-ecological data, a new archive of palaeo-channel sediments was used (Shriek, Rhineland-West, Germany) to conduct an integrated high-resolution multi-proxy analysis. The Modern Analogues Technique (MAT) was used to reconstruct (calibrate) annual, summer and winter temperatures and precipitations at this place between 5600 and 4800 BC (Berger, Ortu, Moussa and Toonen, in press). These results served to calibrate the WorldClim data for the area under study (http://www.worldclim.org), which was used to simulate the spatial variations of temperature and rainfall during the LBK period.

Climatic thresholds

Climatic thresholds analogous to those for current extreme variations in the same region were used. Based on an analysis of the Strasbourg meteorological data (http://www.infoclimat.fr), these thresholds for seasonal temperature and rainfall were considered as positive and negative deviations from the mean and applied to the WorldClim gradient. In this way, these climatic thresholds represent values for humidity, drought, heat and cold related to the mean seasonal climate of each ‘pixel’ in the simulation.

3. The Obresoc socio-natural model

A model of the multi-agent type (MAM) was built up from ethno-archaeological inferences on the palaeoenvironment together with partial intermediate, decontextualised anthropological and ethno-historical models. The MAM model includes partial intermediate models of the following:

An LBK house

This is an individual farm, representing a household typically composed of an extended family divided into nuclear families (named also ‘Eastern European family type’ in historical demography: Polla 2003, 2006). The household is the agent in the model. Several households (LBK houses) in the same km² pixel form a hamlet.

An agricultural system

This includes cereal crops and livestock (cattle, goats / sheep and pigs), supplemented by hunting and gathering and dairy products; economics (cultivation on the best fertile patch, initial slash-and-burn in primary forest followed by intensive farming using manure, expected production/consumption ratio in equilibrium with the household demography plus a small food security margin). The economic output as a whole is constrained by the available workforce. When food production does not cover needs: i) due to a bad harvest, in which case intensified gathering will tend to compensate for the shortage of grain, up to the maximum level of output allowed by the available workforce; ii) due to insufficient meat production, in which case hunting will tend to compensate for the shortage of meat, with a weight of game determined by the number of hunters, time available, local game density and random game capture. Hunting tends to compensate for livestock as long as the size of the household’s mini-herd has not reached the human/animal equilibrium level through annual population growth. Dairy products form a general diet supplement.

A relationship between climate change and performance of the agricultural system

Crops and livestock are sensitive to climatic variations. These positively or negatively affect the capacity of different soils to support vegetation. This fertility factor oscillates, depending on the season on the climate curve, between hydromorphy and drought according to combined temperature and rainfall limits, which are determined by comparisons with hydrological estimations/simulations. Yields from cereal crops and herbaceous plants in general therefore vary as a proportion of these variations in ‘fertility’.

To determine the quality of a patch / pixel, a score was determined by regressing the dummy ‘site’ variable (coded 1/0) indicating the presence / absence of an archaeological site from a database of 6575 geo-referenced sites, against a subset of real or estimated geoclimatic variables (7 on 20). This subset was selected by stepwise regression from a large sample of pixels on the map. The geographic distribution of the best patch score is shown in Figure 1. Due to the low resolution of the European soil-map, mid-range valleys could not be evaluated as potential best patches. This is a problem particularly in the West, where Lorraine and Champagne, two areas with a largely attested LBK occupation, do not offer good best-patches on the map. When looking for a locality to settle, the agents of the model will thus prefer, contrary to the historical reality, to target westerner pixels in the Paris Basin. In this way they will tend to accelerate the expansion towards western France.

Beyond the routine situation, extreme events are considered: very wet and cold autumns, icy and snowy very long winters, excessively wet or dry springs, very hot and dry or unusually wet summers. These meteorological events, which can seriously affect the survival of humans and animals, are simulated by setting the seasonal temperature and/or rainfall beyond the limits required by the biological characteristics of plants and defined by historical investigations.

Population dynamics

These correspond to the inputs/outputs in the population, as well as marriages, with a mechanism governing the demographic expansion and intensification of the settlement. Demographic data (mortality, fertility) are

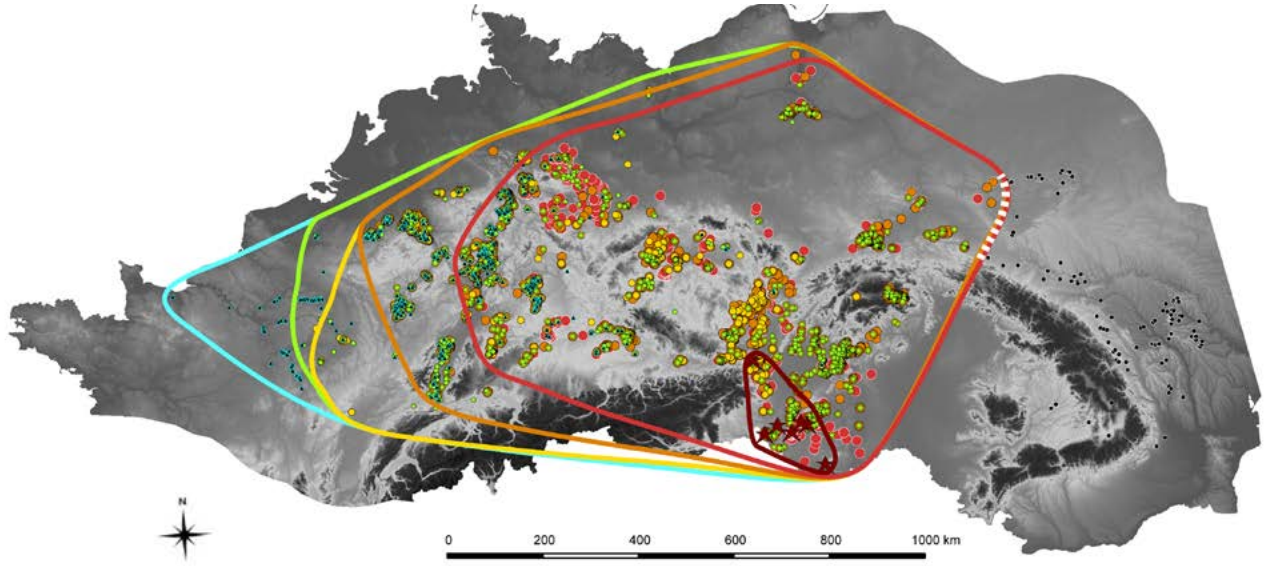


Figure 1: Geographic distribution of the best patch score in the LBK area. The scores are in a colour scale, from green (highest score) to brown (lowest score). Continuous dark brown is on the area not colonized by the LBK.

taken from historical preindustrial demography in the same central European region. Structural remains of residential Mesolithic hunter-gatherer (HG) settlements, which could reflect a certain population density on contact with the LBK expansion, are few in number (Jochim 2008), suggesting a very low HG population density. The essential cause of this very low HG population density is that the density of secondary biomass, i.e. ungulate prey (deer, aurochs, wild boar), is also very low in the moderately dense open mixed oak primary forest (Kreuz 2008) that was colonized by the LBK (from which is deduced a minimum population estimate of about 50,000 people over 1.9 million sq km; see also Rozoy 1978).

It has long been known that LBK colonists chose their settlements, in preference to all other seemingly possible settlements, first of all on criteria for optimum farm production (Childe 1929), accessibility and comfort, which can be summarized as choosing the 'best patch', a population ecology concept (Sutherland 1996) introduced by Shennan (2009) for LBK settlements. Other 'attractors' have also probably weighed in colonization directions and management, including access to certain critical resources like flint, hard rocks and salt. But their modeling has not yet been made. LBK settlements were mainly established in forest clearings in primary forests of sparse oak, where the population density of hunter-gatherers was very low, as indicated above. It was not against a background of land shortage that the LBK expansion took place, either during the initial phase of the pioneer settlement or during the subsequent increase of the population, but rather in a context of apparent land abundance. Against this background of abundant land, few anthropological assumptions of a non-economic nature have been put forward to explain the relatively fast initial geographical expansion / dispersion of LBK settlement. The assumption we made is that of dispersion as a mechanism for avoiding interpersonal conflict as population density increased in the immediate vicinity, also called 'scalar stress' in social anthropology (Carneiro 1967, 1978; Johnson 1982; see also Rappaport 1968). Population saturation of the best patch in the Obresoc model, or the limit of its carrying capacity, is not determined by economic saturation but is the result of 'scalar stress'. When scission / dispersion occurs (see below), the new migrant domestic household will be looking for a best patch within its radius of geographical perception.

In a nuclear family in an agrarian community, a specific and long identified source of tension is the pressure of consumer mouths to be fed (C) on the activity of producers (P) (Chayanov 1923). The larger the number of consumers, the greater the amount of work that the producers must provide, causing pressure on producers that generates the psychological strain of so-called 'scalar stress' among members of a population unit. The physical and psychological cause of household splitting is the refusal to participate in the additional workload engendered by the nuclear family of the younger child, relatively to the work required to sustain the elder's nuclear family. Counter-intuitively, Chayanov's scalar stress in an extended family and the ethno-historical data on colonizers both select as pioneers, not young single people or young newly married first-time parents about 17-25 years old, but nuclear families made up of able bodies rather than mouths to be fed: typically, mature couples whose years of childcare are nearly over, with 2-3 adolescents aged 15 + and no young children. In the model, the strength of Chayanov's scalar stress is measured by the C / P ratios within nuclear families of an extended family. When Chayanov's ratio C / P = 1, then there are as many consumers as producers in the nuclear family and tension within the nuclear family is at a minimum. When C / P > 1, then there are more consumers than producers and Chayanov's stress rises within the nuclear family. But this stress also weighs on the other nuclear families in the extended family, especially on nuclear families that are now composed of producers only. These families do not want to provide additional work for those who still have immature children. In this case, when Chayanov's ratio for a nuclear family is C / P = 1, it splits away from the household, with its agricultural possessions of crops and seeds in demographic proportion. These splits occur in late summer, when the food and seed harvest is over. The last nuclear family in an extended family stays with the founding parents if they are still alive.

Emergence of a headman at the hamlet scale

As the population increases, influential figures begin to appear, known as Headmen or Big Men, along with a segmentation of society by lineage and the formation of clans and sub-clans based on kinship (for a summary see: Lewellen 2003; Wiessner and Tumu 1998). In accordance with the logic of feasting (Dietler and Hayden 2001), a hamlet with a headman has two advantages in the model: storage and demographic attractiveness. Where storage is concerned, each household in a 'big man hamlet' can accumulate the potentially uneaten portion of its annual harvest, including the security margin, for up to three years at a stretch. Households in an ordinary hamlet (with no headman) do not store produce but consume their entire harvest. The effect of storage is to enable households in a 'big man hamlet' to act as shock absorbers in the event of an agricultural food crisis, which benefits others in their neighbourhood. While the effect of storage in a 'big man hamlet' is to buffer food crises in its neighbourhood, it also strengthens forms of clientelism.

One of the underlying issues in the emergence of politics through a headman is the growing number of supporters of the latter, meaning that the headman will eventually become a demographic attractor. The issue here is one of political demography. Because the personality traits driving the appearance of a headman in a locality could not be modelled on an individual level, the rule was to consider that when a hamlet's population reached an unusual or critical size, with an unusually high number of households (houses), it would become a 'headman' pixel/hamlet. The critical level is located in the upper 5% of the cumulative probability density of households/houses in the hamlet, which is itself determined by a Poisson function with an archaeological parameter.

An ethno-historical model of social dislocation caused by a short-term environmental crisis producing a famine

A typical pattern of responses to famine was formalized from ethno-historical and de-contextualized sources (Corbett 1988; Cutler 1984; Hartog 1981; Lachiver 1991). To summarize, the pattern is expressed in three stages: 1) an insurance mechanism (reduction of the current consumption level, transfers of dependent individuals – children and the elderly – between related or neighbouring households; 2) disposal of productive assets (animals, agricultural tools, land); 3) deprivation: distress migration of the entire household if it still exists. Of these three stages, the second was omitted due to the very low technical level of LBK agriculture (digging sticks, adzes) and the plentiful land available.

4. Modelling dynamics

A household agent must:

- ensure economic production to meet the household's energy needs (kcal), which are determined by its demographic structure;
- ensure reproduction of the population (by finding spouses);
- cope with Chayanov's scalar stress, formalized by the ratio C / P (see above), possibly by splitting up the household. This occurs when a nuclear family departs, which in turn determines the construction of a new house;
- cope with Johnson's scale-related-social-stress, formalized by a density-dependent probability of settling together (see above) if a household should split up. Then either this new nuclear family will be 'allowed' to densify the hamlet or will colonize / densify a region. Within a radius of perception, the new nuclear family settles in the pixel most favourable to agricultural development ('best patch');
- respond, if necessary, to a variety of typical climatic crises, either stochastically local or systemic on a larger geographical scale (regional to subcontinental) that have generated famines in experimental historical contexts in Europe.

The LBK population 'grows and multiplies', densifying existing hamlets or establishing new hamlets in the geographical directions and areas of the zones most favourable to the agricultural system determined by the 'best patch' areas, via Chayanov's and Johnsons's scalar stress. The LBK society expanded geographically, driven by population growth and regulated according to the social rules already described. The colonizers moved from a pioneer front to settle territories, extending their area of expansion towards the best patch and densifying the population in the area already occupied.

5. Partial results and concluding remarks

To date, exploration of the Obresoc model is at an early stage. This is why the results of the project are still incomplete and mainly orientated towards the question of the sensitivity of the agricultural and demographic model to reconstructed climatic variations.

First of all, sixteen simulations with artificial mean temperatures and precipitations displayed around the reconstituted continental means, from one extreme (cold, wet) to the other (heat, drought), showed the sensitivity of the intensive agricultural system to long periods of environmental conditions that are unfavourable to agriculture (especially hot/dry and cold/wet). These adverse environmental conditions can reduce the final population size by a factor of 6 within 600 years, relatively to its optimum size under favourable agricultural conditions, with consequences in terms of decelerating geographical expansion and the development of social complexity (not shown).

A sensitivity analysis (SA) against a mean climatic restitution was then performed with the simulated output variables (demographic, economic and anthropological) to assess their sensitivity to changes in the values of key parameters used as input (Cereal intake: 65-75 %, manure, cows/ha/yr: 1-10 cows, yield, kg/ha: 500-1500 kg, cow/human ratio: 0.5-2.0 cows). The SA (not detailed) produces a hierarchy of parameters with a potential impact on the space-time expansion of the cultural area and LBK populations: it shows the major role in the Obresoc model of manure, whose variation influences over 70% of the variation in agricultural, demographic and social variables, but has no influence on hunting and gathering.

Aware of these two systemic characteristics of the Obresoc model, its exploration in a situation of historical climate simulation endeavoured first of all to define the conditions of parameters enabling an approximate reproduction of the spatial expansion attested archaeologically, as well as its tempo. As a first approach, the degree to which the Obresoc general model fits the observed data was measured by visual inspection of the geographical dispersion of simulated localities relative to the dispersion of archaeological sites, via the convex hull polygon to determine the limits of expansion. This inspection concerned the general expansion, at the end of the simulation, as well as the various stages defined by the typochronological periods subdividing the total duration of the LBK. The areas of dispersion, over the six periods, of both the observed archaeological sites and the presently simulated localities in the 'best' model are shown in Figure 2a and 2b respectively.

This is a model that corresponds to a population of up to one million two hundred thousand people at the end of the simulation. Probably because of the lack in the model of other attraction than farming advantages and certainly because of the approximations of the best patch map in the West (see above), the tempo of the expansion is a bit slow at the begining of the simulation and a bit fast at the end.

When looking now, in figure 3, at the distribution of the household-types produced by this model -i.e the distribution of mono- and multi-families households- we can recognize a fairly good correspondence with the distribution of the four main archaeological house-types, represented by the number of inner rear-parts in a sample of around 700 archaeological house-plans. Under the hypothesis that the number of rear-parts in a LBK house is related to the number of nuclear families sheltered under its roof, from mono- to multi-families households, the correspondence between simulated households and observed archaeological houses possibly gives us a new interesting way of evaluating the 'best' model.

These two broad comparisons between simulated facts and archaeological facts suggest that the adjustments of the retained simulation produce an expansion that is quite close to the one we are trying to model. We will now consider some details of this reconstituted trajectory, mainly from the point of view of the impacts of climate variations.

Figure 4 presents the temporal distribution of the simulated climatic crises. Each year of the simulation cumulates the number of occupied pixels that have been subjected to a particular pattern of weather. The latter corresponds to a

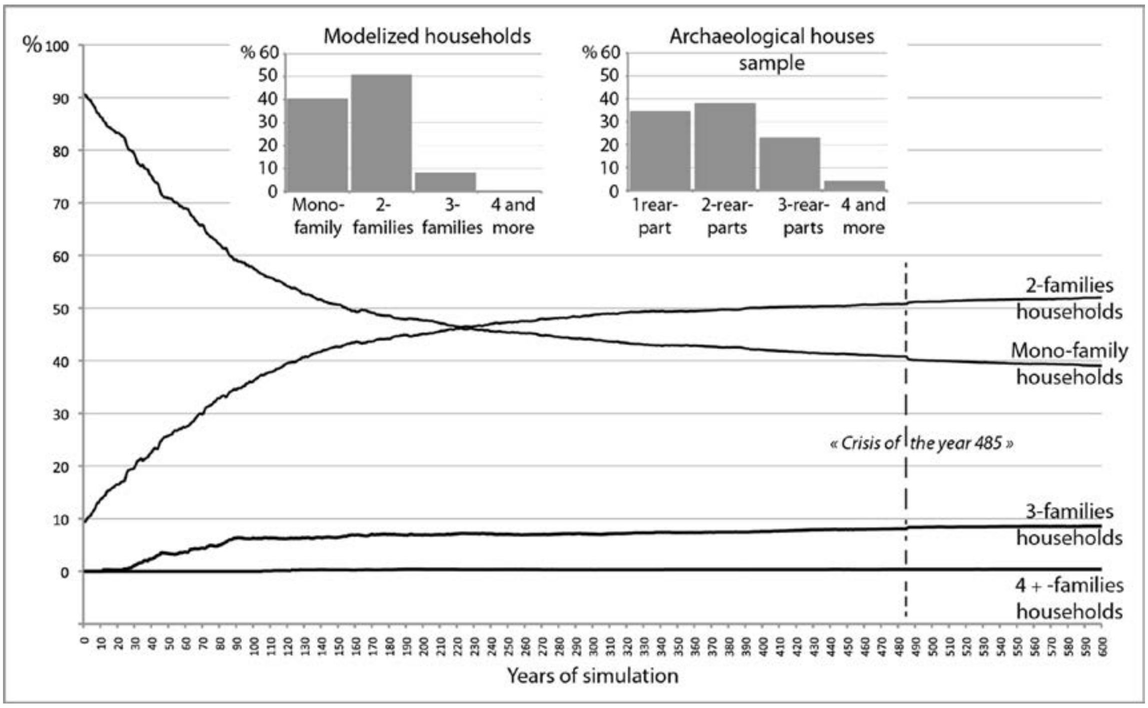


Figure 3: Temporal distribution of household-types and comparison of the histogram of cumulated values with the sample of archaeological houses.

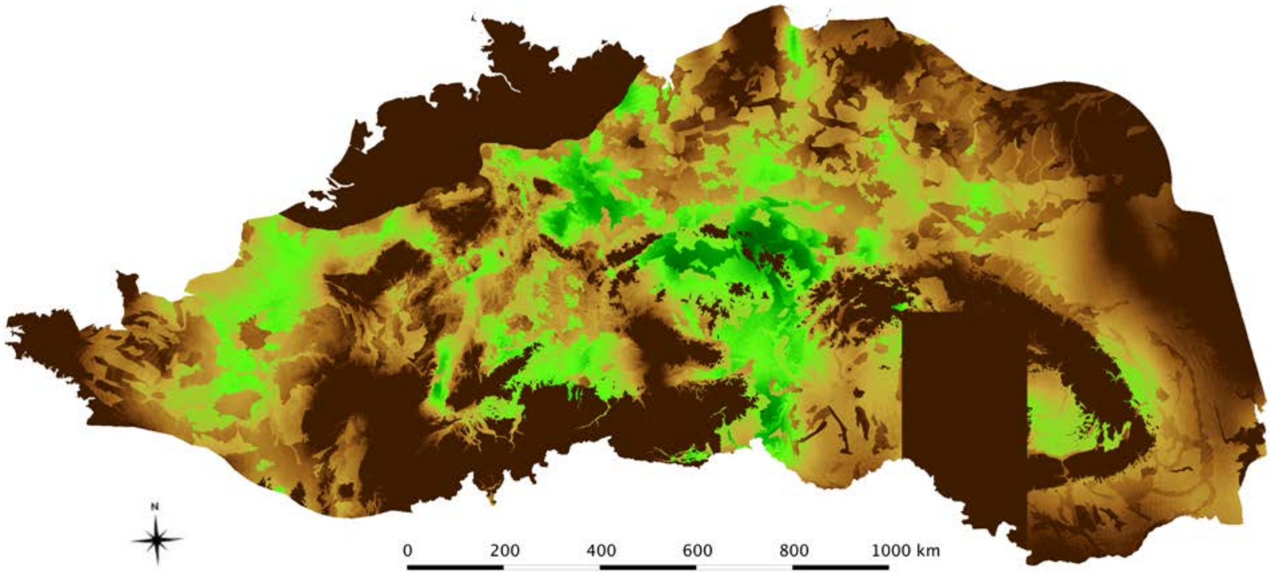


Figure 2a: Areas of the convex hull polygons, superimposed onto each other, representing the dispersion of observed archaeological localities at the end of the five main LBK periods (from dark red- to light blue, at: ca 5550, 5300, 5175, 5125, 5025, 4950 BC). For reasons of difficulties in accessing archaeological data the eastern part of the LBK area (23,5°E and beyond) is not taken into account.

