Book of abstracts

Concepts, methods and tools
Ontologies and standards for improving interoperability of archaeological data: from models towards practical experiences in various contexts

(Thursday 24th 8h30-19h Pantheon S6)

Chairs: Anne-Violaine Szabados, Katell Briatte, Maria Emilia Masci, Christophe Tuffery

1: Archéologies et Sciences de l’Antiquité (ArScAn) CNRS: UMR7041, MAE-Maison René Ginouvès. NANTERRE - France
2: Fondation internationale pour le Lexicon Iconographicum Mythologiae Classicae (LIMC) MAE-Maison René Ginouvès. NANTERRE - France
3: DGP – DSIP Ministère de la Culture et de la Communication
4: Scuola Normale Superiore di Pisa (SNS)
5: Institut National de Recherches Archéologiques Préventives (INRAP)
6: Cités, Territoires, Environnement et Sociétés (CITERES) CNRS : UMR7324, Université François Rabelais - Tours

The documentation and data of archaeology - planned or preventive - and of cultural heritage, lead to the need for technical and semantic interoperability. The development and the use of ontologies, standards and languages devoted to archaeology intend to facilitate the integration, exchange and sustainability of cultural information. Some researchers have explored these approaches during the previous CAA (Perth 2013 sessions 4, 5 and 6). Archaeology, and its documentation and data, are multidisciplinary and heterogeneous. In order to make the best use of conceptual reference models and their implementation in standards, a balance must be found between general approaches, which allow promote interdisciplinary exchanges, and specialized ones that preserve the unique aspects of archaeological research (contextualization, uncertainty, incompleteness, evolution analysis, review of dating, different spatial scales, subjectivity methods, observation and or interpretation, allocation, etc.). The use of ontologies and standards for treatment of cultural heritage information (CIDOC-CRM / ISO 21127; Europeana Data Model ...), spatial information (INSPIRE, ISO 19115, GML, GeoSPARQL ...) or languages and thesauri (OWL, SKOS ...) already has a proven basis but it may be insufficient. During their implementation in recording, study and management applications using archaeological data, various standards and conceptual models should ensure that archaeologists can choose the level of representation of information, from operational to continental scales. To encourage the use by cultural heritage specialists of languages and construction principles of ontologies, works specific to their domains have to be shown. This will involve presenting standards, and models, but also tools developed on recognized standards already adopted by the scientific community and cultural institutions. The use and reuse of archaeological data encoded in a standardized way, and metadata widely available also raises several points of discussion: the integrity of archaeological data, the risk of misinterpretation, the quality of recorded data, the need for labeling data, the level of specialization in data description and levels of generality imposed by interoperability, etc. The purpose of using ontologies and standards is to enable interoperability and a large but controlled diffusion of archaeological and cultural heritage data in the general context.
of wide diffusion of big and open data. The session's goal is to provide practical use cases of ontologies and standards to address these issues. Theoretical aspects too difficult to be transposed will be avoided. Papers may focus on the more practical issues in the construction of ontologies, models and applications, from experience feedback analyzed, and available tools built on such approaches. Finally, communications will enable the various actors of archaeological research and cultural heritage to identify and evaluate, what can be used if they wish to engage in the use of standards for interoperability and wide dissemination of their data.
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 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.

There is no centralised register of archaeological fieldwork in Sweden, making it difficult not only to keep track of what is going on where, but also to know what fieldwork ? if any ? has taken place in connection to a particular site in the national sites and monuments record. Sweden also has no central digital archive for the storage of either archaeological fieldwork data or reports; as such records are now produced digitally, valuable archaeological data is thus increasingly at risk of being lost.

Furthermore, despite the fact that almost all of the data and administrative metadata surrounding archaeological work are digital-born, they are still handled according to analogue paradigms, particularly when information must be shared between different organisations. Sources of archaeological data which are currently made available digitally by various national and local bodies are not typically linked together. This leads to inefficiencies in information transfer, duplication of data and effort, and to information describing the same ‘objects’ being stored in different systems within different organisations.

The DAP programme intends to address these problems over the course of a five-year period, using standardised platform-agnostic data formats and protocols to streamline information transfer between organisations, by releasing 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 by creating a secure digital repository both for the raw data and reports arising from fieldwork and research. We aim to make this information freely available as linked open data. Our overall mapping of the current Swedish archaeological process is complete (although some details remain) and we are currently working on a conceptual model on which our future information architecture will be based. In parallel, we are also working to translate and release our existing (analogue) archaeological taxonomies to SKOS and release them as linked open data authorities, beginning with the Swedish monuments types thesaurus.

This paper will describe the work that we have accomplished so far and the progress we have made, as well as our plans for the rest of the programme. In particular, it will focus on our mapping of the existing archaeological processes in place in Sweden and our plans to digitise them, our ongoing translation of existing controlled vocabularies to SKOS, and our use of the semantic web as a fundamental method.
ing the associated implementation specifications. The use of such a scheme has a number of advantages: savings of analysis and modelling effort, subsequent use of the available software and a high level of semantic and syntactic interoperability. In other cases, an existing conceptual scheme cannot be used as a fall back option. Traditionally a project-specified data model will then be developed and described in a formalized manner (today mostly with UML) in an often protracted, complex process. A highly specialised but more inflexible data model also arises as a result of this analysis.

Independent of whether a standard is used for conceptual modelling, this approach requires that before the implementation of the information system, an agreement is achieved regarding the professional view on the «Universe of Discourse» and its scope. However, this pre-condition is not met if the information system needs to be flexibly used for numerous projects in different areas of application and for user groups with non-harmonised professional perspectives. This situation typically occurs if the information system is part of the service offer in an infrastructure, e.g. for geographic data or archaeological research data. Then a solution will be needed that is, on the one hand, sufficiently general, generic and thereby flexible but, on the other hand, supports the domain-specific standards with the aim of an interoperable exchange of information.

This presentation will firstly contrast properties of specialised and generic models. In doing so, conditions, areas of application, modelling principles and effects on the operation of the information system will be discussed. Subsequently, specific approaches of an archaeological research data infrastructure project (OpenInfRA) and of INSPIRE (Infrastructure for Spatial Information in the European Community) will be explained. Alongside modelling details, this presentation will expand on the effects of the generic approach on the associated database and XML-scheme as well as the application. Finally the consideration of relevant standards will be presented.

linkedARC.net: addressing the standards question in archaeological digital data management using Linked Open Data

Lynam Frank
Trinity College Dublin (TRINITY COLLEGE DUBLIN) Trinity College Dublin, College Green, Dublin 2, Ireland http://www.tcd.ie/

The employment of standards by the archaeological digital data management community is by no means the logical and uncontentious design decision that it might at first appear. Whether one chooses to side with the pro or anti standards camps says a lot about their fundamental beliefs on how archaeology should be practiced. On the pro side, standards promise to greatly ease the load of the archaeologist who seeks out broad-ranging generalities and patterns within the material record of the past. For this group, the archaeological method exists within the positivist scientific space; data is fed into processes and truths are produced at the other end. On the other side of the argument, advocates of the postmodernist archaeological method eschew the very constraint and simplicity that the pro-standards camp craves. Insight, for this second group, can only be arrived at through a process of informed contextual and specific analysis; what works for one material-set by definition cannot explain and provide structure for the material of another point in space or time. Therefore, the standard for the postprocessualist is a potentially deterministic methodological tool that is not to be trusted and one that should be avoided as much as is practically possible.

As an archaeologist, I have struggled with the implications of this methodological reality since entering the discipline. At a philosophical level, I see myself as a hermeneutic archaeologist of the Hodder-Shanks-Tilley school of thought and yet as an every-day practitioner of archaeological digital data management, I am drawn to the structure and clarity that the standard promises, and I would suspect that this internal dilemma is not unique to myself. This paper presents a summary of my digital data management investigations over the last number of years within the context of this debate. Specifically, I am interested in how the principles of Linked Open Data might be implemented to alleviate some of these tensions.

linkedARC.net is a LOD-compliant server architecture that I have designed to implement the theoretical positions generated by my research. Data is housed within a relational database core and is exposed to the digital world as a part of the greater data mesh that we know as the Semantic Web. linkedARC.net’s data is semantically transparent, being structured by public ontologies. A
web app frontend consumes and contributes to the data housed within the linkedARC.net core. The test dataset for the project comes from the excavation and post-excavation records from Priniatikos Pyrgos, a highly complex multi-period site in East Crete. The challenge of designing the web app and the underlying server architecture has been manifest in the provision of a level of structure that the complexity of the data necessitates, while also allowing for the interpretive scope demanded by the researchers active on the project. As such, it is a very practical example of an implementation of cutting edge LOD technology within archaeology that has required a constant attention to the philosophical questions that surround this space.