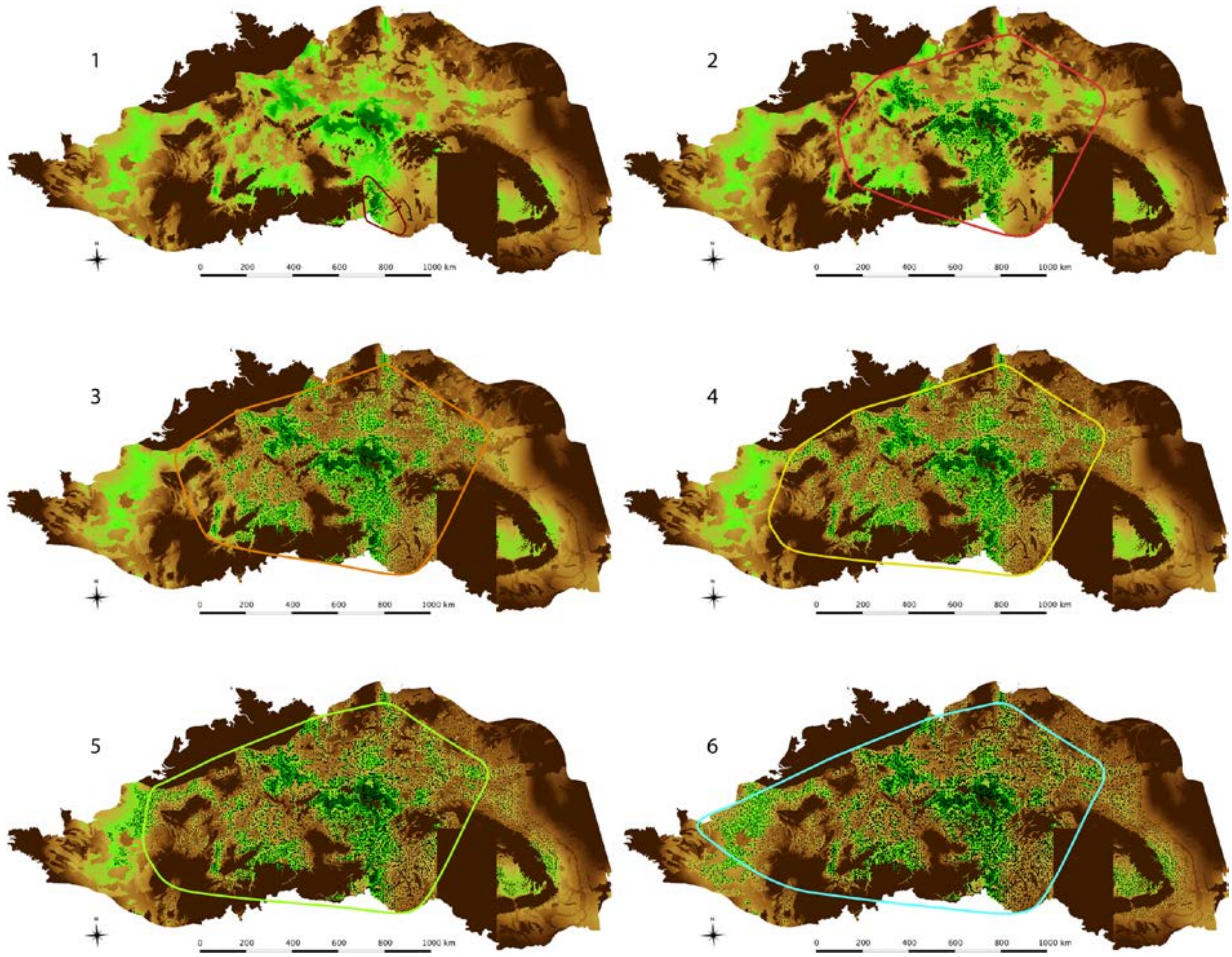


Figure 2b: Dispersion of simulated localities at 5550 BC, 5325 BC, 5175 BC, 5125 BC, 5025 BC, 4950 BC against the convex hull polygons representing the LBK dispersion.

succession of several consecutive seasons with rainfall and average temperatures above or below thresholds unfavourable for agriculture, according to a typology developed from research by Leroy-Ladurie (2003, 2004) on the Little Ice Age.

It appears that only one major crisis is produced by the model, around the 485th year of the simulation (i.e. in terms of the historical chronology of the simulation around 5070-5060 BC). This 'date' corresponds to a period with the coldest and driest winter and spring of the whole sequence. During the same year, summer is wetter and colder than the mean and one of the wettest and coldest summers. One observes that the death rate in this particular 'year' is exceptionally high and that the number of households decreases significantly. One century later, a similar weather pattern apparently produces the same effect, although this is less intense and widespread.

Figure 5 zooms in on the end of the sequence in figure 4, to present the timing of climate and food crises, as well as the temporal distribution of annual numbers of deaths and households. This shows that the main climatic crisis is followed, the next year, by a severe famine and then a period of shortage which bring about the temporary increase in deaths. One also observes that the lowest point in the number of households comes a year later than the highest point in the death rate. This three-year sequence – weather crisis, food and demographic crisis, population crisis – confirms the model's coherence and its overall sensitivity to climate variations.

Figure 6 examines the details of the food crisis and its tempo. Thus, following the weather crisis, the first economic domain affected is agriculture (represented by the average net crop per household, after loss due to harvesting and dehusking of hulled cereals). Probably after a year when agents had to kill high numbers of livestock and reduce their capacity to manure fields, the amount of cattle meat in diet decreases substantially the following year, thus producing additional scarcity. This fall in the availability of cattle meat and cereal products then leads to an increase in hunting and hazelnut gathering. A return to the pre-crisis situation of normal consumption only becomes apparent 6 or 7 years later, with a population that has been cut back by about 20%.

The outputs of the above simulation gradually incorporate space from one year to another. The spatial and social dimensions of the modelled facts, produced during the simulation, have yet to be programmed for output, due to the increase in calculation time caused by storing this information. The relative slowness of each simulation would have been in fact incompatible with preliminary exploration of the model. These outputs, delayed for practical reasons, are now being programmed and should enable an assessment of the spatial component of the simulated phenomena and their effect at individual household scale. For the moment, only a few spatial characteristics are available, concerning the location of weather crises and the composition of hamlets/pixels in terms of population, household and families.

Thus figure 7 maps the hamlets/pixels that underwent the 'year 485' (simulated 5065 BC) weather crisis.

The wide geographical scope of the phenomenon can clearly be seen, as well as vast zones that remain untouched. This spatial variability shows that the climate model does integrate the geographical variability established by the WorldClim database. Amongst the regions affected by the weather crisis, a zoom on central Germany, corresponding to one of the best patches in the simulated LBK territory (Saxony, Saxony Anhalt, Lower Saxony, Thuringia), enables a finer analysis of the population component of the phenomenon.

Figure 8 shows for this region a demographic decrease of around 20%, still perceptible 15 years later, accompanied by a drop in numbers of occupied hamlets/pixels, living households and nuclear families.

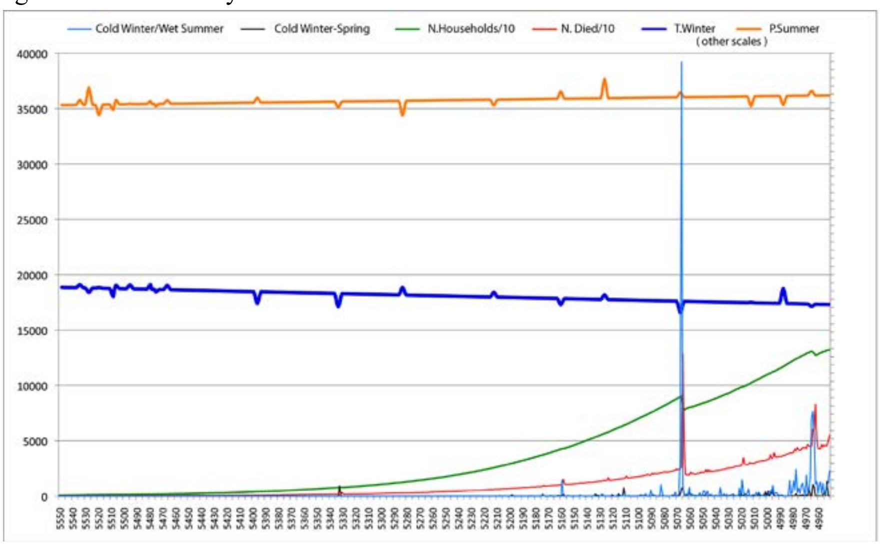


Figure 4: Temporal distribution of climate-weather crises (Cold-Winter/Wet-Summer and Cold-Wint-Spring) along the simulation, in relation to the average continental climatic curves (Twin and Psumm), the number of deaths and the number of households.

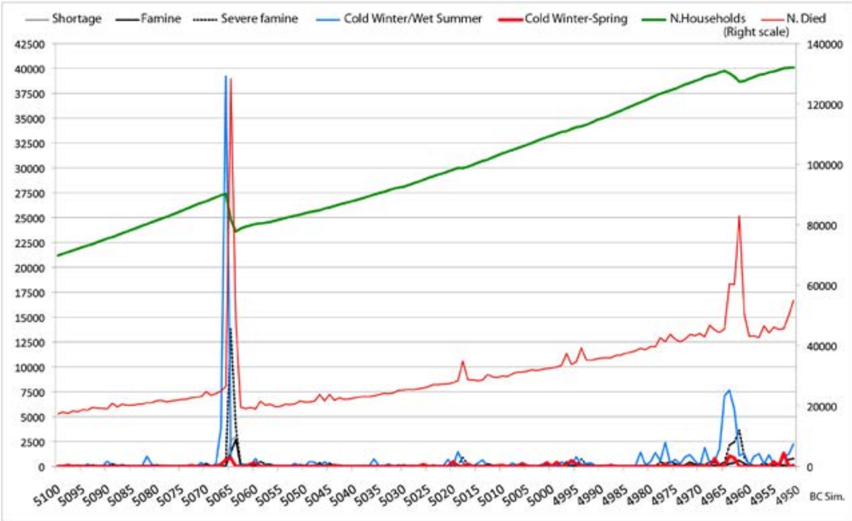


Figure 5: Detail of the end of the simulated sequence: same variables as in figure 4, to which are added the distribution of the number of households suffering from shortage, famine and severe famine.

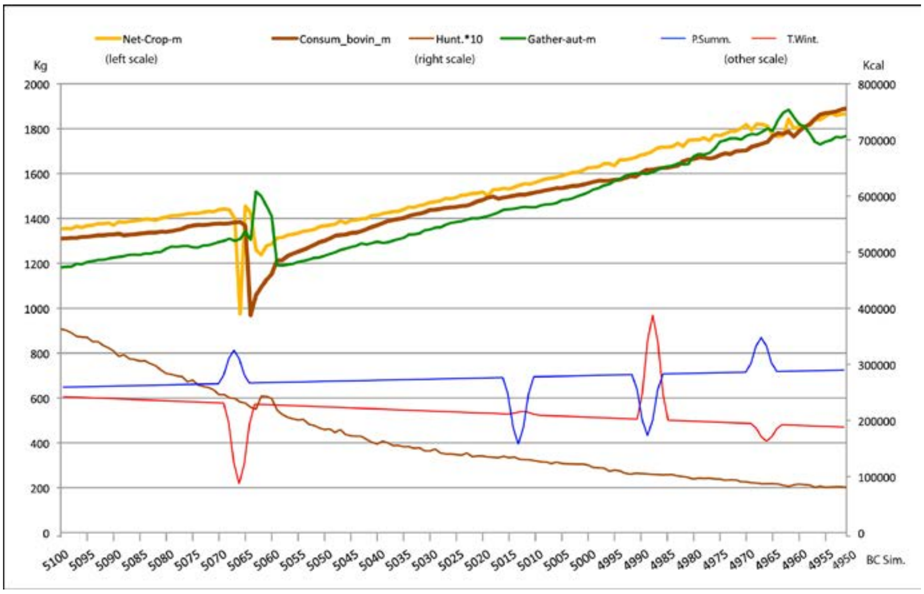


Figure 6: Temporal distribution of values for average net crop per household (in kg), intake in Kcal of domestic cattle meat, hunting and gathering, in relation to the climate curve in the same time window as figure 5.

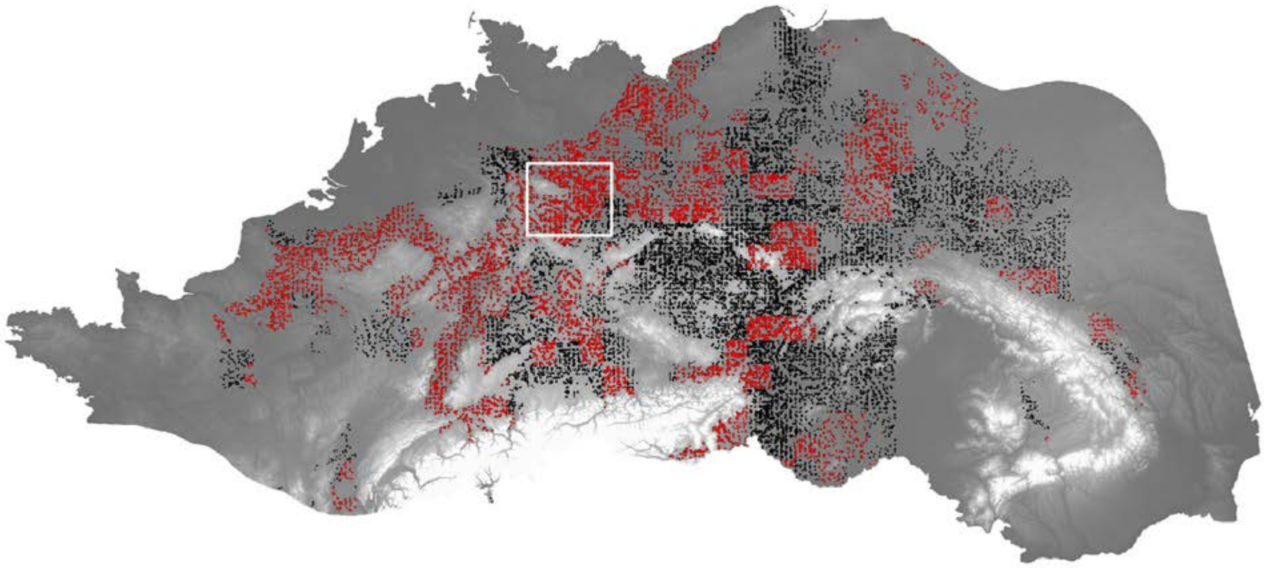


Figure 7: Distribution of hamlets/pixels undergoing food crisis (in red) and location of the rectangular zone tested in detail (central Germany).

One notes, amongst the two simulated products that can be evaluated archaeologically, that the number of households (-25%) more strongly reflects the demographic problem than the number of occupied hamlet-pixels (-6%). The archaeological identification of such a crisis would therefore require reliable data on the distribution of households, not only at site level but especially at regional level.

Although it is not possible at present to portray the finer detail of the economic trajectory followed by the agents of this settlement zone, there is no doubt that the logic described in figure 6 is involved. We can just notice that the simulated crisis seems to have a differential effect according to the different types of households (from mono- to multi-families), as it can be seen in figure 3, and also to the different types of hamlets (from single-farm to villages of nine and more households). The latter is shown in figure 9 where a redistribution of the number of households by hamlet is noticeable during the weather crisis. Mid-range hamlets (from four to eight aggregated households) are indeed proportionnally more hit by the crisis than single-farms or big hamlets.

With a more detailed information output, per household, we would be able to access the variability of situations depending on household size and age, demographic composition, cluster density, the presence or absence of a leader-hamlet, etc… It is under these conditions the Obresoc model should be able to reveal its full potential, at the same time with better definition of its limits and of the improvements required.

As a concluding remark, it appears that the temporal simulation of the Obresoc economic, demographic and anthropological model, through a climate simulation based on prehistoric and present-day observations, gives some indication of the fragility of the system in the face of pronounced climate and weather events. One also notes its resilience if these events are just occasional. Under the conditions of the model and its climate configuration, the crisis produced here strongly disturbs the system without having a profound effect on its development. Due to the timing of the crisis, triggered after the geographical expansion was mostly completed, the effects of this kind of disturbance on the whole geographical distribution cannot be evaluated. This can only be done on a local or regional scale. What would happen if such crises occurred in close succession over a timescale of just a few years or decades ? The magnitude of the effects observed in this simulation suggests that a close succession of crises would have a major impact on the system, as predicted by models of climate and environmental destabilization in simple agricultural societies (Pfister and Brádzil 2006). The history of the LBK in certain regions provides evidence, precisely towards the end of the 6th millennium, for reduction in settlement density and for marked cultural, political or ideological transformations (Gronenborn 2007 ; 2010). The roots of these phenomena could partly lie in a succession of climate and environmental crises, which should now be programmed as a scenario for the simulation so that their effects can be measured. The Obresoc tool enables this kind of experiment.

Acknowledgements

This work has received support from the French National Research Agency with the reference ANR-09-CEP-004-01/ OBRESOC (JP Bocquet-Appel, coordinator).

References

BÁNFFY, E., OROSS, K., (2010). 'The earliest and earlier phase of the LBK' in Transdanubia. In: D. Gronenborn, J. Petrasch (eds), *The spread of the Neolithic to Central Europe International Symposium, Mainz 24-26 June 2005.* Mainz, Römisch-Germanisches Zentralmuseum: RGZM-Tagungen 4, pp. 255-272.

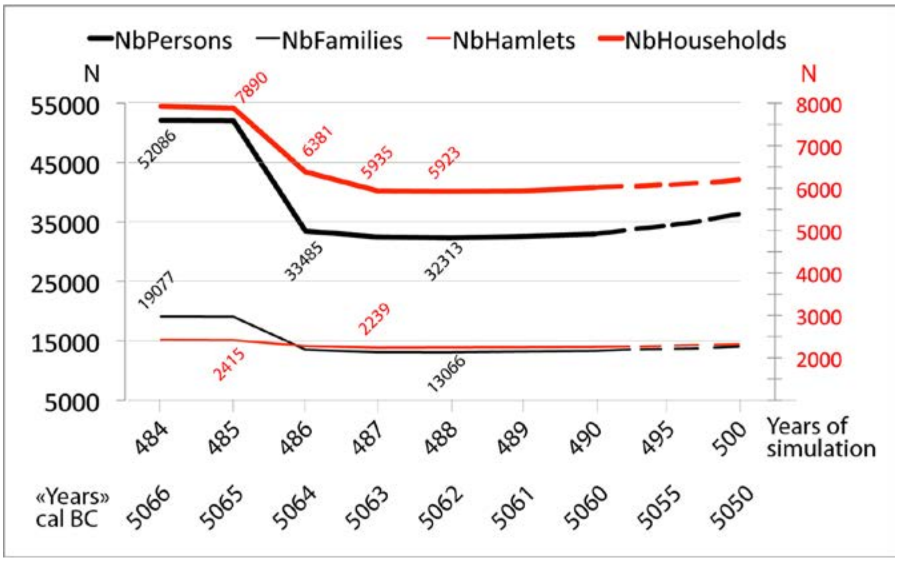


Figure 8: Temporal distributions of numbers of inhabitants, nuclear families, households and hamlet-pixels occupied before and after the '5065' crisis (485th year of the simulation) in central Germany.

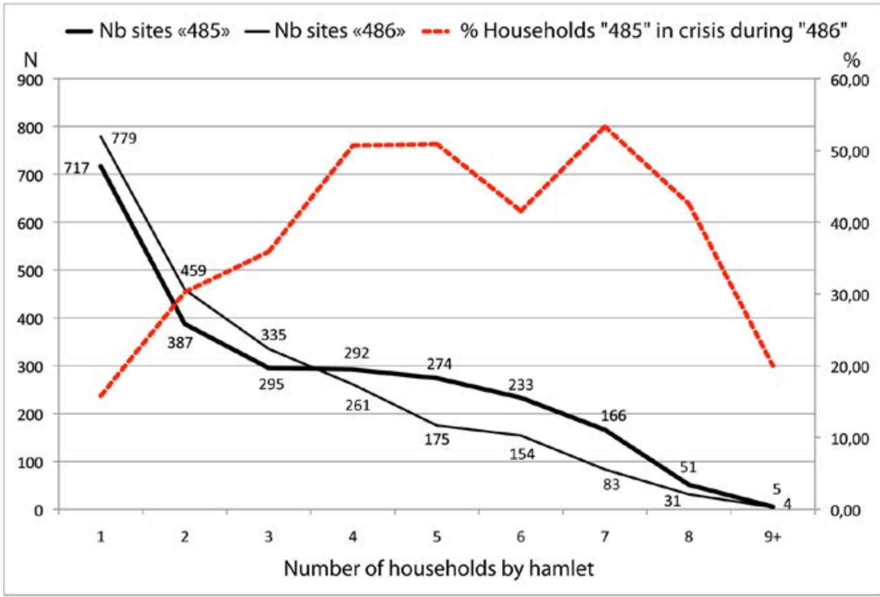


Figure 9: Effects of the '5065' crisis on hamlets household composition in central Germany; relative distribution before and after the crisis (solid lines) and differential effect on the hamlet types (dashed line).