**Dykes of standards supporting polders of data: the practices used in the Netherlands for making archaeological data available and accessible**

*Gilissen Valentijn*

Data Archiving and Networked Services (DANS) Postbus 93067 2509 AB Den Haag [http://www.dans.knaw.nl/](http://www.dans.knaw.nl/)

In the Netherlands, the archiving and publication of archaeological research data has led to the establishment of EDNA, the e-Depot for Dutch Archaeology [http://www.edna.nl/](http://www.edna.nl/), accommodated at DANS (Data Archiving and Networked Services) [http://dans.knaw.nl/en](http://dans.knaw.nl/en).

EDNA is a collaboration between DANS and the Cultural Heritage Agency (RCE). EDNA was set up as a project within the broader organisation of DANS. DANS is an institute of the Royal Netherlands Academy of Arts and Sciences (KNAW) and the Netherlands Organisation for Scientific Research (NWO). DANS serves as the archive for scientific data. Among other services, DANS maintains its on-line archiving system EASY (Electronic Archiving System, https://easy.dans.knaw.nl) which allows researchers to deposit and describe their own datasets.

The mission of DANS is to enhance permanent access to digital research data. DANS stimulates cooperation between data producers and users and does research into long-term accessibility. By participating in international projects such as CARARE and ARIADNE, DANS is continuously improving on the options for finding, accessing and re-using (archaeological) data. The existing infrastructure of the e-depot for Dutch Archaeology allows for sharing of good practices such as long-term preservation, data organisation and data dissemination for accessibility.

Agreements in order to fulfil this mission have been laid down in the KNA: the quality standard for Dutch archaeology. This is not the only standard which servers to structure and strengthen the Dutch data flow. Datasets are described in EASY using the Qualified Dublin Core metadata standard. The archiving process involves file format migration to preferred formats according to DANS policy with attention to file format aspects for long-term preservation and long-term accessibility. DANS continuously strives for international certification, by means of applying the (self-) audit and assessment of the TRAC (Trusted Repositories Audit & Certification) and by following the guidelines of the DSA (Data Seal of Approval). Additionally, DANS participates in the implementation of a Dutch national protocol of exchanging excavation data.

These standards and practices make for the foundations of a successful running service in which durable archiving and unlocking of all digital documentation of the archaeological research is ensured. By 2013, EASY provides online access to more than 20,000 datasets: 17,000 archaeological reports and 3,000 large datasets consisting of data of excavations and explorations (photos, GIS, data-tables, drawings). Both the research descriptions and all data can be downloaded via the online archiving system EASY.

**Integration of Archaeological Datasets Through the Gradual Refinement of Models**

*Gonzalez-Perez Cesar, Martín-Rodilla Patricia*

Institute of Heritage Sciences, Spanish National Research Council (Incipit, CSIC) [http://www.incipit.csic.es](http://www.incipit.csic.es)

The problem of integrating disparate archaeological datasets has received a great deal of attention lately. This is a relevant problem, and we do not dispute the advantages of being able to query and process information that is uniformly accessed from a wide array of distinct sources. However,
we observe that most, if not all, of the approaches to integrate archaeological datasets comprise extremely ad hoc efforts that suffer from one of the following two issues, and sometimes both. The first issue that we have observed is that, in order to integrate different datasets, a new, distinct model capable of catering for all of them is created. This results in the paradox of generating yet one more unique, non-standard view on the data. The second issue that we have observed is that, when carrying out the integration, individual datasets are mapped to a reference model such as CIDOC CRM [1], bridging the large gap that exists between the specific implementation (often at the database level) and the conceptual expression, which is costly and error prone.