BARTON, M. (2013). 'Stories of the past or science of the future? Archaeology and computational social science', in A. Bevan and M. W. Lake (Eds.), *Computational approaches to archaeological spaces.* Walnut Creek: Left Coast Press, pp. 151-178.

BENTLEY, A. (2007). *Mobility, specialization and community diversity in the Linearbandkeramic: isotopic evidence from the skeletons.* Proceedings of the British Academy 144, pp. 117-140.

BERGER, J.F., ORTU E., MOUSSA R. and TOONEN W. . (tbp). 'Precipitation, temperature and climate variation during the LBK period as estimated from the Rhine fluvial archives (the Shriek core)'. In: J.-P. Bocquet-Appel, R Moussa and J. Dubouloz (eds). *Model-based approaches to Europe's Neolithic First Farmers The Linearbandkeramik's trajectory.* NY: Springer, in press.

BINTLIFF, J. (1991). 'The contribution of an Annaliste/structural history approach to archaeology.' In: Bintliff J (ed). *The Annales School and Archaeology*. Leicester U Press: Leicester, pp. 1-33

BRAUDEL, F. 1958. 'Histoire et sciences sociales : la longue durée'. *Annales*, oct. 1958, pp. 727-731.

CARNEIRO, R.L. (1967). 'On the relationship between size of population and complexity of social organization.' *Southwestern Journal of Anthropology* 23, pp. 234-243.

CARNEIRO, R.L. (1978). 'Political expansion as an expression of the principle of competitive exclusion.' In Cohen R and Service ER (eds), *Origins of the state: The Anthropology of Political Evolution*. Philadelphia: ISHI publ., pp. 205-223.

CHAYANOV, A.V. (1923). *The Theory of peasant economy.* P Parey: Berlin. 1986: Madison: The University of Wisconsin Press.

CHILDE, G. (1929). *The Danube in prehistory*. Oxford: Clarendon Press.

CORBETT, J. (1988). 'Famine and Household coping strategies.' *World Development* 16(9), pp. 1099-1112.

CUTLER, P. 1984. 'Famine forcasting; Prices and peasant behavior in Northern Ethiopia.' *Disaster* 8(1), 48-56.

DIETLER, M. and HAYDEN, B. (2001). *Feasts*. Tuscaloosa: The University of Alabama Press.

EPSTEIN, J.M. (2006). *Generative Social Science.* Princeton University Press, Princeton.

GRONENBORN, D. (2007). 'Climate change and socio-political crises: some cases from Neolithic Central Europe.' In: T. Pollard and I. Banks (eds.). *War and Sacrifice: Studies in the Archaeology of Conflict.* Leiden/Boston: Brill, 13–32.

GRONENBORN, D. (2010). 'Climate, crises, and the ' Neolithisation ' of Central Europe between IRD-events 6 and 4.' In D. Gronenborn, J. Petrasch (eds), *The spread of the Neolithic to Central Europe International Symposium, Mainz 24-26 June 2005*. Mainz, Römisch-Germanisches Zentralmuseum: RGZM-Tagungen 4, pp. 61-80.

HARTOG DEN, AP. (1981). 'Adjustment of food behavior during famine. In Robson JK (ed): Famines. Its causes, effects and management.' *Food and nutrition in history and anthropology* vol 2. New York: Gordon and Breach Science publishers, pp. 155-162.

KOHLER, T.A. (2012). 'Complex systems and archaeology.' I Hodder (ed), *Archaeological Theory Today.* Cambridge UK: Polity Press, pp. 93-123.

JOCHIM, M.A. (2008). 'The Mesolitihc of the Upper Danube and Upper Rhine.' In G. Bailey and P. Spikins (eds), *Mesolithic Europe*. Cambridge U Press, pp. 203-220.

JOHNSON, G.A. (1982). 'Organizational Structure and Scalar Stress.' In C Renfrew, M Rowlands and BA Segraves-Whallon (eds), *Theory and Explanation in Archaeology. The Southampton conference.* New York: Academic Press, 389-431.

LACHIVER M. (1991). *Les années de misère*. Paris: Fayard.

LAKE, M.W. (2010). 'The uncertain future of simulating the past.' In A. Costopoulos, M W Lake (Eds.), *Simulating change: archaeology into the twenty-first century*. Salt Lake City: University of Utah Press, pp. 12-20.

LAKE, M.W. (2014). 'Trends in Archaeological Simulation.' *Journal of Archaeological Method and Theory* 21, pp. 258–287.

LE ROY LADURIE, E. (2004). *Histoire humaine et comparée du climat.* Paris: Librairie Arthème, Fayard.

LE ROY LADURIE, E., ROUSSEAU, D., VASAK, A. (2003). *Les fluctuations du climat de l'An Mil à aujourd'hui.* Paris: Fayard.

LEWELLEN T.C. (2003). Political anthropology: An introduction. Third editions. Wesport: Praeger publ.

MCGLADE, J. (2014). 'Simulation as Narrative: Contingency, Dialogics, and the Modeling Conundrum.' *Journal of Archaeological Methods and Theory* 21, pp. 288-305.

PAVÙK, J. (2004). 'Early Linear Pottery Culture in Slovakia and the Neolithisation of Central Europe.' In: A. Lukes, M. Zvelebil (eds), *LBK Dialogues: Studies in the formation of the Linear Pottery Culture.* BAR International Series, vol. 1304, Oxford: Archaeopress, pp. 71-82.

PFISTER, C. and BRÁDZIL, R. (2006). *Social vulnerability to climate in the 'Little Ice Age': an example from Central Europe in the early 1770s. Climate of the Past 2*, 2006, pp. 115–129.

POLLA, M. (2003). 'Peasant and hunter households in Oulanka, northern Russia, 1710-1910.' *The History of the Family* 8, pp. 163-181.

POLLA, M. (2006). 'Family systems in central Russia in the 1830s and 1890s.' *The History of the Family* 11, pp. 27-44.

PREMO, L.S. (2007). 'Exploratory agent-based models: Towards an experimental ethnoarchaeology'. In JT Clark and EM Hagemeister (Eds.), *Digital discovery: exploring new frontiers in human heritage. CAA 2006*. Computer applications and quantitative methods in archaeology. Budapest: Archeolingua Press, pp. 29–36.

PREMO, L.S. (2010). 'Equifinality and explanation: the role of agent-based modeling in postpositivist archaeology.' In A Costopoulos and M Lake (Eds.), *Simulating change: archaeology into the twenty- first century*. Salt Lake City: University of Utah Press, pp. 28–37.

RAPPAPORT, R.A. (1968). P*igs for the Ancestors: Ritual in the ecology of a New Guinea People.* New Haven, CT: Yale University Press.

RAILSBACK, S.F. and GRIMM, V. (2012). *Agent-based and individual-based modeling: a practical introduction*. Princeton: Princeton University Press.

ROZOY, J.G. (1978). *Les Derniers Chasseurs. L' Épipaléolithique en France et en Belgique. Essai de synthèse.* Chez l'auteur, 26 rue du Petit-Bois, F-08000 Charleville.

SCHWARTZ, D., ERTLEN D., DAVTIAN G., DUBOULOZ J., GAUVRY Y., VYSLOUZILOVA B., and BERGER J.F. (in press). 'The state of soil cover in central Europe at the end of the Early Holocene and estimation of LBK soil fertility.' In: J.P. Bocquet-Appel, R Moussa and J Dubouloz (eds). *Model-based approaches to Europe's Neolithic First Farmers The Linearbandkeramik's trajectory.* Springer: NY, in press

SHENNAN, S. (2009). 'Evolutionary demography and the population history of the European Neolithic.' *Human Biology* 81 (2), pp. 339-345.

SUTHERLAND, W.J. (1996). *From Individual Behavior to Population Ecology.* Oxford, U.K.: Oxford University Press.

SWEITZ, S.R. (2005). *On the periphery of the periphery: Household archaeology at hacienda Tabi, Yucatan, Mexico.* PhD. Texas A&M University, Tamu, USA.

THUCYDIDE. *Histoire de la guerre du Péloponnèse.* Traduction par J. Voilquin et J. Capelle. Librairie Garnier : Paris, vol 1 : 1936, vol 2 : 1948. http://remacle.org/bloodwolf/historiens/thucydide/table.htm

WIESSNER, P. and TUMU, A. (1998). *Historical vines.* Washington and London: Smithsonian institution press.

WOBST, H.M. (2010). 'Discussant's comments, Computer Simulation Symposium, Society for American Archaeology.' In A Costopoulos and M Lake (Eds.), S*imulating change: archaeology into the twenty- first century*. Salt Lake City: University of Utah Press, pp. 9-11.

Explaining the Adoption of 'Innovations' in Western Europe During Bronze Age: Towards an Agent-Based Simulation

Juan A. BARCELÓ, Florencia DEL CASTILLO BERNAL, Giacomo CAPUZZO, Berta MORELL, Joan NEGRE

Quantitative Archaeology Laboratory, Department of Prehistory, Universitat Autònoma de Barcelona, Spain
uanantonio.barcelo@uab.cat, florenciadelcastillo@hotmail.com, giacomocapuzzo@gmail.com, bertamorell@hotmail.com, negreperez@gmail.com

Abstract

In this contribution we present the design phase of an ABM implemented to understand how cultural identities and cultural standardization may have emerged in a prehistoric sedentary early complex society. The aim of the model is to explain how diversity and self-identification may have emerged in the small-scale societies of our prehistoric past. The model explores some possible consequences of the theoretical assumptions about cultural diversity and the process of ethnogenesis. We have not modelled the decision process from the point of view of individuals, but at the level of the population, modelling social dynamics as a set of factors that constrain social actions. The agent does not decide as a rational individual, but probabilities for action at each cycle are taken into account as soon as local conditions change. It is a way for analysing the consequences of actions without regarding the cognitive causes of action.

Keywords: Agent-Based Models, Late Bronze Age, Cultural Transmission, Ethnogenesis

Introduction

How Cultural Consensus Emerge ?

In past hunting-gathering societies, technology and knowledge about how to interfere with natural productivity were limited. In such conditions, humans were conditioned by their environment and, because of their extreme dependence to local carrying capacity and diminishing returns from labour; they were obliged to constant geographical mobility, given that they could not restore what they extracted from nature. Thus, the more people interacted participating in collective strategies, the better, because it contributed to increase labour efficiency. On the other hand, sedentary societies substituted territorial mobility by sedentariness, and agriculture. The relevance of technology (metal tools) was higher, and the nature of social interaction (both exchange and war) more complex. But also in those circumstances ethnicity appeared as a long-term cultural standardization that influenced the updating of social identities and the possibilities of economic cooperation.

As a consequence of different forms of social interaction and exchange between related individuals, cultural consensus may emerge, and human groups can aggregate into greater groups that exhibit culturally homogeneity, what affected social reproduction, increasing similarity in the long run, and reducing the risk of conflict with others groups identified as different (out of the new cultural consensus).

The research presented here is based on a previous model simulating Hunter and Gatherer identities dynamics (Del Castillo, 2012; Del Castillo *et al.*, 2014). One of the main characteristics of that model was the way we dealt with the formalization of the Cultural Identity concept. We adopted the classical Axelrod approach using a vector of X dimensions that represents an organized list of meanings, values, beliefs and symbols inherited at birth, learnt within the evolving group, modified all along the life of the agent and transmitted to the new generation (Axelrod, 1997). It should be taken into account that such identity is in constant renegotiation and updating. Consequently, it is recalculated at each time on the basis of the binary vectors containing all we know about material culture at that particular moment. Agents calculate at each time a Similarity Threshold on the basis of the percentage of consensus needed, depending on how much they need food or tools from others to survive. The more at risk they found themselves, the less tolerant to the others difference (Barceló *et al.,* 2009, 2010).

But Cultural similarity is not enough, especially when the economy is more complex than hunting and gathering, and it is based on delayed consumption (production of surplus, storing, capital accumulation). In the present version of our model, even in the case of a similarity above the similarity threshold, the agent with food or technology in excess should decide whether the proposed exchange has long term benefits: the more the actual cultural consensus, the less risks of being attacked later. We have modelled this decision in terms of a variation of the classical Prisoner's Dilemma. In case there are no agents with a cultural similarity above the actual value of similarity threshold to exchange food and/or technology, the agent may steal what culturally different agents have produced. The agent selects the nearest agent (in terms of cost-weighted distances) with the maximum cultural difference, and attacks depending on its own human and technological strength.

The Bronze Age Model we propose here designs an artificial sedentary society whose agents (territories and not individualized households) produce food with agriculture and herding practices, manufacture metal instruments for their work, exchange with other territories or steal what others may have produced. Each agent also has its own identity, inherited at birth, learnt within the evolving group, modified all along the life of the agent and transmitted to the new generation. It has been built as a hypothetical explanation of the process of adoption of a new funerary practice (cremation burials in urn) as an innovation adopted by different groups (Barceló *et al.,* 2014, Capuzzo *et al.*, 2014). From the end of the Middle Bronze Age onwards, human communities across Europe and on some places of the northern basin of Mediterranean began to adopt a new funerary practice characterized by the burial of cremated body parts, frequently inside an urn. We consider this change can be analysed in terms of the diffusion of a standard, according to Weitzel *et al.* (2006) sense of the term: *standard* is used to refer to any technology or product incorporating specifications that provide for *compatibility* or cultural consensus. Cultural consensus theory assumes that cultural beliefs are learned and shared across people. The challenge to this view is that it assumes social mechanisms through which members of a group can identify how much they share (Romney *et al.* 1986, 1996; Romney 1999). Consequently, instead of *assuming* that agents have common identity traits based on membership to an already existing 'ethnic' group, agents may ask themselves, as to what extent they 'believe' they are similar to those others in the same neighbourhood, and queried if the outcomes of those values are perceived to be similar. Cultural consensus should be considered as a relevant property of a social system that enables social agents to 'somehow go together' and makes them subject to a network effect. Hence, cultural compatibility standards enable agents (Barceló *et al.,* 2010, Del Castillo *et al.*, 2014). We use then 'cultural consensus' in the sense of active *standardization* as the implementation and use of a standard to interact with a communication partner. The theoretical bottom-line argument for standardization cultural processes is that the discrepancy between individual (at the level of the agent or the household) and collective (network wide) gains leads to coordination problems. The basic question underlying the agent cultural change decision is whether the costs of building a new cultural consensus through standardization exceed the benefits.

The problem is that the social (or even economic) benefits of building and maintaining a given cultural consensus are hardly quantifiable after the adoption. While the increased cultural similarity can lead to direct savings due to faster, more frequent and predictable communication, cultural consensus may also induce more strategic benefits: avoiding conflict and increasing the flow of goods and labour among culturally similar agents. Cultural similarity could be explained as the result of learning as a social mechanism of cultural transmission (Boyd and Richerson, 1985). This explanation understands culture as the information transmitted from one individual to another through teaching, imitation or other forms of social transmission. So, if we accept the notion of culture in these terms, the adoption of a particular cultural trait could be the consequence of multiple interactions between individuals or social groups over time and space.

2. Building the model

2.1 Entities, state variables and scales

Agents are modelled as a series of regions, each one reflects the statistical mode of all archaeological sites within a determined buffer zone around some topographic central

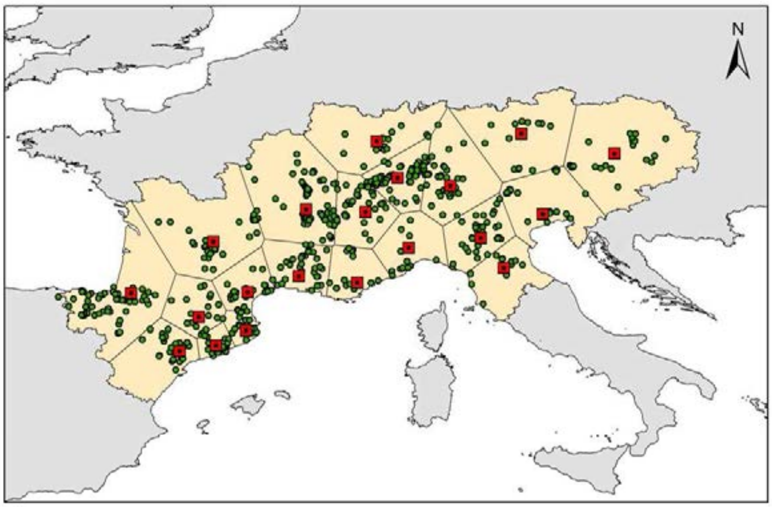


Figure 1: Map showing 20 agents (squares) corresponding to 20 different geographical regions identified with Thiessen polygons. Small dots represent 14C-dated archaeological sites included in the EUBAR database (Capuzzo et al. 2014) (Software: ESRI 2011. ArcGIS Desktop: Release 10. Redlands, CA: Environmental Systems Research Institute).

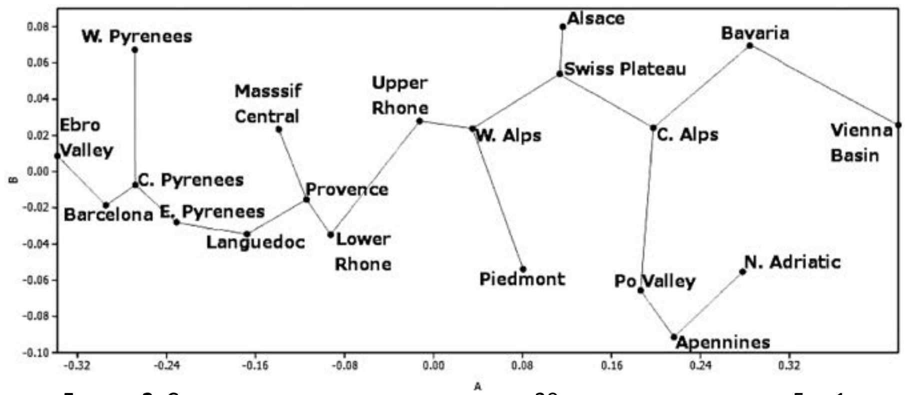


Figure 2: Cost weighted distances among the 20 agents represented in Fig. 1.

point. There are 20 agents in the simulation, and this number is constant all along the simulation (Figures 1 & 2).

We have created an anisotropic model of the environment in which the only that matters are the cost-weighted distances between regions. Distances based in cost-weighted models try to define the least costly path to reach each known point using the path with least accumulated travel cost. An inter-point cost-weighted distances matrix allows in this way to understand the spatial correlation between the elements of our point pattern distribution in a more trustworthy way. We have estimated such cost-weighted distances in terms of modern secondary roads in the current transportation network. From this, a symmetrical matrix of all pairwise cost-weighted distances has been created and processed by a Multi dimensional scaling (MDS) algorithm to perform the re-ordination of the point pattern into a new framework. That means that all agents are in a sense 'neighbours' within the global area considered for study. We have enough archaeological evidence that in Bronze Age sites situated 1000 km far away could be connected by some form of exchange or they had the risk of being attacked, but not all movements within such area were equally probable.

Agents attributes are:

The simulation uses the following external parameters, which should be initiated at start-up (table 1):

The simulation uses the following external parameters, which should be initiated at start-up (table 2):

2.2 Process overview and scheduling

All the agents are involved in four kinds of activities:

- Produce food
- Exchange food
- Exchange tools
- Steal food and tools

Surviving is the first process in the agent schedule at the beginning of a new time-step. Agents have a surplus from previous productive acts, and they PRODUCE FOOD. Production of food is also affected

LABOR UNITS (l_i)	A Poisson distributed parameter counting the aggregated quantity of labour from all groups.
SURVIVAL THRESHOLD ($\bar{e}_i$)	Survival of agents depends on the amount of food so, a survival threshold should be calculated in terms of the quantity of calories all agents need to be able to live a season long (six months).
IDENTITY	A vector.
TECHNOLOGY (b_i)	Is the aggregated efficiency of labour obtained when increasing the number of manufactured tools. It starts at 1 (lack of tools) and has not an upper maximum.
ENERGY (e_i)	Produced food, expressed in kilocalories.
SURPLUS (s_i)	The difference between energy produced and energy consumed and it is stored for later use.

Table 1: Agents Attributes.

INTERNAL CHANGE RATE (IRC)	A random value (from 0 to 1, usually very small) defined in analogy to the probabilities of internal change (invention, mutation, catastrophe, sudden change).
DEMOGRAPHIC VARIABILITY	λ of the Poisson distribution of the number of labour units within each agent.
LOCAL DIFFICULTY FOR PRODUCING FOOD (h_i)	A Poisson distributed parameter counting the quality of soil and the availability of water and temperature at each time-step: the poorer quality of soils and the scarcer is water. This parameter is initiated at start up (a random number following a Poisson distribution whose λ is a free parameter selected by the user at the beginning of the simulation), and changes every time-step, in such a way that at odd cycles (warm season) it is the half that at even cycles (cold season).
NUMBER OF NEW TOOLS CREATED AT THE END OF A GIVEN CYCLE	A user selected number of agents, from specific locations (mining regions) have the chance of producing a user selected constant number of new tools at the end of each cycle.

Table 2: External Parameters.

by *diminishing* marginal returns relative to the variations in the local difficulty of producing food and the effects of robbery. Food produced and not consumed at the present time-step is converted into surplus.

When an agent needs food or tools and some degree of cultural consensus already exists between agents in this area, the decision whether EXCHANGE or not is taken according a variation of the Prisoner's Dilemma. If an exchange is decided, the half of the actual value of surplus, or the half of the non-necessary tools are transmitted. The agent in need receives these quantities from all agents in the environment with a similarity in identity higher than a similarity threshold. The received quantities of food and tools are weighted negatively by the cost-distance separating them.

One agent STEALS another if 1) it is in need of food or tools, 2) they have appropriately dissimilar identities, that is to say, if some existing cultural consensus is below a critical threshold. Consequently, the current value of each agent identity vector influences the probabilities of cooperating or conflict within the current time-step. When exchange is successful, the current value of the identity vector changes adaptively to fit the newly built cultural consensus. That is to say, to decide if people in a region cooperated with people from another region without moving. If their respective identities are 'sufficiently' common, they decide to cooperate, and the probability of success in survival increases. If identities are too different, people do not cooperate there is a growing probability that they can enter in conflict stealing what they have produced and accumulated so far.

'Identity' is socially built by agents through a local imitation process, which evolve, change and adapt to local features at each time step. We define two main mechanisms for this identity process.

Internal change, supposed to be random at the scale of a population.

Adaptive, trying to fit individual identity to collective identity if economically advantageous.

2.3 Subprocesses

One time cycle in the simulation roughly represents what happened in a region during 1 season; two cycles or ticks represent one year. Nine processes are responsible for all system dynamics: agents work for surviving and they use positive (exchange) or negative (robbery) interaction flows to compensate for circumstantial interruptions in survival. Consequently, they need to *identify* other agents and act accordingly, thus, identity evolves and updates, as a result of interaction.

SURVIVE: In this simulation, survival threshold ($\bar{e}_i$) is fixed for all the simulation, although not any agent has the same survival threshold. It is a multiple of the number of labour units in the agent, assuming an individual needs an average of 730 kilocalories per year (2000 calories per day), and one time step (cycle or 'tick') in the simulation roughly represents what an agent is able to do in six months, $\bar{e}_i = (365 * l_i)$.

Each agent begins by using existing surplus, produced at the end of the previous time-step. If stored surplus (s_i) is equal or greater than the survival threshold ($\bar{e}_i$), the agent survives. If stored surplus is not enough, the agent should produce food (e_i). After survival, the amount of food that is not consumed is converted into surplus.

In case $e_{i(t)} = \bar{e}_i$, agents survive. In case $e_{i(t)} > \bar{e}_i$, then the difference $e_{i(t)} - \bar{e}_i$ is converted into surplus. In case $e_{i(t)} < \bar{e}_i$, the agent should look for alternatives sources of food (EXCHANGE or ROBBERY). If after looking for those alternatives the produced food still remains below the threshold ($e_{i(t)} < \bar{e}_i$), the agent loses one of its members (a labour unit), and recalculates the survival threshold ($365 l_i$) until it adapts the survival threshold to the number of labour unit that can be maintained. If after such reduction the threshold is still not attainable, the agent dies, indicating that the region is abandoned and it will not be reoccupied for the rest of the simulation.

Produced food is expressed in energetic units, that is to say in the same units as the survival threshold (kilocalories). Food is obtained by agent i by means of labour $l_i(t)$ with the contribution of its own technology $b_i(t)$, used to compensate the local difficulty of producing food h_i. In this simulation, we assume that any individual can produce 360 kilocalories in a single six-month time-step (2000 calories per day * 180 days).

$$e_i = [360\ l_i(t) / h_i\ ] * b_i(t)$$

(h_i) is a Poisson distributed parameter counting the quality of soil and the availability of water and temperature: the poorer quality of soils and the scarcer is water, the more labour or more technology is needed to obtain resources up to survival threshold). This parameter is initiated at start up (a random number following a Poisson distribution whose λ is a free parameter selected by the user at the beginning of the simulation), and changes every time-step, in such a way that at odd cycles (warm season) it is the half that at even cycles (cold season).

Technology has the role of multiply the effects of human labour. Its smallest possible value is 1, when the agent has no tools at all, and therefore it cannot multiply the results of human labour. Agents are always looking for increasing the number of tools from other agents in the network, assuming that there is no free access to sources of raw material (metal) nor the technical knowledge needed to produce tools. There are not a maximum of technology, given that all food produced in excess can be stored. There is no theoretical limit to the quantity of what can be stored.

EXCHANGE OF FOOD. When $e_{i(t)} < \bar{e}_i$ the agent first looks for obtaining food from neighbours. It searches in all the territory) all culturally similar agents and asks them for the food they have in excess (surplus). If the agent agrees,

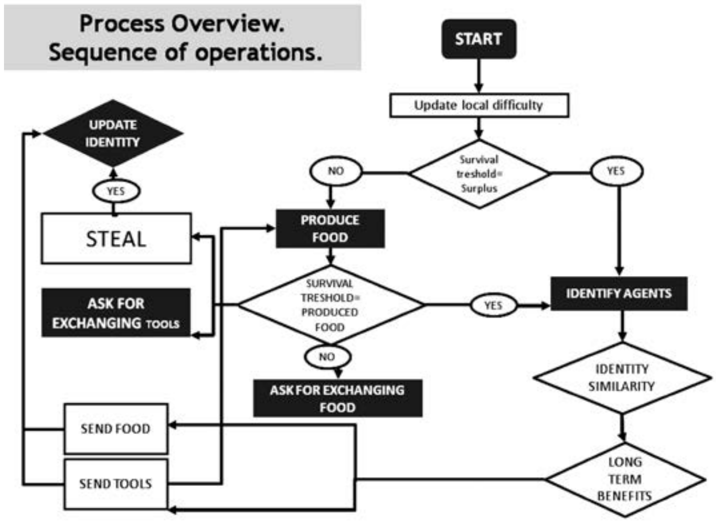


Figure 3: Process overview of the proposed ABM, sequence of operations.

a quantity of food is sent, multiplied by the inverse of the cost-surface weight linking the agent in need and the agent sending food.

Deciding whether sending surplus to a 'friend' in need or keeping it for its own future consumption, is a difficult decision. At first there is a preliminary condition: the agent asking for food should be culturally similar. We will come back to this point later. But the mere fact of having a high enough cultural consensus between both is not enough. It is important to take into account that sending food implies for the sender an important cost, it reduces its surplus and may affect to the probability of surviving at later steps. It can produce, however, a benefit in the long run, because it will increase cultural consensus and decrease the risk of being attacked later in the future. To be rational, the decision whether sending the requested food or refusing the proposed exchange should imply a way to evaluate the advantages or drawbacks of this behaviour. This can be implemented in terms of *utility threshold* to be maximized.

Our simulation implements a somewhat modified iterated Prisoner's Dilemma, agents seriously consider whether or not to maintain the opportunistic decision of getting the highest payoff in the short term or to opt for what seem the best for all agents in the territorial network in the long run; cooperation and high enough cultural consensus. This is because the players know that there is a distinct chance that they will meet again. The likelihood of future interactions thus casts a shadow on the present and arouses the possibility of an altruistic strategy.

Once taken the decision to exchange, the agent with an excess of food send a fixed amount of surplus (50% of actually accumulated surplus). This quantity arrives to the agent in need negatively weighted according to the cost-surface distance. The agent in need receives from all agents in the territory for which it has enough similarity, and it adds all quantities. The part of energy received in this way and exceeding the survival threshold is stored in its surplus and will be available for other agents in a reproduction of the same decision at the next cycle.

Exchange of Technology

When $e_{i(t)} < \bar{e}_i$ the agent first looks for obtaining food from neighbours. But it is also possible to increase technology to compensate local difficulty of producing food. It may have also occurred that the agent has lost its technology because of an attack. When the agent is in need of food, it looks simultaneously for food and technology from neighbours. As a result, it is perfectly possible that at the next cycle it has excess of technology and food for its own survival threshold. In such a case it become an agent in excess and inverse its position in hierarchy.

Technology is difficult to produce. You need a source of raw material, and information for that production. Consequently, the only way of obtaining it is through exchange and/or robbery. It is important to take into account that exchanging technology implies for the sender an important cost: it reduces its actual technology efficiency and may affect to the probability of having success in production of food at later steps, but it may also to increase the probability for an exchange in the opposite direction later. This is also a rational decision.

In the same way as in the case of food exchanging, the agent who should take the decision knows that if it accepts the proposed exchange act, although costly, because it loses some of its technology, is rewarding because the agent that asks for help will help when asked later. Utilities are a bit different in the case of technology, because the weight of rewarding in case of cooperation is assumed to be lower.

If people in region 1 asks for tools to increase their production of food and people in region 2 chooses to send some of its technology, then both agents receive a different reward: the agent in need receives what it needs (but with a global value lower than in the case of food), and the helping agent increases the expectations of cultural consensus with agent in need. The costs for sending technology are higher than in the case of food, therefore the reward of cooperation is significantly lower. If the agent with a surplus in technology refuse the exchange now, when it asks for need sometimes in the future, the region that actually needs to increase its technology will take revenge: cultural consensus will have decreased and the chances of being helped will be very low.

Once taken the decision to exchange, the agent with an excess of technology evaluates the amount of technology it can send. To do that, it calculates the minimum quantity of technology it needs to arrive to its survival threshold, taken into account the local difficulty of producing food

$$\min b_i = \bar{e}_i / [360\, l_i(t) / h_i]$$

The agent then sends the half of the difference between the actual value of b_i and min b_i(t):

TECHNOLOGY TO TRANSFER = $b_i(t) - \min b_i(t)$

This quantity is negatively weighted according to the cost-surface distance and arrives at the agent in need, which receives from all agents in the territory for which it has enough similarity.

Except for exchange, technology can only be acquired or decreased by robbery. In case an agent has been stolen and has lost all its technology, it should open exchange acts with culturally similar agents.

ROBBERY. In case there are no agents with enough cultural similarity to exchange food and/or technology, the agent may steal what culturally different agents have produced. The agent selects the nearest agent (in terms of cost-weighted distances) with the maximum cultural difference, and provided:

$$b_{agent\_in\_need}(t) * l_{agent\_in\_need}(t)) * \textit{weighted distance} >$$

$$b_{most\_different\_agent}(t) * l_{most\_different\_agent}(t)$$

The attacking agent will take all the surplus and technology of the attacked agent.

IDENTIFY-NEIGHBORS. Many procedures depend on the actual degree of cultural consensus between agents. It is calculated in terms of the similarity between the identity of agents. The identity of an agent is expressed in a binary vector including the presence/absence information of cultural elements, which can be also identified in the archaeological record. Given that each agent is in fact an integration of local groups (archaeological sites: settlements, hillforts, workshops, cemeteries, ritual places, etc.) such a vector contains the statistical mode of all attributes present in the buffer zone corresponding to each 'regional' agent. Because we do not have any other information, all attributes have exactly the same relevance in the calculation of similarity.

How big should be this threshold to allow cooperation between agents or increase the probabilities of being attacked? If an agent is able to produce enough food by itself and survive, it does not bother to identify other agents in the territory. Otherwise, agents calculate the percentage of consensus needed, depending on how much they need food or tools from others to survive. The greater the ease with which an agent obtains needed resources, the more predisposed to help at no cost. This is because the more cooperation today, the more expectations to cooperate in a more or less near future. As a result of collaborative work, current identities tend to approximate, increasing expectations of future collaboration. In the same way, the more survival depends on exchange, the more tolerant to others difference. We assume there is an expectation of the benefits of cooperation: the more benefits coming from exchange, and the more disadvantages coming from conflict and robbery, the higher the tolerance to the other difference in identity, and the lower the percentage of consensus needed. A similarity threshold (*ST*) expressing the degree of independence an agent has on others is defined in terms of the amount of extra-resources needed for survival, given the local amount of resources.

Where h_i represents the local difficulty of obtaining resources, and e_i food produced, as in equation (1). The higher this value, the less tolerant to the other difference and the more difficult is to create a link of positive interaction. If the current measure of identity similarity between agents *i* and *j* is above the calculated value of $ST_{(t)}$ then there is option to decide whether exchange or not, according to expected benefits (Prisoner's Dilemma). In case the current measure of identity similarity between agents *i* and *j* is below the calculated value of $ST_{(t)}$ in case the agent *i* is in need of food or technology, it will attack agent *j*. This parameter is the exactly opposite to Schelling's tolerance (1978): the higher the number of needed common features to build cultural consensus s, the lesser the tolerance with 'the other' difference. The easier to build cultural likelihood, that is, the less common traits are needed, the higher the tolerance.

UPDATING IDENTITY. Cultural consensus emerges by combining the identities and values of interacting agents in an emergent group. Therefore, once the agent exchanges for increasing the chance of its own survival, the identity vector used to define the possibilities of cooperation is updated towards the statistical mode of the exchanging agents identity. With a fixed probability level (95%) each agent copies the statistical mode of identities within the group. There is an additional source of identity change, implemented at start-up and fixed for all the simulation, in form of an Internal Change Rate (IRC). This is a random value (from 0 to 1, usually very small) defined in analogy to the probabilities of internal change (invention, mutation, catastrophe, sudden change). Then on every tick, and with a fixed probability level determined as an external parameter, the identity vector mutates.

Although the simulation does not implement reproduction, assuming the stability of population, we consider that what future generations inherit from their forefathers is not a fixed identity, but an evolving communality. We assume that the higher the cooperation between people, and the higher the cultural consensus among them, the higher the probability that reproductive couples be formed within groups. The idea is that once a social network has emerged and survival of agents has been assured, hybridization mechanisms began to act because inherited identities (ethnicity) should be modified to maintain the newly built consensus.

3. Experimenting with the model

This Bronze Age Model was designed to explore possible hypothesis about the effects of some social and economic factors on the emergence of cultural diversity. We are interested in studying the possible impact and relevance of:

- Local difficulties and food production
- Environmental adaptation and Climatic Change as causal factors for cultural standardization during Late Bronze Age
- Conflict and banditry as a consequence of increased difficulties in the production of food
- Alliance and reciprocity to compensate for the increased difficulty in producing food

The simulation is also designed to study the effects of implementing non-Euclidean distances in our model. We assume that space varies in a heterogeneous way, so the movement on its surface has a directionally dependent behaviour (Negre, 2014). Therefore, we intend to explore different cost-weighted distance matrices to explore how socially mediated connectivity affects the emerging pattern of cultural standardization. At Start-up, cultural vectors can be very different among all agents (random initialization), or all agents can share the same culture; in so doing we can analyse the degree of regionalization during the previous cultural phase as a factor to explain the degree of cultural standardization during Late Bronze Age.

Furthermore, the effects of territorial ranking and power can be experimented by varying some start-up parameters and considering alternative worlds. For instance, when at start-up we select the population (labour units) of agents, we can create different scenarios:

One agent has a higher population than others, and dominates the production of food and the challenges of attacking or being attacked

All agents have equal ranking and equal population

Only one agent has the possibility to produce tools because it is the only one controlling the source of raw material. The rest of agents must obtain tools through exchange with culturally similar agents, or steal the tools accumulated by culturally different agents

Reduced number of agents (higher than 1) monopolize raw material and technology reproduction

All agents have equal access to the possibilities of tool production

Conclusions

As any other society, European Late Bronze Age territories were constituted by local groups of people connected to one another by overlapping arrays of social ties that together configured a social network (Capuzzo 2014). Social interaction, and hence, the flow of people, goods and ideas, depended upon each territorial network of intercommunity contact or its network of social communication and that the configuration of this network was primarily dependent on the presence of various social barriers which may have impeded, diverted and channelled communications.

Our key theoretical assessment is to consider the emergence of ethnicity as a long-term process of group formation where cultural standardization influences the updating of social identities and the possibilities of economic cooperation. We also consider this kind of standardization as a consequence of different forms of social interaction where cultural consensus may emerge. In this scenario aggregation of human groups also emerges affecting social reproduction by increasing both similarity in the long run and intergroup affinity.

We view the emerging complex of 'Urnfield culture' as a process of reproducing identity from generation to generation. As a result, some people arrived to share some knowledge and some behaviours because they have learnt from the same people. However, what is learnt at birth and during childhood is progressively modified during life when interacting with other people with different knowledge, behaviours and believes. Social reproduction does not take place in an empty social world, but it should be built in the present through the social and political selection of prospective partners. That is, ethnogenesis and identity formation emerge as result of the contradiction between social inertia (knowledge inheritance) and cultural consensus built during cooperation and labour exchange.

The key of our perspective is that any shared traits among social groups, their behaviour, their beliefs, and their language, the products of their work and/or the material or immaterial results of their actions should be contingent to the social interaction process that generated those traits. In so saying, we follow a constructive approach to 'ethnicity' and the study of cultural diversity. That means that the way Bronze Age people took economic, social and political decisions is what configured people clustering at different scales. In other words, the question is '*why* groups of people are the way they are' in terms of *how* they acted within a social aggregate their previous activity contributed to build. The complex interplay of social actions, people and the consequences of their actions explain the degree of cultural consensus and standardization by showing how social aggregation fit into a causal structure, that is to say, a vast network of interacting *actions* and *entities*, where a change in a property of an entity dialectically produces a change in a property of another entity. Cultural standardization at the end of Late Bronze Age was probably the result in a change in the way social agents interacted in their economic and political activities. Our simulation intends to explore alternative scenarios to understand the social mechanism that generated those changes.

References

AXELROD, R. (1997). The Dissemination of Culture: a Model with Local Convergence and Global Polarization. *Journal of Conflict Resolution* 41, pp. 203-226.

BARCELÓ, J.A., DEL CASTILLO, M. F., MAMELI, L., MORENO, E.J. (2009). The Computer Simulation of Social Dynamics and Historical Evolution. The case of 'Prehistoric' Patagonia. *Proceeding of the 2nd Workshop on Social Simulation and Artificial Societies Analysis (SSASA'08)*. CEUR Workshop Proceedings, Barcelona, Spain. Available online http://CEUR-WS.org/Vol-44.

BARCELÓ, J.A., CUESTA, J.A, DEL CASTILLO, F., GALÁN, J.M., MAMELI, L., MIGUEL, F., SANTOS, J. I. AND VILÀ, X. (2010). Patagonian Ethnogenesis: towards a computational simulation approach. *Proceedings of the 3rd World Congress on Social Simulation*. Kassel: Center for Environmental Systems Research, University of Kassel. pp.1-9.

BARCELÓ, J.A. AND DEL CASTILLO, F. (2012). Why hunter and gatherers did not die more often? Simulating prehistoric decision making. In: Graeme Earl, Tim Sly, Angeliki Chrysanthi, Patricia Murrieta-Flores, Constantinos Papadopoulos, Iza Romanowska and David Wheatley (eds). *Archaeology in the Digital Era*. Amsterdam: Amsterdam University Press.

BARCELÓ, J.A., CAPUZZO, G. AND BOGDANOVIĆ, I. (2014). Modeling Expansive Phenomena in Early Complex Societies: the Transition from Bronze to Iron Age in Prehistoric Europe. *Journal of Archaeological Method and Theory*.21 (2), pp. 486-510.

BARCELÓ, J.A., DEL CASTILLO, F., DEL OLMO, R., MAMELI, L., MIGUEL QUESADA, F.J., POZA, D. AND VILÀ, X. (2014). Social Interaction in Hunter-Gatherer Societies: Simulating the Consequences of Cooperation and Social Aggregation. *Social Science Computer Review* 32 (3), pp. 417-436.

BOYD, R. AND RICHERSON, J. (1985). Culture and the evolutionary process. Chicago: The University of Chicago Press.

DEL CASTILLO, F (2012) *Modeling Ethnic Heterogeneity and Cultural Diversity of Patagonian Hunter-Gatherers. Approaches from Computer Simulation and Agent Based Models.* PhD. Thesis. Spain, Department of Prehistory. Autonomous University of Barcelona.

DEL CASTILLO, F., BARCELÓ, J.A., MAMELI, L., MIGUEL, F. AND VILA, X. (2014). Modeling Mechanisms of Cultural Diversity and Ethnicity in Hunter-Gatherers. *Journal of Archaeological Method and Theory* 21 (2), pp.364-384.

CAPUZZO, G. (2014). *Space-Temporal Analysis of Radiocarbon Evidence and Associated Archaeological Record: from Danube to Ebro Rivers and from Bronze to Iron Ages*. PhD Thesis. Spain, Department of Prehistory. Autonomous University of Barcelona.

CAPUZZO, G., BOARETTO, E. AND BARCELÓ, J.A. (2014). EUBAR: A Database of ^{14}C Measurements for the European Bronze Age. A Bayesian Analysis of ^{14}C-dated Archaeological Contexts from Northern Italy and Southern France. *Radiocarbon* 56 (2), pp.851-869.

NEGRE, J. (2014). Non-Euclidean Distances in Point Pattern Analysis: Anisotropic Measures for the Study of Settlement Structures in Heterogeneous Regions. IN: J.A. Barceló and I. Bogdanovic eds. *Mathematics and Archaeology* (forthcoming). Enfield: Science Publishers (CRC Press).

ROMNEY, A. K. (1999). Cultural consensus as a statistical model. *Current Anthropology* 40, pp. 103-115.

ROMNEY, A., KIMBALL, S., WELLER, J., BATCHELDER, W. H. (1986). Culture as consensus: A theory of culture and informant accuracy. *American Anthropologist* 88. pp. 313–38.

ROMNEY, A. KIMBALL, S., BOYD, P., MOORE, C.C., BATCHELDER, W.H. AND BRAZIL, T. (1996). Culture as shared cognitive representations *Proceedings of the National Academy of Sciences, U.S.A.,* pp. 4699–4705.