In this paper we propose an alternative approach to archaeological dataset integration, which introduces the notion of the gradual refinement of models. In order to integrate a collection of disparate datasets, we propose that a conceptual model of each dataset is first created; these models must be extensions of a given abstract reference model such as CHARM [2], so that they capture the peculiarities of each dataset, but at the same time they share a common conceptual framework. Then, these particular models are examined for similarities, and more abstract models are generalised from those that are the closest. This can be repeated as necessary until, eventually, a hierarchy of models is finally obtained, and a number of views on the data are generated at different levels of abstraction, each one encompassing more datasets the more abstract it is.

The paper will illustrate this approach through the integration of three real datasets: a subset of the SIA+, the 20-year old, half-million object dataset that Incipit CSIC uses as their regular archaeological database; the «Camiño Primitivo» database, a purpose-built store also by Incipit CSIC for the cataloguing of sites and associated ethnographic elements along the Way of St. James in Spain; and the «Town and Country in Roman Essex» dataset offered by the Archaeological Data Service of the University of York, a Roman settlements database from Essex in the UK. A particular model will be presented for each of the three datasets. Then an abstract particular model will be generalized from the two datasets belonging to Incipit CSIC, since it is assumed that, having been created within the same organisation, they should share a significant number of conceptual commonalities.

Integrated querying and information processing will be demonstrated at the top level of abstraction, involving the three datasets, as well as at an intermediate level of abstraction, which involves only the two datasets at Incipit CSIC but provides much richer and more specific semantics.

Building comprehensive management systems for cultural ? historical information

Bekiarí Chryssoula 1*, Doerr Martin 2, Agelakis Dimitris 3, Karagianni Flora 4

1 : Institute of Computer Science, Foundation for Research and Technology - Hellas (FORTH-ICS)
N. Plastira 100, Vassilika Votýno, GR 700-13, Heraklion, Crete

2 : Institute of Computer Science, Foundation for Research and Technology - Hellas (FORTH-ICS)
N. Plastira 100, Vassilika Votýno, GR-700 13 Heraklion, Crete, Greece
www.ics.forth.gr/isl

3 : Institute of Computer Science, Foundation for Research and Technology - Hellas (FORTH-ICS)
N. Plastira 100, Vassilika Votýno, Heraklion GR-700 13, Greece

4 : European Centre for Byzantine and Post-Byzantine Monuments (EKBMM)
Leoforos Stratou 2, 546 40, Thessaloniki, Greece
www.ekbmm.gr

* : Corresponding author

Nowadays the digital scientific documentation of archaeological sites, places and museum objects in digital form is a reality. Still a very small part of the basic information is exploited by the public. The reasons are many. One of them is based on the fact that in most digital repositories about museum objects and standing monuments, the basic scientific information stored in them is only a fraction of the Information found in old reports, catalogues, photos and grey literature. For instance, a dairy tells parts of the day to day story of an excavation. A restoration study of a historical building usually contains the archaeologist’s or architect’s interpretation of the building expressed as a story based on the restored structures. A lot of human effort is needed in order this information to be documented and stored in digital repositories. But even if such an effort is achieved, there is a lack of information systems to support in a uniform, transparent and consistent way the lifecycle of information found, produced and used by professionals, scholars and interested public, which deals with all aspects of human history and activity, including products, art objects,
science, thoughts and ideas. The «classical» cultural information systems mainly address the
collection management task. They provide functions well defined by various standards (Spectrum
MDA, Midas, etc..) but they don’t support any particular presentation and « telling story» concept.
They are normally based on RDBMS, and there are dozen commercial providers.
This paper describes the requirements for data organization and management for knowledge col-
clection and promotion along with methods of user interaction, and presents a system design and
functions that supports the whole life cycle of information and knowledge including the primary
recording and scientific documentation of cultural objects and sites as well as data interlinking,
presentation and the knowledge enhancement.
Practical considerations are presented about the data models and procedures which are based
on CIDOC CRM and other standards proposed by semantic web for collection, interlinking, pre-
sentation and enhancement of knowledge in a transparent. Also workflows requirements for ad-
ministrational and scholarly processes are discussed.
The proposed guidelines and considerations have been evaluated and verified by domain experts
during the implementation of a comprehensive management system about Byzantine Culture.
This information system has been implemented by FORTH and has been installed in European
Centre for Byzantine and Post Byzantine monuments, in Byzantine Museum in Thessaloniki and
in Athens. It supports the scientific documentation for byzantine monuments and provides digital
content based in archives, diaries and publications for promoting Byzantine culture via interactive
information services. The e-services it offers are designed to inform teachers, researchers,
schoolchildren and interested members of the general public about monuments, personalities,
historical events, achievements and aspects of daily life in Byzantium. All site content is viewable in
both English and Greek.