WEITZEL, T., BEIMORN, D. AND KÖNIG, W. (2006). A Unified Economic Model of Standard Diffusion: The Impact of Standardization Cost, Network Effects and Network Topology. *Management Information Systems Quarterly* 30 (8), pp. 489-514.

Spatial Interaction Simulation Methods for Ancient Settlement Distributions in Central Italy

Taylor OSHAN,[1] Carson FARMER,[2] Eoin O'DONOGHUE[3]

[1]Centre for GeoInformatics, University of St Andrews, UK
[2]City University of New York, Hunter College, USA
[3]National University of Ireland, Galway, Ireland
tayoshan@gmail.com, carson.farmer@gmail.com, eoin.odonoghue@nuigalway.ie

Abstract

The research presented here aims to review existing spatial interaction theories pertaining to ancient Mediterranean civilization and then to explore quantitative methods for analysing relative interaction amongst early Etruscan states within a geographic network structure. Three conceptually unique techniques will be employed in order to simulate networks of relative interactions amongst a distribution of sites in the region of Etruria: an agent-based model, a radiation analogy model, and a Hamiltonian gravitational model. We strive to demonstrate how each can be leveraged for different research questions pertaining to the development of a system of settlements. By testing a variety of models we more clearly identify unique features of each, which make it more amenable for particular applications.

Keywords: Spatial Interaction Modelling, Etruria, Networks

Introduction

Human Landscapes are created by the necessary struggle of people to meet their social and economic needs (Eckbo, 1969). Thus, as societies have settled and developed urban communities, human culture intensified, which in turn magnified the relationship between the people and the space they inhabit. The result is that the landscape is created by human decisions whether it was their intent to do so or not (Eckbo, 1969; Anschuetz, Wilhusen & Scheick, 2001, p.161). These decisions result in connections between people, places, and groups, producing a network of spatial interactions. Through spatial analysis it is possible to simulate representations of these spatial interaction networks. Consequently, they can be used to enrich our perception of a region and therefore help answer broader questions about how a place becomes defined through the larger landscape within which it exists (Tuan, 1976).

The study area will first be introduced, including historical developments and theories pertaining to spatial interaction. Next, we select the models which are best suited for simulating spatial interaction. We then test each model by collecting network visualizations and metrics of their results, given identical inputs. Finally, we present an overview of how each model can best be applied to develop more specific research questions. The limitations of the models are highlighted with the goal of better understanding how to navigate the selection of an appropriate spatial interaction model, despite the fact that they often produce results which can be interpreted in a similar fashion.

1. Ancient Italy and the Etruscans

This research will focus on settlements associated with the Etruscans, an important Mediterranean civilization that dominated much of central Italy from the seventh to fifth centuries BCE (Figure 1), but are arguably visible in the archaeological record from the thirteenth century.

There is evidence for the control of the production, distribution and consumption of natural resources as early as the sixth millennium BCE (Malone, 2003). This is most apparent through finds of decorated Neolithic pottery and the practice of animal husbandry (Robb & Kostalena, 2012; Tagliacozzo, 2005). The practice of controlling resources served as the basis for the growing economic, political, and social competition that was to be the catalyst for the development of organized communities. The limited evidence from the Italian Chalcolithic (roughly the mid-fourth to second millennium BCE) suggests an increase in agricultural productivity, and new distinct funerary customs appeared, including divisions of graveyards according to social hierarchies (Cardarelli, 1992).

By the Final Bronze Age there was significant intensification that lead to a shift from a landscape of hamlets and farms to one that included proto-urban villages. Throughout this period, population growth is evident with the appearance of new settlements. Expanded territorial control gave rise to an increase in larger residences, hierarchical rankings, and economic activity. While this period still had a significantly distributed population of inhabitants, it marks the beginning of the process of nucleation and the formation of chiefdoms (Barker and Rasmussen 1998, 44-

49). Characterized by a dramatic increase in the process of nucleation, the Iron Age (950-750 BCE),[1] shows signs of further societal development. Settlement numbers decrease considerably though those that prevail increase in population, size, and influence. It is in this period that the region under consideration begins to resemble a collection of central places, which lend themselves agreeably to the basic assumptions involved in spatial interaction modelling.

The Etruscans became a distinct cultural group at least as early as the Final Bronze Age (c.1300-1150 BCE) through to the Roman conquest of Italy in the final centuries BCE (Riva, 2010). Indeed, early Roman, Latin, and Etruscan material culture is often indistinguishable, and it is possible to argue for a shared cultural koine throughout much of Central Italy (Bradley, 2000; Cornell, 1995, p. 163-165; Roth-Murray, 2007). From the seventh and sixth centuries BCE it is possible to recognise early city-states similar in concept, and perhaps even socio-political organisation, to those in existence elsewhere in the Archaic Mediterranean world (Torelli, 2000). These city-states incorporated sophisticated linguistic, political, military, and urbanising systems into dynamic socio-political entities. The traditional geographic extent of Etruria at its greatest extent included territories as far south as modern-day Campania to the Po Valley in north-east of the Italian peninsula (Spivey & Stoddart, 1992, p. 21-38). However, the heartland of the Etruscans roughly equates to the contemporary region of Tuscany and northern Lazio.

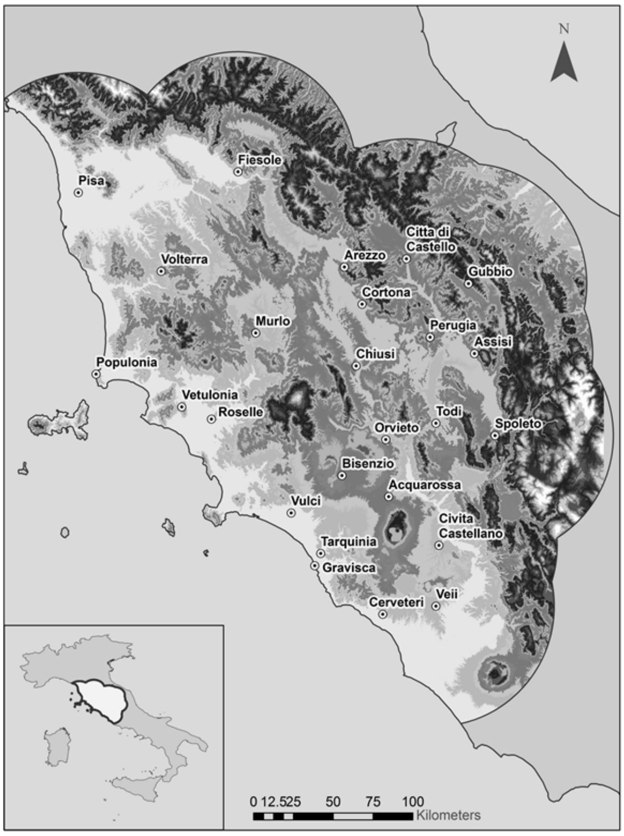


FIGURE 1: Study sites and the regional topography.

2. Spatial Interaction Theory Within Etruria

Interaction within regions seems to enjoy popularity as a theory of drivers of growth within civilizations (Kowalewski, 2008, p. 226; Binford, 1983, p. 380). Colin Renfrew (1975) provides one of the most complete theories of interaction, in the form of trade, as a driver for the development of early states. While Renfrew credits trade as the mechanism behind the spatial interaction, he also discusses how these interactions rely on the interdependence of material and spiritual aspects of human culture and inevitably leads to communication and information flows as well. He suggests that habitual exchange leads to central places, since it requires organization and security and it implies some assumed criteria of value. Competing centres often result in specialization so that demand for the trade of goods between central places increases, thus promoting greater interaction.

According to Renfrew, trade between domains within a region is the least studied despite the idea that its 'effect must have been to produce and maintain the uniformity of culture or civilization as a whole' (1975, p. 18). He also states that 'Civilization implies the development of a highly structured and differentiated society, with specialist production (craftsmen), a permanent controlling organization disposing of a significant proportion of produce (government), and a developed, explicit set of shared beliefs (cognitive structure), sometimes with large aggregations of population' (35). This suggests that those sites which have remains indicative of increasingly complex societies would have likely participated in more inter-site trade amongst the other central places. Similarly, Izzet cites increased encounters with 'others' as a driving factor in attitude changes towards ethnic identity. In this process, the differentiation of the self is an effect which takes place only after discovering 'others' (2007, p. 210). Physical growth and cultural development are therefore inextricably linked and, to a certain degree, controlled through the common process of interacting through space.

3. Spatial Interaction Theory and Modelling

Spatial interaction encompasses an array of behaviours, such as communication or movement, which occur over geographic space as a result of a decision process (Fotheringham & O'Kelly, 1989). In each application a trade-off is considered between the costs incurred from interaction and the benefits accrued through doing so. Early spatial interaction models were developed using an analogy to gravitational forces from the field of physics (Roy & Thill, 2004). In the 1970's a new framework for gravity-based interaction modelling, based on maximizing entropy, was introduced. An aggregate flow between two locations is considered a macro state which can be satisfied

[1] Chronological estimates for Italy, before 600 BCE, is a topic fraught with many challenges and is constantly changing with the gradual emergence of new data from radiocarbon testing, dendrochronological analyses, and other methods. Therefore, the dates provided here are subject to change. They largely follow the traditional chronological range provided by Riva (2010, p. 2-3).

by various micro states, though the most likely macro state is one which can be described by the highest number of micro states (Fotheringham, Brundson & Charlton, 2002). Despite the analytical gains of the maximum entropy models, criticisms arose due to the fact that it was still based on a physical analogy and that many of the applications lacked a behavioural interpretation (O'Kelly, 2004).

Consequently, more contemporary efforts have sought to derive models based on spatial information processing. Fotheringham (1983) developed the competing destinations model to accommodate the fact that during the decision-making process an individual will be faced with a unique spatial distribution of alternative locations for every destination under consideration. The attributes of each alternative can contribute to how and whether or not an individual considers it at all. As a result, it is entirely possible for decisions to be made that may not achieve the maximum possible benefit for an individual since they may not even evaluate every choice within a system. More recently, Simini et al. (2012) present the radiation model as an entirely new framework which deviates from the gravity law concept (which has been at the core of most previous models). A notable feature is that the model is parameter free and therefore requires no previous flow data for calibration. For each possible origin and destination pair, flows are predicted by comparing the benefits of interacting with one proposed destination to the benefits of interacting with each other site within a radius from the origin equal to the distance from the origin to the proposed destination. Thus the model depends not directly on the distance variable but rather the aggregate benefits (e.g. population of a city) available at all other alternative choices between the origin and destination, lending itself to the analogy of radiation and absorption processes. Given the simplicity of the radiation model and the fact that, at present, there is no known research in which this model has been employed in the field of archaeology, it will therefore provide an interesting comparison against a gravity law based model.

4. Models in Classical Archaeology

Renfrew and Level's XTENT model (1979) employs a rigorous theoretical framework in order to mathematically model the effects of space within an archaeological region by dividing territory between respective central places, similar to weighted Thiessen polygons.

Ultimately, their model describes hierarchy through competition, by assigning political boundaries primarily based on settlement sizes, rather than evaluating how space mediates interaction within settlement networks. For this reason, the model will not be used in this research, which seeks to directly compare relative spatial interaction; however, Stoddard and Redhouse's (2011) extension of the XTENT model to Etruria provides insight into territorial dynamics of the region under inspection.

A simple model which has been employed in archaeology (Terrell, 1977; Broodbank, 2000) for measuring connectivity is proximal point analysis (PPA), in which each site is connected to its K nearest neighbours so that distance is not considered in the decision making process. In contrast, a fixed radius model considers two sites connected if the distance between them is less than model parameter D (K and D are set by the analyst). In a sense, both of these models are unrealistic because they only account for a limited set of scenarios (Rivers, Knappett & Evans, 2013). Additionally, these simple models do not designate flow directions for the links; either two sites are simply connected or they are not, which excludes the option for non-reciprocal relationships, which could be expected if a settlement was economically or politically subservient to a more powerful one. These models were therefore not applied in this case study.

Gravity models have historically been deployed within classical archaeology (Clarke, 1977). One maximum entropy implementation by Rihll and Wilson (1991) offered an improvement on other simpler ones by adding constraints and parameters, though it was not without flaws. The model has been criticized that it inherently creates networks that are 'supply side' such that larger cities attract the interactions of the nearest smaller sites, never allowing small sites to connect with each other. In effect, you get a network reminiscent of a star pattern (Knappett, Evans, Rivers, 2008) which is helpful for identifying key nodes but does not represent a realistic network.

Ariadne, a model developed by Evans et al. (2009), improves upon past gravity-based model implementations. They employ a 'Hamiltonian' energy function, H, consisting of four terms, which takes the form,

$$H = -kR - \lambda E + jP + \mu T \qquad (1)$$

where R is the benefit of exploiting local resources, E is the benefit of maintaining links, P is the cost of supporting local populations and T is the cost of maintaining links. Each term has a parameter which controls the effect it has within the model. Expanding the terms give H the following form,

$$H = -k\sum_i S_i v_i(1-v_i) - \lambda\sum_{ij} V(\frac{d_{ij}}{D}) * (S_i V_i) * (S_j e_j) + j\sum_i S_i v_i + \mu\sum_i S_i v_i e_{ij} \qquad (2)$$

where S_i is the given size of a site, v_i and e_{ij} are outputs of the model representative of the weight of each settlement and interaction link in the network, and $V(d_{ij}/D)$ provides a gravity function based on the distance between two nodes (d_{ij}) and an average maximum journey length (D). H provides a tool for balancing the cost and benefits of both 'supply' and 'demand' between site interactions. Using optimization techniques, a minimum value is obtained for H and the model output values (v_i and e_{ij}) can be read off and visualized as a network (Knappett, Evans, Rivers, 2008, p. 10-24)

The additional output of interaction 'strength' (e_{ij}) makes it possible to create a network in which all nodes can participate, yet closer sites will have strong links and farther sites will have weaker ones. Unlike Rhill and Wilson's model, smaller neighbours are free to interact with each other rather than being limited to the hierarchical 'star' (Evans, Rivers & Knappett, 2012, p. 12), and was therefore selected as the gravity law based model to simulate network interactions within central Italy.

Model three, inspired by the Travellersim model, introduces a third unique framework into this research while also building off of the Rhill and Wilson framework.[2] Authors Graham and Steiner (2008) propose a method for observing the emergence of territories from the perspective of the individual. The general motivation for this agent-based approach is that it is not truly possible to know the exact relationship between sub-systems within a larger system. Instead we can assign very simple rules to individuals and then let them interact, thus developing the more complex relationships observed at the macro level. Within each iteration of a model run, the agents consider 3 settlements within their local (20km) vision and then choose to travel to the most 'attractive' destination. Attractiveness is calculated using the equation:

$$Attractiveness = \frac{importance^{\alpha} * e^{-\beta D}}{visitors^{\alpha} * e^{-\beta D}} \qquad (3)$$

where the importance is the current importance of a site, α is a parameter to represent the benefits from resources, β is a parameter to represent the cost of communicating over space, and D is the distance from the agent to the settlement whose attractiveness is being calculated. After all agents have moved, settlements then update their importance based upon the interaction they received using the following calculation:

$$Importance = V * \sqrt{\frac{\sum I}{v}} \qquad (4)$$

where V is the cumulative total number of visitors who have arrived at a settlement, I is the importance values from each of the settlements an arriving agent had originated from for a given iteration, and v is the number of visitors at each settlement for the current round.

In this manner the authors were able to successfully 'grow' fuzzy territories and visualize the spread of influences over the region. Furthermore, by tracking the individual visits of the agents it is possible to produce a weighted directed network similar to the radiation and Ariadne model. In this respect, Travellersim is not limited like Rhill and Wilson's model or the XTENT model despite their shared interest in hierarchy and territory. The bottom-up approach demonstrated in Travellersim provides an alternative to the more top-down perspectives of the Ariadne and radiation models. It was therefore used as the basis to construct the third model for this research.

[2] http://figshare.com/articles/Travellersim_Agent_Based_Model/91976

Name	Number	Size (ha)	GettyLon	GettyLat
Veii	0	185	42.0333	12.4
Cerveteri	1	160	42	12.1
Vetulonia	2	100	42.85	10.9667
Populonia	3	150	42.9833	10.4833
Volterra	4	100	43.4	10.85
Chiusi	5	50	43.0167	11.95
Bisenzio	6	35	42.57	11.87
Acquarossa	7	30	42.4833	12.1333
Perugia	8	32	43.1333	12.3667
Gravisca	9	24	42.2	11.7167
Gubbio	10	20	43.35	12.5833
Assisi	11	20	43.0667	12.6167
Citta di Castello	12	20	43.45	12.2333
Tarquinia	13	150	42.25	11.75
Vulci	14	126	42.4167	11.5833
Roselle	15	40	42.8	11.1333
Murlo	16	10	43.15	11.3833
Pisa	17	20	43.7167	10.3833
Orvieto	18	85	42.7167	12.1167
Arezzo	19	32	43.4167	11.8833
Cortona	20	30	43.2667	11.9833
Civita Castellano	21	26	42.2833	12.4167
Fiesole	22	30	43.8	11.2833
Todi	23	20	42.7833	12.4
Spoleto	24	20	42.7333	12.7333

Table 1: Sites used in the study and their attributes.

5. Methodology

5.1 Data

For all of the chosen models, the necessary data is a selection of settlements, their geographic coordinates, a measure of their size, and the distance between each pair. Stoddart and Redhouse (2011) provide a list of 25 Etruscan settlements, which co-existed within the time period 900 BCE to 600 BCE, along with estimates of their urban extent. Their data provides the best available starting point for spatial interaction modelling in this context because they have assembled site size estimates by drawing from various archaeological sources as well as personal experience when no estimation existed (p. 166). The sites, as presented in Table 1, were all given spatial context via the Latitude and Longitude attributes from the Getty Thesaurus of Geographic Names Online.[3] If an entry existed for the ancient site as opposed to the modern one, then the ancient site was used, but if only one entry existed (modern or ancient), then that was utilized. For Bisenzio, which was not listed in the thesaurus, coordinates were estimated visually using its known location via Google Earth.

To incorporate the natural terrain into the distance calculations, methods similar to those utilized by Stoddart and Redhouse (2011, p. 165) within the XTENT model will be leveraged to promote consistency. Ultimately though, the method employed here must differ since this research seeks to incorporate the terrain by deriving least accumulated cost paths between sites whereas they created cost distance rasters for the region.

Distances between sites were derived by first calculating a slope raster from an ASTER[4] 30m resolution digital elevation model (DEM), which was then clipped to a buffer of two times the average nearest neighbour distance (straight

[3] http://www.getty.edu/research/tools/vocabularies/tgn/
[4] http://asterweb.jpl.nasa.gov/

line) of the settlements and to the natural coast line of the Italian peninsula (Stoddart & Redhouse, 2011, p. 167). The slope raster was then reclassified by differing energy 'costs' (Minetti et al., 2002) to traverse each cell using generalized slope ranges based on the direct relationship between increases in energy costs and slope (Table 2) to produce a cost raster. This was then used as input to derive the least accumulated cost paths in origin-destination matrix format (Etherington, 2011). These calculated values are all longer than the straight-line distances since they tend to closely gravitate towards areas with lower slopes. They represent the routes (Figure 2) which will require the least energy to be exerted without greatly deviating from the shortest possible path.

5.2 Simulating Networks Using Spatial Interaction Models

5.2.1 Radiation model

Code for the radiation model[5] was developed following Simini et al. (2012) with one exception. The term T_i, which is supposed to be a proportion of the total population originating at each site, instead considers the settlement size estimate at a given location. Since M_i in the original equation is also equal to the origin estimate, we apply a simplified version of the radiation model:

$$T_{ij} = \frac{(M_i^2 * N_j)}{(M_i + S_{ij})(M_i + N_j + S_{ij})} \qquad (5)$$

Where T_{ij} are the predicted flows, M_i is the origin size estimate, Nj is the destination size estimate and N_{ij} is the sum of location sizes that occur within a radius from M_i equal in length to the distance between M_i and N_j.The code formats the data into an origin-destination matrix and then carries out equation (5) for each pair [*i, j*] of locations, resulting in a matrix of relative spatial interaction between them. Since many flows between sites will have near-zero values the network is filtered by removing all edges with a weight that is less than one percent of the largest edge weight value.

[5] https://github.com/tayoshan/RadiationModel.git

Name	Number	Size (ha)	GettyLon	GettyLat
Veii	0	185	42.0333	12.4
Cerveteri	1	160	42	12.1
Vetulonia	2	100	42.85	10.9667
Populonia	3	150	42.9833	10.4833
Volterra	4	100	43.4	10.85
Chiusi	5	50	43.0167	11.95
Bisenzio	6	35	42.57	11.87
Acquarossa	7	30	42.4833	12.1333
Perugia	8	32	43.1333	12.3667
Gravisca	9	24	42.2	11.7167
Gubbio	10	20	43.35	12.5833
Assisi	11	20	43.0667	12.6167
Citta di Castello	12	20	43.45	12.2333
Tarquinia	13	150	42.25	11.75
Vulci	14	126	42.4167	11.5833
Roselle	15	40	42.8	11.1333
Murlo	16	10	43.15	11.3833
Pisa	17	20	43.7167	10.3833
Orvieto	18	85	42.7167	12.1167
Arezzo	19	32	43.4167	11.8833
Cortona	20	30	43.2667	11.9833
Civita Castellano	21	26	42.2833	12.4167
Fiesole	22	30	43.8	11.2833
Todi	23	20	42.7833	12.4
Spoleto	24	20	42.7333	12.7333

TABLE 2: Cost classification for each slope range

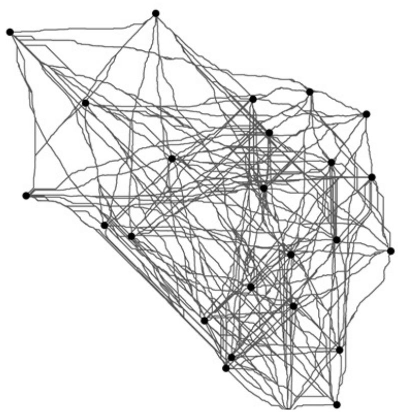


FIGURE 2: Accumulated least-cost paths between all sites. Each dot represents a site, while the lines are the routes which have the lowest path between sites, considering the regional topography.

5.2.2 Ariadne Model

Based on conclusions aggregated by Bevan (forthcoming) the maximum travel speed over Mediterranean terrain in the ancient world for a pedestrian would have been 5km/h, assuming they were not carrying any significant load. In contrast, horses would have reached speeds of 4km/h - 8km/h when carrying packs and 10km/h - 30km/h when utilized solely for riding. If we take a full day's travel to be about 12 hours, then the maximum distance travelled by a pedestrian in a day is approximately 60km, while pack horses could have neared 100km and riding horses could have journeyed well beyond that. It is apparent that these two different travel modes would impose varying spatial limitations and therefore the model will be run with *D* values of both 60km and 100km in order to compare their impact on the results. A distance of 5km was defined as the point at which very close settlements are considered as one entity though the site distribution for this study is unaffected since all settlements are separated by values larger than this threshold.

Ariadne's four parameters *k*, λ, *j*, and μ must be adjusted in order to find a network which best represents a theoretical scenario. Drawing on the experience of previous research (Knappett, Rivers, & Evans, 2011) we will initiate parameterization using values of k = 1.0, λ = 4.0, j = -2.0, and μ = 0.1. For our purpose, we searched for a network which represents two different scales (short vs. long) which are differentiated by their strength (strong vs. weak). For the system assuming walking distance restrictions, the initial parameters will be used to produce networks with ideal conditions. It is evident that employing the same initial values using a *D* value of 100km will be too dense and permit too many strong long distance relationships. Doubling *k* or halving λ gives more accurate representations and therefore, by increasing *k* from 1 to 1.5 and decreasing λ from 4 to 3 it was possible to take a balanced approach to settling on parameters to produce a preferable interaction system.

Due to the stochastic nature of Ariadne it is necessary to take an average of many model runs (25 in this research).[6] This ensures that fluctuations which may occur between different model runs are accounted for within this analysis. All networks produced from this model were filtered to remove near-zero edges using a cut-off value produced by dividing the average edge weight by seven times its standard deviation. This method was chosen, rather than the one employed for the radiation model, due to the fact that the distribution of network edges for different Ariadne model runs was always similar but that the maximum edge weight value could vary enough so that filtering using a cut-off based upon it would create very different networks.

5.2.3 Agent-based Model

To take an agent-based approach to simulating spatial interaction for this study region, some changes[7] were imposed upon the Travellersim model introduced earlier. First, the local vision was increased from 25km to match the journey limitations proposed for use in the Ariadne model (60km and 100km). This change reflects the additions of horse related transportation and a general increase in spatial awareness that would have been assumed to accompany intensified relationships within the landscape in this time period. For this research, only the longer distance constraint could be utilized due to the fact that the short distance constraint resulted in networks which *a priori* excluded some settlements from all others. In contrast, all other model variations resulted in one connected component. Next, the attractiveness of all settlements were calculated each round rather than randomly selecting three from those within a human agent's vision. The order in which agents moved was then selected at random to maintain the stochastic nature of the model and an average over many networks was used for interpretations.

Parameterization of the agent model includes choosing the logistical variables of how many human agents should start at each settlement agent and how many iterations the model should be run for, as well as model variables such as the difficulty of communication and benefits derived from resources. Testing the simulations with varying numbers of agents or iterations holding other parameters constant demonstrated that the model was insensitive to these logistical inputs. Five agents and 10 iterations (per agent) were chosen to ensure enough complexity[8] in each model run without utilizing unnecessary computing time. The model variables were tested in a similar manner; values for the range .01 to 5 were tested for each parameter holding the other one constant. At the given study region scale, the system structure proved extremely resilient to these deviations, so values of 1 were employed for both variables. Networks produced from the agent-based model do not result in near-zero values for edge weights and therefore do not need to be filtered.

5.3 Network Analysis Metrics

Metrics for analysing aspects of networks created by each model in this study are reviewed below.[9]

5.3.1 Average Node Degree

The degree of a node is the measure of how many other nodes it maintains relationships with, with a higher degree being associated with higher connectivity. By taking the average of all nodes it is possible to approximate overall connectivity within a network.

5.3.2 Diameter

The diameter is a measure of the shortest path (number of links traversed) between the most distant nodes in a network. It is a helpful measure because more connected networks will have a structure which results in a smaller diameter.

5.3.3 Transitivity/Average Clustering

A clustering coefficient, which measures transitivity, is most easily described in terms of social networks, in which it is the mean probability that a friend of a friend is also your friend. In network terminology, this relationship is referred to as a triangle and the coefficient measures the ratio of all triangles present in a network compared to the total possible triangles (Newman, 2003).The coefficient will provide insight into the clustered nature of the generated interactions.

5.3.4 Density

A network's density is defined as the ratio between the number of edges present and the maximum number of possible edges the network could have with a given set of nodes. A completely disconnected graph has a density of zero while a density of one corresponds to a network in which every node is connected to every other node.

5.3.5 Modularity

This measures the structure of a network by defining smaller communities or modules within the larger system. Higher modularity occurs when there are dense connections between nodes within communities but sparse connections among those between communities. Communities are suggestive of settlements which would have been more strongly associated in terms of economic, political or social relationships. Specifically, Louvain's community detection algorithm (Blondel et al., 2008) will be used to calculate modularity and define communities.

[6] The Ariadne software could not be automated, and therefore time constraints restricted more models runs from being used. The same number of model runs was averaged for the agent-based model to maintain consistency across stochastic models.

[7] https://github.com/tayoshan/AgentModel.git

[8] Visually, it was possible to detect a feedback mechanism within 10 iterations, such that all agents exclusively traveled between a few of the larger settlements.

[9] For further description please refer to Newman (2010).

6. Results and Discussion

The radiation model produces networks with strongly spatially clustered communities (Figure 3). This is also evident due to the fact that it has the highest modularity for all of the models (Table 3). Since the model does not require any parameters, this model is ideal for comparing results across regions. For example, we could compare networks from Etruria against contemporary sites in Ancient Greece to investigate differences in community structure. In contrast, its lack of parameters excludes it from being useful for examining how components of a larger system would have affected the area of interest.

Both Ariadne models (60km and 100km variations) show higher metrics values (Tables 4 & 5) than the radiation model and higher diameter and transitivity values than the agent-based model, given the prescribed parameters. Since those parameters can easily be changed we cannot conclude that the model systematically produces more connected networks; however, we can see that despite being more highly connected (via higher metric values) than the radiation model, that it has a lower modularity value. Furthermore, we can see that between the two Ariadne model variations, increasing the average trip distance from 60km to 100km results in higher metric values except for modularity. Allowing settlements to interact further afield seems to affect the overall structure of their interactions. This is illustrated particularly well by the settlements in the north-east of the study region (Figures 4 & 5). Initially, they appear to interact mostly amongst themselves but when the spatial restriction is expanded, they cease to interact with themselves in favour of their larger but longer-distance neighbours. Within the context of this research these diverging interaction landscapes are analogous to populations with and without access to horses. It could be interpreted as evidence towards decaying community structure as societies became wealthier and access to horses increased. The Ariadne model lends itself nicely to testing how various spatial scales differently condition interactions but its parameterization process can be tenuous and this model can only be recommended when there is strong evidence to support a particular scenario. It is well equipped to ask questions such as ‘how does the system change when it suddenly becomes more costly to support the population’ or ‘how does the system change with the introduction of new transportation technology’. In contrast, it would be challenging to compare two regions, each with their own set of parameters, using this model and is consequently better suited to conduct analysis over a single study area.

It is interesting that the agent-based model results in networks (Figure 6) which visually look as modular as the radiation model networks yet its modularity value is much lower (Table 6), closer to that of the Ariadne models. The strength of the communities may be weaker than they appear since the settlements have a similar strength of interactions with those outside of their community as with those within their community. This could be due to the

FIGURE 3: COMMUNITIES AND INTERACTIONS FROM RADIATION MODEL NETWORK. EACH SITE IS REPRESENTED AT A CIRCLE (AND NUMBERED ACCORDING TO TABLE 1), WHERE THE SIZE DENOTES ITS ESTIMATES INPUT SIZE AND THE COLOUR SIGNIFIES IT COMMUNITY AFFILIATION BASED ON ITS INTERACTIONS. THE INTERACTIONS ARE REPRESENTED BY THE EDGES INCREASE IN SIZE IN PROPORTION TO THE MAGNITUDE OF INTERACTION ON THE EDGES. THE ARROWS DENOTE THE DIRECTION OF THE INTERACTION.

FIGURE 4: COMMUNITIES AND INTERACTIONS FROM ARIADNE MODEL (60KM) NETWORK. EACH SITE IS REPRESENTED AT A CIRCLE (AND NUMBERED ACCORDING TO TABLE 1), WHERE THE SIZE DENOTES ITS ESTIMATES INPUT SIZE AND THE COLOUR SIGNIFIES IT COMMUNITY AFFILIATION BASED ON ITS INTERACTIONS. THE INTERACTIONS ARE REPRESENTED BY THE EDGES INCREASE IN SIZE IN PROPORTION TO THE MAGNITUDE OF INTERACTION ON THE EDGES. THE ARROWS DENOTE THE DIRECTION OF THE INTERACTION.

Metric	Diameter	Average Degree	Transitivity	Density	Modularity
Radiation Model	4	10.08	0.4108	0.21	0.4555

TABLE 3: METRICS FOR RADIATION MODEL NETWORK.

Metric	Diameter	Average Degree	Transitivity	Density	Modularity
Ariadne (60km)	3.12	12.8576	0.6982	0.2679	0.3143

TABLE 4: METRICS FOR ARIADNE MODEL (60KM) NETWORK.

FIGURE 5: Communities and interactions from Ariadne model (100km) Network. Each site is represented at a circle (and numbered according to Table 1), where the size denotes its estimates input size and the colour signifies it community affiliation based on its interactions. The interactions are represented by the edges increase in size in proportion to the magnitude of interaction on the edges. The arrows denote the direction of the interaction.

FIGURE 6: Communities and interactions from agent-based model network. Each site is represented at a circle (and numbered according to Table 1), where the size denotes its estimates input size and the colour signifies it community affiliation based on its interactions. The interactions are represented by the edges increase in size in proportion to the magnitude of interaction on the edges. The arrows denote the direction of the interaction.

Metric	Diameter	Average Degree	Transitivity	Density	Modularity
Ariadne (100km)	2.92	13.5872	0.7134	0.2831	0.2904

TABLE 5: Metrics for Ariadne model (100km) Networks.

fact that every settlement started with the same number of traveller agents and therefore many interactions are the result of the demand that all agents move every iteration of a model run. For future work, it may be appropriate to scale the number of starting agents by the size of each site. Doing so would be more representative of the contemporary landscape and would likely lower the overall connectivity and increase modularity, yielding metrics more similar to the radiation model. It should also be noted that in the agent-based model results, some nodes located between two communities could randomly shift their affiliation amongst different runs. These small fluctuations, due to the stochastic nature of the model, could have been predisposed to conflict given their positioning amongst their neighbours, something which would not become apparent in the radiation model due to its deterministic nature.

Conclusion

It is widely known that no mathematical representation of a social phenomenon can be perfect, but that they can be helpful and offer insight given a defined research question. We have reviewed the historical context and theoretical underpinnings to justify spatial interaction modelling and highlighted an array of methods for doing so. It can be seen that building models can aid many different types of regional studies. Furthermore, the minute intricacies among models may only become apparent when compared against each other. Often, models seem interchangeable since they can all produce a network of relative spatial interactions. We emphasize the need to carefully select a model based on its ability to facilitate the investigation of specific hypotheses. In future work we plan first to design experiments to test detailed questions within Etruria and then to subsequently further develop the appropriate models for carrying out the analysis. It will be equally important to collect additional data which can be used to derive interaction relationships between settlements to verify the simulated data. It would be interesting to identify both the regions which adhere to the assumptions within the models as well as those which deviate from them, signalling more complex phenomena and perhaps gaps within current spatial interaction theories.

References

ANSCHUETZ, K. F., WILHUSEN, R. H. and SCHEICK, C. L. (2001). An archaeology of landscapes: Perspectives and directions. *Journal of Archaeological Research*. 9 (2). p.157-211.

BARKER, G. and RASMUSSEN, T. (1998) The Etruscans. Oxford: Blackwell.

BEVAN, A. (forthcoming) Travel and interaction in the Greek and Roman world. A review of some computational modelling approaches, In: Dunn, S. and Mahoney, S. (eds.). *Digital Classicist.* London: Wiley-Blackwell (Supplement Bulletin of the Institute of Classical Studies).

BINFORD, L. R. (1983). *Working at Archaeology.* New York: Academic Press.

BLONDEL, V. D., GUILLAUME, J.-L., LAMBIOTTE, R. and LEFEBVRE, E. (2008). Fast unfolding of communities in large networks. *Journal of Statistical Mechanics: Theory and Experiment* 10. P10008 DOI:10.1088/1742-5468/2008/10/P10008

BRADLEY, G. (2000). States, cities, and tribes in central Italy. In: Herring, E. and Lomas, K. (eds). *The Emergence of State Identities in Italy*. London: Accordia. pp. 109-129.

BROODBANK, C. (2000). *An Island Archaeology of the Early Cyclades*. Cambridge: CUP.

CARDARELLI, A. (1992). Le Età dei metalli nell'Italia settentrionale. In: Guidi, A. and Pipemo, M. (eds). *Italia Preistorica*. Roma: Laterza. pp. 366-419.

C-LARKE, D. L. (1977). *Spatial Archaeology*. New York: Academic Press.

CORNELL, T. J. (1995) *The Beginnings of Rome: Italy and Rome from the Bronze Age to the Punic Wars (c. 1000-264 BC).* London and New York: Routledge.

ECKBO, G. (1969) *The Landscape We See*. New York: McGraw Hill.

ETHERINGTON, T. R. (2011) Python based gis tools for landscape genetics: Visualizing genetic relatedness and measuring landscape connectivity. *Methods in Ecology and Evolution*. 2 (1). p. 52-55.

EVANS, T. S., RIVERS, R. J. and KNAPPETT, C. (2012) Interactions in space for archaeological models. *Advances in Computer Systems*. 15 (01n02).

FOTHERINGHAM, A. S., BRUNDSON, C. AND CHARLTON, M. (2002). *Quantitative Geography*. London: Sage Publications Ltd.

FOTHERINGHAM, A. S. AND O'KELLY, M. (1989). *Spatial Interaction Models: Formulations and Applications*. London: Kluwer Academic Publishers.

FOTHERINGHAM, A. S. (1983). A new set of spatial interaction models: The theory of competing destinations. *Environmental and Planning A*. 15 (1). p. 15-36.

GRAHAM, S. and STEINER, J. (2008). Travellersim: Growing settlement structures and territories with agent-based modelling. In: Clark, J. T. and Hagemeister, E. M. (eds) (2008) *Digital Discovery: Exploring New Frontiers in Human Heritage. Papers from the 34thAnnual Conference of Computer Applications and Quantitative Methods in Archaeology (CAA), Fargo, United States, April 2006,* Budapest: Archaeolingua.

IZZET, V. (2007). *The Archaeology of Etruscan Society*. New York: Cambridge University Press.

KNAPPETT, C., EVANS, T. AND RIVERS, R. (2008). Modelling maritime interaction in the Aegean bronze age. *Antiquity* 82. pp. 1009-1024.

KNAPPETT, C., RIVERS, R. and EVANS, T. (2011). The Theran eruption and Minoan palatial collapse: New interpretations gained from modelling the maritime network. *Antiquity.* 85 (329). p. 1008-1023.

KOWALEWSKI, S. A. (2008) Regional settlement patterns. *Journal of Archaeological Research*. 16. p. 225-285.

MALONE, C. A. T. (2003) The Italian Neolithic: a synthesis of research. Journal of World Prehistory. 17(3). p. 235-312.

MINETTI, A. E., MOIA, C., ROI, S., SUSTA, D. and FERRETTI, G. (2002) Energy cost of walking and running at extreme uphill and downhill slopes. Journal of Applied Physiology. 93 (3). p. 1039-1046

NEWMAN, M. E. J. (2003) The structure and function of complex networks. *SIAM Review* 45. p. 167-256.

NEWMAN, M. E. J. (2010) *Networks: An Introduction*. Oxford: Oxford University Press.

O'KELLY, M. E. (2004) Isard's contributions to spatial interaction modeling. *Journal of Geographical Systems* 6 (1). p. 43-54.

RENFREW, A. C. and LEVEL, E. V. (1979) Exploring dominance: Predicting polities from centres. In: Renfrew, C. and Cooke, K. (eds) *Transformation: Mathematical approaches to Culture Change*. New York: Academic Press.

RENFREW, C. (1975) Trade as an action at a distance: Questions of interaction and communication. In: Savbloff, J. A. and Lamberg-Karlovsky, C. C. (eds) *Ancient Civilizations and Trade*. Albuquerque: School of American Research – Univeristy of New Mexico. pp. 3-59.

RIHLL, T. and WILSON, A. (1991) Modelling settlement structures in ancient Greece: New approaches to the polis. In: Rich, J. and Wallace-Hadrill, A. (eds) *City and Country in the Ancient World*. London: Routledge. pp. ,59-95

RIVA, C. (2010). *The Urbanisation of Etruria: Funerary Practices and Social Change, 700-600 BC.* Cambridge: Cambridge University Press.

RIVERS, R., KNAPPETT, C. AND EVANS, T. (2013). Network models and archaeological spaces. In: Bevan, A. and Lake. M. (eds) *Computational Approaches to Archaeological Spaces*. Left Coast Press Inc.

ROBB, J. and KOSTALENA, M. (2012). In small things remembered: Pottery decoration in Neolithic Southern Italy. In: Jessen, M., Johannsen N. and Jensen, H. J. (eds). *Excavating the Mind: Cross-Sections Through Culture, Cognition and Materiality.* University of Aarhus Press. pp. 161-181.

ROTH-MURRAY, C. (2007). Elite interaction in Archaic Etruria: Exploring the exchange networks of terracotta figured frieze plaques. *Journal of Mediterranean Studies* 17(1). pp.135-160.

ROY, J. A. and THILL, J.-C. (2004) Spatial interaction modeling. *Papers in Regional Science*. 83. p. 339-362.

SIMINI, F., GONZALEZ, M. C., MARITAN, A. and BARABASI, A. L. (2012). A universal model for mobility and migration patterns. *Nature* 484. pp. 96-100.

SPIVEY, N. J. and STODDART, S. (1990). *Etruscan Italy: An Archaeological History.* London: Batsford.

STODDART, S. and REDHOUSE, D. I. (2011). Mapping Etruscan state formation. In: Terenato, N. and Higgins, D. C. (eds) *State Formation in Italy and Greece: Questioning the Neoevolutionist Paradigm.* Oxford: Oxbow Books.

TAGLIACOZZO, A. (2005). Animal exploitation in the early Neolithic in Central-Southern Italy. *Antropologia-arkeologia* 57. p. 429-439.

TERRELL, J. E. (1977). *Human biogeography in the Solomon Islands. Fieldiana Anthropology* 68(1), pp. 1-47.

TORELLI, M. (2000). The Etruscan city-state. In: Mogens, H.H. (ed) *A Comparative Study of Thirty City-State Cultures: An investigation conducted by the Copenhagen Polis Centre*. Copenhagen: Royal Danish Academy of Sciences and Letters. pp. 189-208.

TUAN, Y.-F. (1976). Humanistic geography. *Annals of the Association of American Geographers* 66 (2). pp. 266-276.

Interactions and Network Analysis of a Rock Art Site in Morro do Chapéu, Bahia, Brazil

Carlos ETCHEVARNE,[1] Grégoire VAN HAVRE[2]

[1]Laboratory of Archaeology, Federal University of Bahia, Bahia Arqueológica Research Group, Brazil
[2]Post-Graduation in Archaeology, Federal University of Pernambuco. Bahia Arqueológica Research Group, Brazil
etchevarnebahia@gmail.com, gvanhavre@gmail.com

Abstract

Figurative rock art in northeast Brazil offers a large collection of human representations. In the region of Morro do Chapéu, in the State of Bahia, many scenes present interacting figures with patterned shapes. In this paper, we present a general method based on networks for the analysis of specific traits in relation to defined variables. Here, morphological attributes are used to investigate the behaviour of those different shapes at the Toca do Pepino, a sandstone rock shelter located on the Ventura River valley. We expose how we built a 2-mode network and analysed centrality measures and partitions. This method allows us to propose an interpretation for this figurative rock art in northeast Brazil as a system of graphic patterns shared by different groups.

Keywords: Interactions; Network Analysis; Rock Art; Brazil.

Introduction

Recent rock-art studies show that single lines of evidence, based on either style or chemical properties, cannot account for the great variability of motives in every site. The same classificatory problems that appeared when stylistic types were confronted to 14C dates, in the early 1950s, could then be repeated with a general scheme based on physical-chemical composition of pigments. The solution lies then, not only in the development of new techniques of analysis, but also in our ability to intertwine their results.

In this article, we propose to use network analysis as a mean to draw and calculate relations between various layers of evidence. We present a study of human representation from the Toca do Pepino, in Morro do Chapéu (Brazil), where we consider the presence of different patterns interacting.

The idea of analysing relations between archaeological objects starts at the premise that there were real relations between them. Depending on the interest of the researcher, these relations might range from direct individual contact to more general consideration, like the influence of texture on shape. Each problem requires a specific set of variables, and every result needs an interpretation.

The depiction of social interactions in the rock art from the Chapada Diamantina region in Brazil is best illustrated in one particular scene, identified at the Toca da Figura (Figure 1). This sandstone rock shelter is located on the top of the southern bank of the Ventura River, approximately 30 kilometres to the east of the city of Morro do Chapéu (Etchevarne, 2007; 2009). The scene is a simple line of 6 individuals bearing spear-throwers, painted in yellow and divided in two groups of three by a protrusion in the rock. These individuals have different head shapes, which can also be found, separately, in other scenes from this site and others in the region.

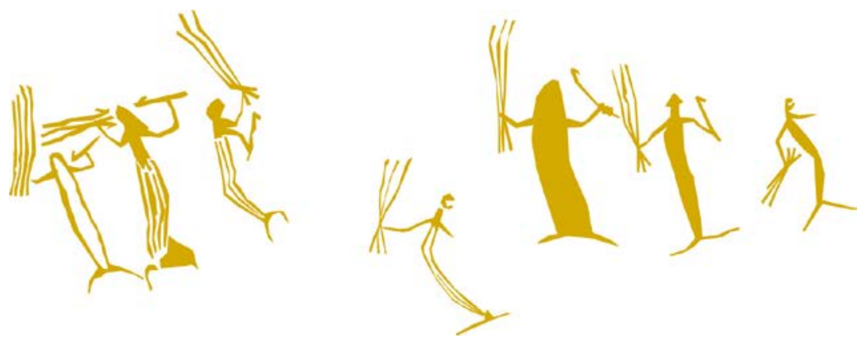


Figure 1: Interaction scene between six individuals painted in yellow ochre at the Toca da Figura, Morro do Chapéu, Bahia (Brazil).

Generally, this kind of representation – lines of individuals – always appear in a single format, and all the individuals are painted the same way. Changes can eventually be found in their respective positions, or in the level of details. For an example, individuals at one end of the line can be more or less detailed, allowing us to propose an orientation, even when they are represented facing the observer. Our main question, in this case, was then to determine if the difference was random or purely graphic, or if we could identify any cultural value.

1. Method

Our attempts at classifying this material with a taxonomic method faced various problems. First, anthropomorphic figures have been defined on the basis on a few essential traits: presence of a trunk, a head, and superior and inferior members (Guidon, 1984; Pessis, 1987). While these criteria allowed us to consider the figures depicted in the interaction, it also gathers others with which they could hardly be compared. Second, many of these motives have unique characteristics, and should then be classified in separate types. As a result, some types would have only one single figure. Considering such a high variability, we needed a finer definition and an instrument capable of analysing each individual figures through a single set of variables.

We proposed a new classification system for the analysis of a neighbouring rock shelter, called Toca do Pepino. This site is located on the top of the northern side of the Ventura River, two kilometres from the first one. It's a rather large sandstone site with more than 50 meters of extension and a thousand different paintings, mainly in red and yellow. As it is naturally well protected from rain and direct sun light, the paintings are in better conditions.

We started isolating a more specific set of motives by adding one more essential trait to those already defined for anthropomorphic figures: the representation of direct manipulation of an object. This allowed us to work with a smaller but more precise database, and showed two direct advantages: first, we can more safely compare the individuals, as we are sure we're talking about human representations (but see Ottoni & Izar, 2008). Second, we also selected a specific way to represent reality on rock surfaces, by detailed figuration. At the Toca do Pepino, we identified a total of 129 human figures. Compared to sites from other regions in Northeast Brazil, where few dozens anthropomorphic figures were found, it is a significant quantity (Kesterling, 2007; Costa, 2012).

The second step was to define a layer of analysis. For practical reasons, we couldn't consider the possibility of studying the chemical properties of the pigments used for each one of these figures. While this remains one of our future objectives, we had to concentrate on different, more accessible variables. In this respect, morphological analysis allowed us to describe each figure based on common graphical elements. Shape still is an important factor in the identification of indigenous groups in Brazil, as many use specific objects or body ornamentation such as paintings and haircuts in several occasions (Ribeiro, 1989; Wobst, 1977; Wiessner, 1983; Schortman, 1989).

We classified all these figures with five binary variables, resulting in ten different attributes: presence or absence of the neck, of an angle at the elbows, of an angle at the knees, and the shape of the trunk (volume or line). As a result, there were 16 possible combinations (Figure 3). Every figure has a specific composition of four attributes, that can be shared by others, and every site has different proportions. If we can't resolve the case of figures with unique characteristics,

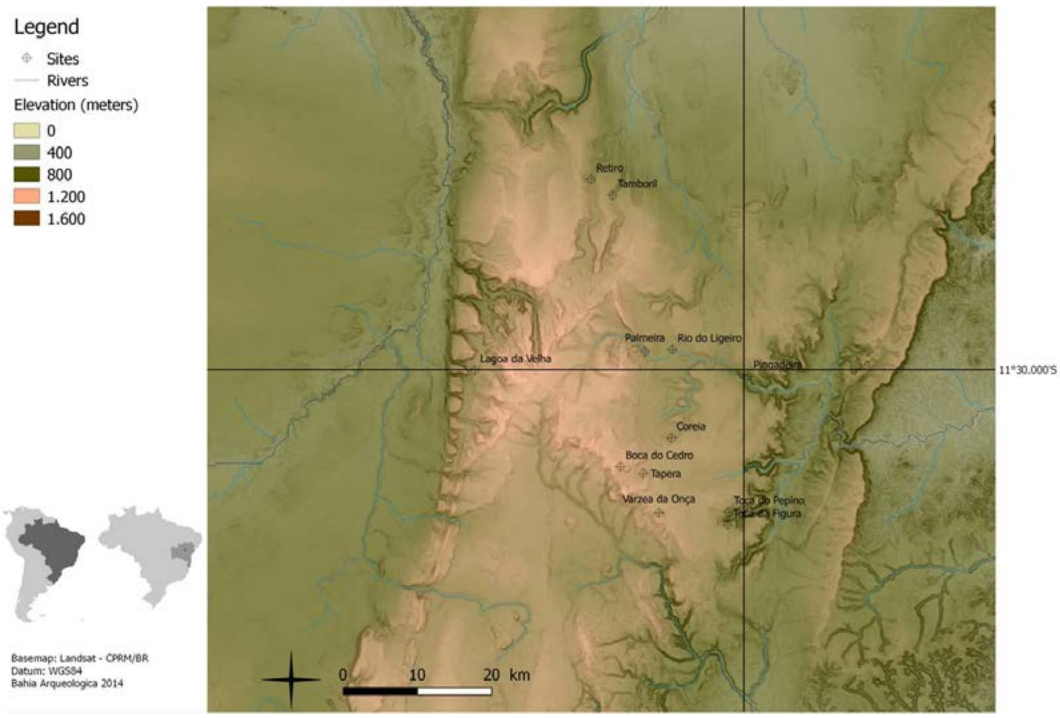
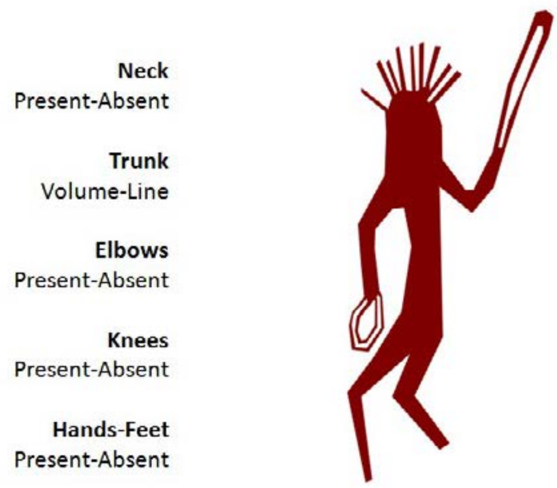


Figure 2: Map of the different rock art sites in the region of Morro do Chapéu.



Type 1	1010101010
Type 2	1010101001
Type 3	1010100110
Type 4	1010100101
Type 5	1010011010
Type 6	1010011001
Type 7	1010010110
Type 8	1010010101
Type 9	1001101010
Type 10	1001101001
Type 11	1001100110
Type 12	1001100101
Type 13	1001011010
Type 14	1001011001
Type 15	1001010110
Type 16	1001010101
Type 17	0110101010
Type 18	0110101001
Type 19	0110100110
Type 20	0110100101
Type 21	0110011010
Type 22	0110011001
Type 23	0110010110
Type 24	0110010101
Type 25	0101101010
Type 26	0101101001
Type 27	0101100110
Type 28	0101100101
Type 29	0101011010
Type 30	0101011001
Type 31	0101010110
Type 32	0101010101

FIGURE 3: 5 BINARY VARIABLES ARE USED TO DEFINE 32 POSSIBLE TYPES OF HUMAN FIGURES.

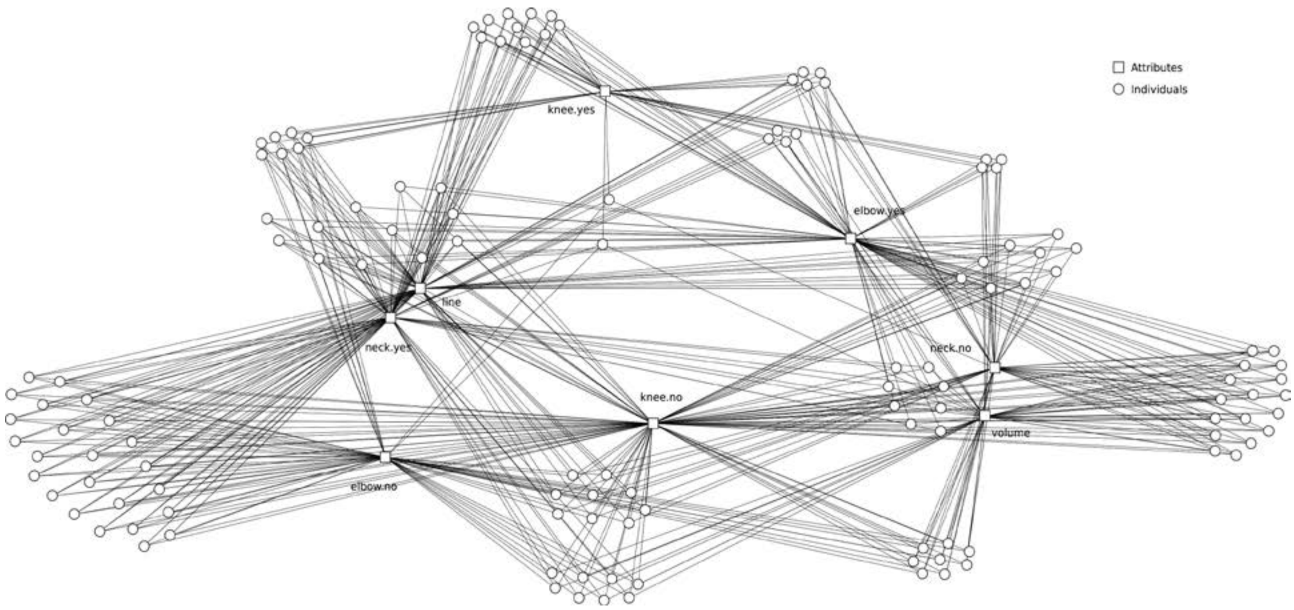


FIGURE 4: MORPHOLOGICAL 2-MODE NETWORK.

their classification is now based on features common to all. As a consequence, these combinations can also lump together or split typological classes, depending on the share of morphological characteristics in their definition.

Other layers based on binary variables can be related to different aspects of the data. Here, morphology is linked to the different choices made by prehistoric groups to represent the human body. Some of them may have defined rules for a correct representation, in the way Greek and Renaissance artists established a mathematical proportion between the head and body sizes. Others, meanwhile, may have discarded such need. As we already identified how different head shapes have been used to represent different individuals, we can now analyse their relation with the general morphology and look for eventual patterns.

We mounted a network with two different modes: one for the 8 possible attributes and the other for the 128 individuals (Figure 4). In 2-mode networks, also called affiliation networks, edges can only exist between vertices of two different modes. There won't ever been a link between two individuals, nor between two attributes (Latapy *et al.,* 2008). In this case, while every individual must necessarily have five edges, the number for attributes may range from 0 to 128. As a consequence, degree measure need not to be applied to individuals.

This network gives us a good representation of the morphological variability of the human figures from the Toca do Pepino. On the graph, most individuals are clustered in components, but some show unique composition and are then isolated. Every site analysed this way will show a different topology, based on the idiosyncratic characteristics of each paintings and their overall distribution.

Toca do Pepino				
Attribute	Degree	%	Betweenness	Closeness
neck.yes	75	59	0	0,02020893
neck.no	53	41	0	0,01569912
volume	50	39	0	0,01526388
line	78	61	2	0,02079348
elbow.yes	70	55	2	0,02028981
elbow.no	58	45	0	0,01675374
knee.yes	33	26	0	0,01129451
knee.no	95	74	6	0,02378063

Figure 5: Results of some centrality measures.

Applying centrality measures to two-mode networks is not simple, as most algorithms are not made for this purpose. In order to obtain satisfactory results, we had to project the data into a weighted one-mode network where attributes are directly related through weighted edges, depending on the number of instances they share in the original data (see Opsahl & al. 2010; Newman 2001; Borgatti & Everett 1997). Betweenness and closeness measures could then be applied, and their results compared to the degree distribution from the first network (Fig. 5).

At first sight, we might think that all the centrality measures work together: as an attribute with high degree is related to many individuals, it should also have a high betweenness because it works as a bridge for many shortest paths. Nonetheless, this tendency is not completely linear, and the actual configuration of all the individuals may produce different results. Two attributes with the same degree can then show different betweenness and closeness. This is best exemplified if we look at the relation between the degree and the closeness of two positive attributes, neck and elbow: the first one has a higher degree, but a lower closeness and betweenness. Interestingly too, the most central combination of attribute – characterized by a neck, linear trunk, elbows but no knees – is not the most numerous on the network. Only 12 individuals share this combination, while the biggest components have, respectively, 24 and 17 individuals. Nonetheless, if we consider all the combinations present on the network, this archetype is the most reachable.

This particular aspect of the topology shows the first advantages of this application of network analysis. It allows us to bypass artistic impressions, and confront a simple counting of instances when defining archaeologically relevant attributes, as it offers a more precise description of the material.

2. Partitions

Beyond its topological properties, this bimodal network allows us to analyse the dispersion of particular elements, traits or characteristics present at the site. Before investigating the behaviour of head shapes, we chose different illustrative categories of data: the location on the site, and a couple of objects depicted.

xFirst, the structure of the rock shelter naturally divides the space in different areas. The two main surfaces, where human figures are located, are separated by about 5 meters. The first area is the longest and most protected from rain, with thicker sandstone layers where paintings could be made more easily. The largest paintings were also registered in this part. The second area is protected by fallen blocks, forming small platforms. Paintings were also made on the blocks.

We added these informations as a partition. On the initial network, we represented the distribution of the 128 figures according to the area where they are located. We could then analyse the variability of the site as a whole (Figure 6).

Both panels are fairly distributed on the whole network. There is a small concentration of volume figures on the first panel and linear figures on the second, leading us to think to a locational concentration of each attribute. Nonetheless, there still is a fair number of the opposite attribute on each panel. The hypothesis of two subsets of individuals with distinct morphologies located on each panel is therefore not possible.

Second, we selected two particular kinds of weapons carried by some individuals: spear-throwers and slings. Spear-throwers are one of the most common weapons observed on human figures in north-east Brazil, and their use was still well attested after the arrival of the Portuguese. Today, some modern indigenous group, in the Alto Xingu region, use them as ceremonial weapons. During the Kwarup, men from different ethnic groups throw arrows at each other (Galvão, 1950). Slings, on the other hand, are ethnographically unknown among Brazilian indigenous nations. Meanwhile, their identification in rock paintings doesn't seem to be doubtful: different movements and positions are in accordance with the manipulation of slings. We could then expect a large presence of spear-thrower, and a more specific representation of slings, maybe restricted to well define types (Figure 7).

Indeed, spear-throwers are distributed on the whole network. No particular combinations of attributes show a massive presence of this kind of weapon. As we proposed too, slings are much more concentrated, and mostly linked to linear figures. This would link them to a particular morphological pattern in rock art that need to be further investigated at other sites.

These two first examples show how analyses of dispersion of specific features on a network, in this case based on morphological attributes, can be useful for our understanding of past societies. As the Toca do Pepino rock shelter is quite rich in human representation, the density of the network makes our results less evident. On the other hand, had the site less figures, this method would highlight possible patterns with more precision.

Finally, we needed to address the main problem we introduced in this article. Two main head shapes were identified on the first scene from the neighbouring Toca da Figura rock shelter. A third one cannot be used at present because most of the individuals represented with at the Toca do Pepino do not manipulate any objects. The two

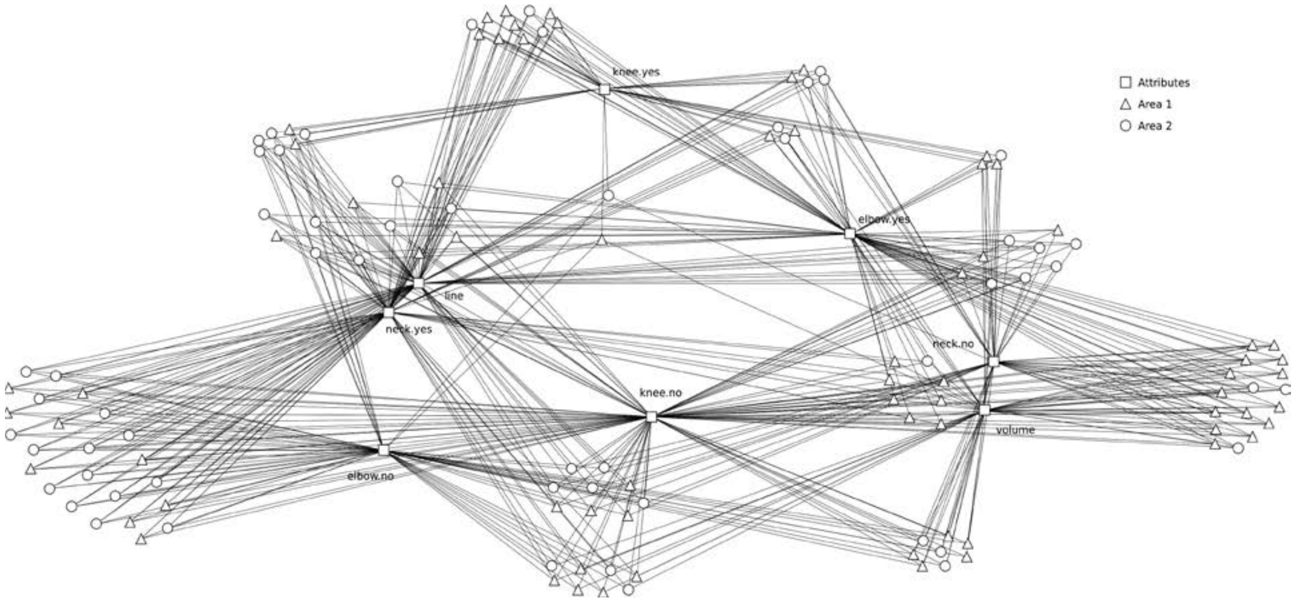


FIGURE 6: 2-MODE NETWORK WITH PARTITIONS ACCORDING TO THE DISTRIBUTION BY AREAS.

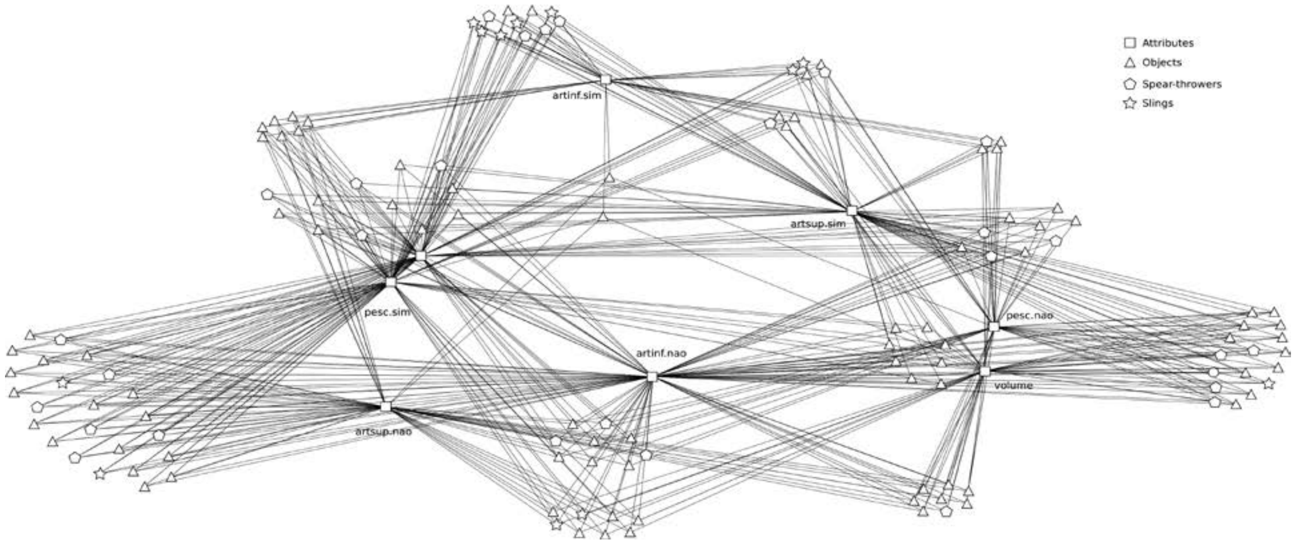


FIGURE 7: 2-MODE NETWORK WITH PARTITIONS ACCORDING TO THE USE OF WEAPONS.

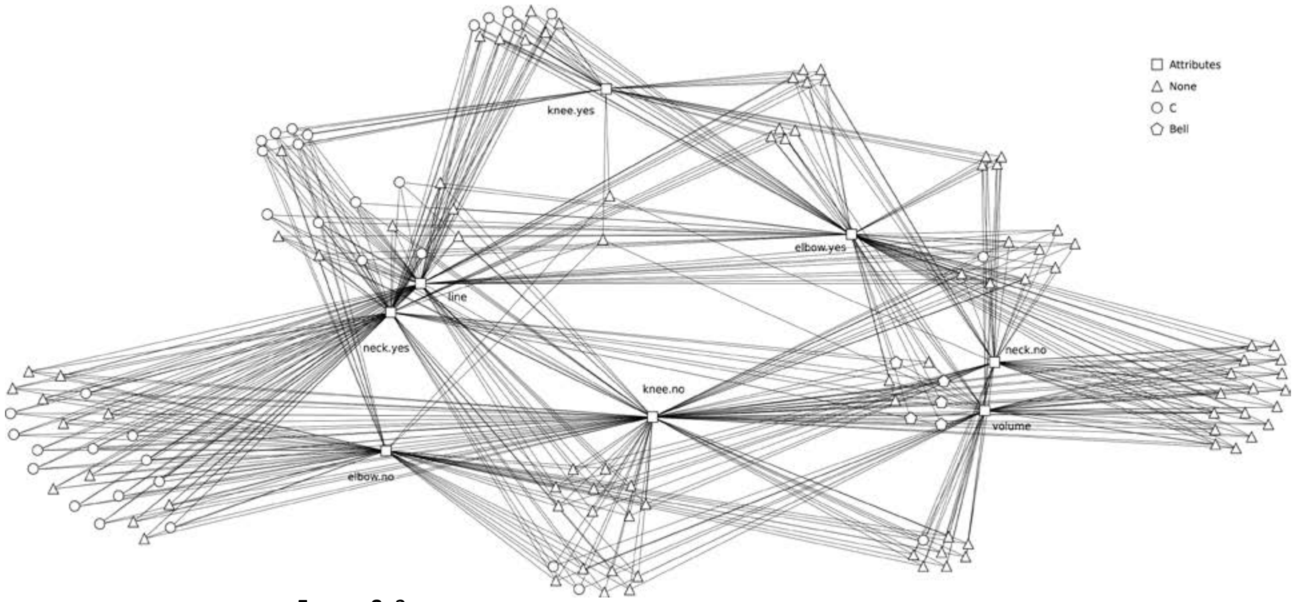


FIGURE 8: 2-MODE NETWORK WITH PARTITIONS ACCORDING TO HEAD SHAPES.

main types, C shaped and triangle shaped, are well attested (Figure 8).

Again, their respective dispersion shows interesting results. The C shaped individuals have a great variability while the triangle head shape is highly patterned. Its reduced number of occurrences could be a problem hadn't it be identified by other mean than counting.

Before discussing these results, we wanted to be absolutely sure that the network analysis wasn't flawing our observations in some respect. Balancing them with other techniques seemed an opportunity to show their robustness. This led us to test against another statistical tool that would eventually have a higher level of confidence. The same data was then used with a Multiple Correspondence Analysis (MCA), an equivalent to Factor Analysis for qualitative data (Le Roux & Rouanet 2010). As a geometric method, it would also allow us to measure the differences between the two modes of elements – the attributes and the figures.

Our data didn't need much editing in order to be fit for a MCA test, as individuals and attributes were already well defined. An MCA graph typically shows two dimensions at a time, and places each element in relation to each other. The cloud can present clusters or trends, depending of each situation. Results are interpreted in terms of these possible structures, and of the opposition between positive and negative values on each dimension (Figure 9).

At the Toca do Pepino, two main axes were considered, as they accounted for more than 50% of the total variability of our data. These dimensions surprisingly show a distribution very similar to our network, except that the latter seems to have been shrinked during the energizing process. Different from it, though, all the individuals are clustered on a single coordinate, for geometry reasons. This makes the reading a bit difficult.

We added three types of partitions as supplementary variables, so that their position on the graph would not influence the localization of the main attributes and individuals. We then considered the position of each one in relation to the whole graph.

The first partition, organizing the individual between two areas, confirms our observations. Even if there's a tendency, for each one, to be related to specific attributes, mainly volume without neck or line with neck, they are localized close to the centre of the axes. The second partition, for different kinds of weapons, shows again that spear-throwers are common to different combinations, while slings are more concentrated. According to the last partition, finally, the opposition between C and triangle individuals is located along the first axis. We can compare this distribution to the first partition, where both areas are also located along this axis. The relation between C shaped individuals and the attributes of line and neck is clear.

Multiple correspondence analyses offer an alternative geometric method to investigate the characteristics of our data. While it shows results similar to those we obtained with network analysis, there is an emphasis on structural oppositions between the attributes. Meanwhile, as we are analysing the distribution of specific elements, it lacks the possibility to identify individuals and apply measures. Both methods show that, while the spatial distribution between the two panels at the Toca do Pepino can't be related to any specific use, other elements, like triangle head shapes and slings, are in closer relation to certain morphological traits.

Considering the scene we described at the beginning of this article, the different heads shapes represented in a single line show different morphological patterns. One of these shapes is related to a specific set of morphological attributes, that is never changed, or only slightly. The other can be represented with a range of combinations, their only commonality often being the head shape itself.

Discussion

In this last part of the article, we will address two points. First, we may offer a hypothesis for the question

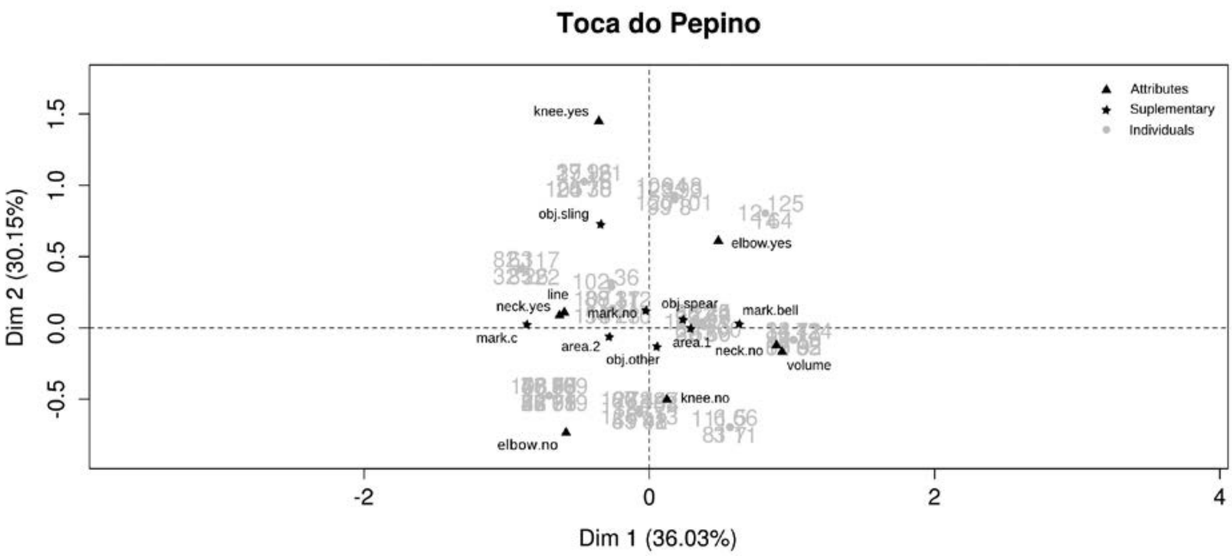


Figure 9: MCA analysis of the same data.

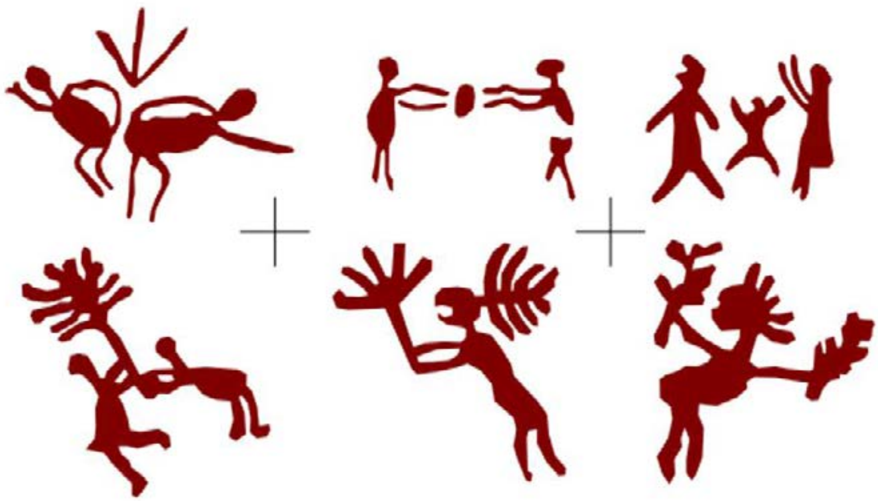


Figure 10: Examples of emblematic scenes identified at distant sites Northeast Brazil (adapted from Martin 2007).

formulated at the beginning of the text. The interactions are not just occurring between two different graphic shapes, but also between two patterns of relationships with morphological variables. What does that mean? The question now concerns their extension: how is this particular phenomenon, observed at a specific location, related to space? Is it restricted to the Ventura Valley? And, most importantly, who is interacting? Second, we will present a few limitations and prospects concerning the use of network analysis in rock art studies.

Academic research in the north-east region of Brazil began in the 1970s, first in the Serra da Capivara and later in the Seridó, with the huge work of Niède Guidon, Anne-Marie Pessis and Gabriela Martin. Many of these sites present particular sets of individuals that are highly patterned and recurring at distance, not just inside each region but in both of them. The authors called these scenes emblematic, and considered them as a marker for the definition of a Nordeste Tradition. In this context, head shape variability was also identified as a more local identity marker (see Pessis, 2003; Martin, 2007; Etchevarne, 2007; Figure 10).

The identification of emblematic scenes and the recurrence of particular head shapes mostly depend on long term research, as one particular scene can be identified at a site and remain isolated until the same configuration is found at another one. That is what actually happened with new research in the Chapada Diamantina. Previously known scenes were identified again in Ventura, and that allowed us to expand the scope of this phenomenon south of the São Francisco river (Figure 11).

The first scene is characterized by two individuals, side-by-side, and possibly a third smaller one. In this case, the scene is particularly detailed. One of the individuals is a woman wearing clothes, and the other is carrying an unidentified object. Both heads are triangle shaped. The second scene wasn't identifiable until recently. It was previously known at one site in the Serra da Capivara (Martin & Guidon, 2010), but has then been observed in six locations in the Chapada Diamantina. Despite variability in some details, the scene is repeatedly organized around the same elements: animals are circled by humans in front of a geometric design interpreted as a fence or a net. There are two representations of this kind at the Toca do Pepino, and the individuals show C shaped heads. A third scene, while in worse condition, was also identified at the Toca da Figura. As prospecting grows, new examples are found in many other places in the municipality of Morro do Chapéu (Figure 12).

These scenes, along with the results of the analysis exposed in the present article, make us rethink the nature of the social contacts among the population living in the whole north-east region during prehistory. Instead of a linear diachronic dispersion, we can consider the possibility of a synchronic interaction sphere.

A second point concerns network analysis as an instrument. We have been mapping the variability of our material according to a specific set of binary qualitative variables. Centrality measures were then used to extract a kind of information that simple counting couldn't give us, the articulation between variables and attributes in an archaeological dataset. Second, we analysed the dispersion of elements on the network, and verified their relationship to specific components. Compared to MCA, with which it shares some elements, network analysis is not only able to offer information about variability. It can also be used to quantify the level of concentration or dispersion of additional material – like, in this case, head shapes. This method must then be expanded to consider other categories of variables. Ultimately, a multilayer network can be mounted where the variability of a single element could be traced in different sets of variables. Such network could help us to bridge the gap between style and physical-

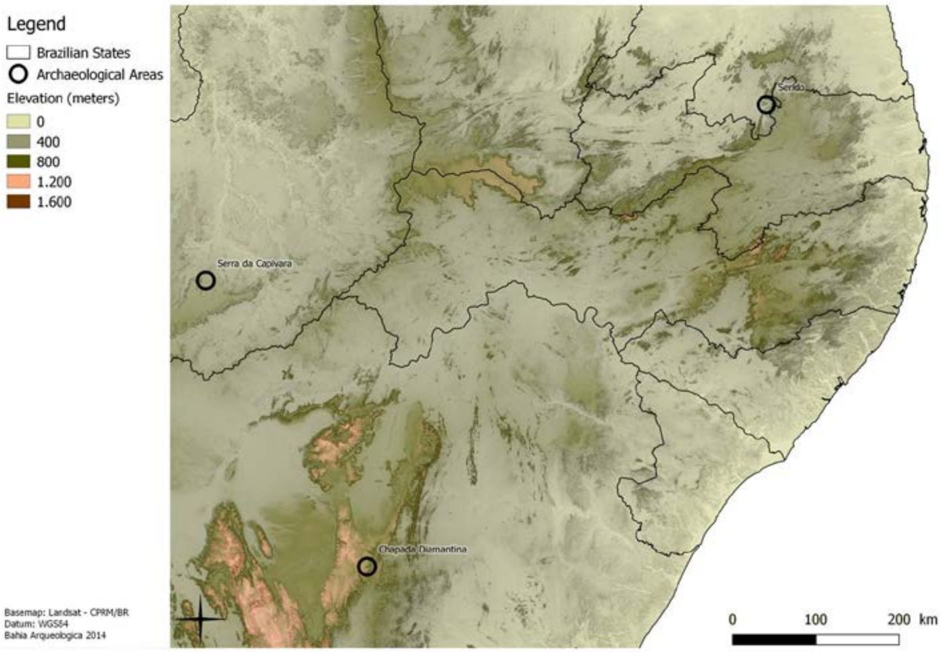


Figure 11: Map of Northeast Brazil with the three main regions where figurative rock art is studied.

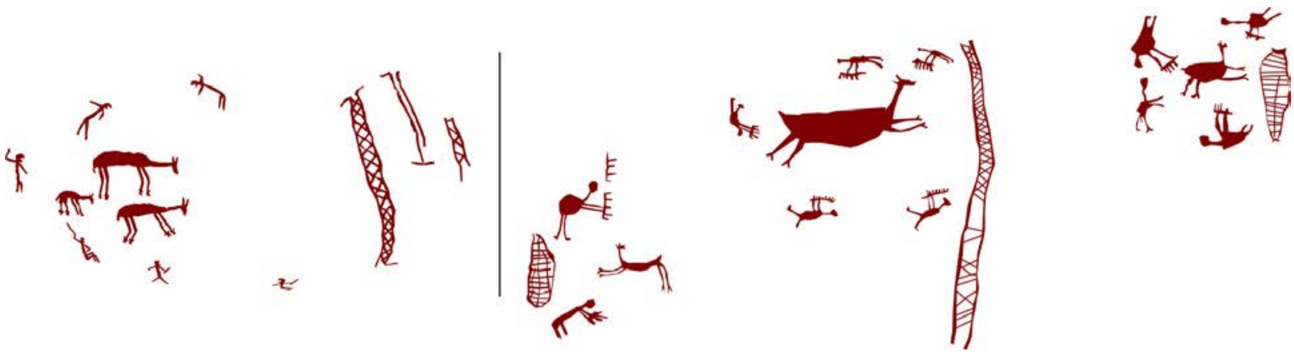


Figure 12: These hunting scenes with fences or nets are patterned in different sites (left: Toca do Pepino, Morro do Chapéu; right: Toca do Estevão III, Serra da Capivara).

chemical properties in rock art studies (for multilayer networks, see Boccaletti *et al.,* 2014; Kivelä *et al.,* 2013).

Conclusion

At intra-site level, we proposed to use network analysis to study the variability of paintings. The application of a 2-mode network based on morphological attributes allowed us to define different patterns that need to be crossed with other variables, such as symmetry and physical-chemical properties. At inter-site or regional level, a network can also be drawn upon the identification of emblematic scenes. It mostly depends, though, on the prospect of new sites with human figures represented, and on the excavation of other remains that could be directly dated. It would also be greatly aided by further developments in network analysis, such as measures for cohesion and fragmentation in a multilayer context.

References

BOCCALETTI, S., BIANCONI, G., CRIADO, R., DEL GENIO, C. I., GÓMEZ-GARDEÑES, J., ROMANCE, M., SENDIÑA-NADAL, I., WANG, Z. and ZANIN, M. (2014). The structure and dynamics of multilayer networks. *Physics Reports* 541 (1). DOI: 10.1016/j.physrep.2014.07.001.

BORGATTI, S. P., EVERETT, M. G. (1997). Network analysis of 2-mode data. *Social Networks* 19, pp. 243-269

COSTA, C. (2012). *Representações rupestres no Piemonte da Chapada Diamantina (Bahia, Brasil)*. Doctoral Thesis. Coimbra: Universidade de Coimbra.

ETCHEVARNE, C. (2007). E*scrito na pedra. Cor, forma e movimento nos grafismos rupestres da Bahia.* Rio de Janeiro: Versal.

ETCHEVARNE, C. (2009). As particularidades das expressões gráficas rupestres da tradição nordeste, em Morro do Chapéu, Bahia. *Clio Arqueológica* 29, pp. 41-59.

GALVÃO, E. (1979). O uso do propulsor entre as tribos do Alto Xingu. In: Galvão, E. (ed.). *Encontro de sociedades.* Rio de Janeiro: Paz e Terra.

GUIDON, N. (1984). L'art rupestre du Piauí dans le contexte sudamericain. Une première proposition concernant méthodes et terminologie. Doctoral Thesis. Paris: Université de Paris Panthéon-Sorbonne.

KESTERLING, C. (2007). *Identidade dos grupos pré-históricos de Sobradinho – BA*. Doctoral Thesis. Recife: Universidade Federal de Pernambuco.

KIVELÄ, M., ARENAS, A., BARTHELEMY, M., GLEESON, J. P., MORENO, Y. and PORTER M. A. *Multilayer networks.* arXiv 1309.7233. Available at: http://arxiv.org/abs/1309.7233 [Accessed: 19 July 2014].

LATAPY, M., MAGNIEN, C., DEL VECCHIO, N. (2008), Basic notions for the analysis of large two-mode networks. *Social Networks* 30 (1), pp. 31-48.

LE ROUX, B. and ROUANET, H. (2010). *Multiple correspondence analysis.* London : SAGE Publications.

MARTIN, G. (2007) *Pré-história do Nordeste do Brasil.* Recife: Editora da Universidade Federal de Pernambuco.

MARTIN, G. and GUIDON, N. (2010). A onça e as orantes. Uma revisão das classificações tradicionais dos registros rupestres do Nordeste do Brasil. *Clio Arqueológica* 25 (1), pp. 11-30.

NEWMAN, M. E. J. (2001). Scientific collaboration networks. II. Shortest paths, weighted networks, and centrality. *Physical Review E* 64, 016132. DOI:10.1103/physreve.64.016132

OPSAHL: T., AGNEESSENS, F. and SKVORETZ, J. (2010). Node centrality in weighted networks: Generalizing degree and shortest paths. *Social Networks* 32 (3), pp. 245-251.

OTTONI, E. and IZAR, P. (2008). Capuchin monkey use of tools: overview and implications. *Evolutionary Anthropology* 17 (4), pp. 171-178.

PESSIS, A.-M. (1987). *Art rupestre préhistorique: premiers registres de la mise en scène.* Doctoral Thesis. Paris: Université de Paris X Nanterre.

PESSIS, A.-M. (2003). *Imagens da Pré-História. Parque Nacional Serra da Capivara.* São Paulo: Gráfica Takano.

RIBEIRO, B. (1989). *Arte indígena, linguagem visual.* São Paulo: Editora da Universidade de São Paulo.

SCHORTMAN, E. (1989). Interregional interaction in prehistory: the need for a new perspective. American *Antiquity* 54 (1), pp. 52-65.

WIESSNER, P. (1983). Style and social information in Kalahari San projectile points. *American Antiquity* 48 (2), pp. 253-276.

WOBST, M. (1977). Stylistic behavior and information exchange. In: Cleland, C. (ed.). *For the director, research essays in honor of James B. Griffin.* University of Michigan Museum of Anthropology Anthropological Papers 61. Ann Arbor: University of Michigan.