To Boldly or Bravely Go? Experiences of using Semantic Technologies
for Archaeological Resources

May Keith
English Heritage (& University of South Wales) (EH)
Fort Cumberland, Fort Cumberland Road, Portsmouth. PO4 9LD

Research on the uses of Semantic Technologies for archaeological data presentation, search and
reporting has advanced considerably in recent years. This has included developing the
technologies for semantically enhanced cross-searching enabled by using the CIDOC CRM
ontology with archaeological extensions such as the CRM-EH developed at English Heritage.
This work, in particular during the Semantic Technologies for Archaeological Resources project
(STAR project http://hypermedia.research.southwales.ac.uk/kos/star/), focused on some of the
most common systems for archaeological recording in the UK, each using a version of the Single
Context Recording methodology introduced by the Museum of London in the early 1990s. In
addition this work was able to incorporate information in the semantic cross-search from free
text descriptions derived from a corpus of archaeological reports (oasis.ac.uk/) that had been
mapped to the same common ontologies, using Natural Language Processing techniques.
At a conceptual level much of the archaeological data recorded using the single context recording
methodology are inter-related by identifying and documenting single units of stratigraphic mate-
rial as recognised and interpreted by the record maker.
Underlying the work of mapping ‘single context’ data from different archaeological database sys-
tems to a common ontology there may be even more fundamental questions about how to share
data derived using different archaeological recording methodologies.
One of the associated issues that emerged from the STAR work was the need to improve the
syntactic integrity of archaeological data sets to make best use of the semantic technologies for
interoperability. Practical work in developing the use of the W3C SKOS standard for incorporat-
ing controlled terminologies with the CIDOC CRM in the STELLAR project (http://hypermedia.re-
search.southwales.ac.uk/kos/stellar/) and subsequent SENESCHAL project (http://hypermedia.
research.southwales.ac.uk/kos/SENESCHAL/), has led to the development of SKOS based ver-
sions of national heritage domain controlled vocabularies and the publishing of these as Linked
Open Data via the HeritageData.org web site (http://www.heritagedata.org/ blog/).
This paper will be based upon practical experiences of modelling the single context recording
system at English Heritage and mapping it to other ‘single context’ based systems. It will also present recent work on identifying conceptual commonalities that may exist in different archaeological recording methodologies, whether ‘single context recording’ or otherwise, along with practical challenges based on experiences of trying to integrate, or simply search across, data from different archaeological recording systems. In addition it will introduce the work to date on developing Heritagedata.org and suggest opportunities for sharing and aligning further archaeological vocabularies using SKOS and Linked Open Data technologies.

**Aligning the Academy with the Cultural Heritage Sector through the CIDOC CRM and Semantic Web technology.**

Oldman Dominic 1*, Rahtz Sebastian 2*  
1: British Museum (BM)  
www.britishmuseum.org  
2: IT Services, University of Oxford  
http://www.oucs.ox.ac.uk/  
*: Corresponding author

CLAROS[1] is a research initiative led by the University of Oxford’s e-Research Centre, and funded by a small grant from Oxford’s Fell Fund, that demonstrates the aggregation of nine European datasets, largely relating to Greek and Roman antiquity, by building a semantic web implementation based on use of the CIDOC CRM ontology. The aim of CLAROS was to provide a platform for experimentation with interfaces for world art, innovative searching techniques such as shape analysis, and testing RDF databases.

The ResearchSpace[2] project is based at the British Museum, London and is funded by the Andrew W. Mellon Foundation. Its aim is to create a shared infrastructure and service that supports collaborative research. This includes the development of communication and research tools, access to digital methods, and the harmonisation of data to create a knowledge environment akin to a modern day Digital Wunderkammer, transforming individual institutional knowledge into a network of connections and meanings that extend the questions that we can ask and the research directions that are available. Much of this data is likely to be available for general reuse as Linked Open Data.

The CLAROS project established the credentials of the CIDOC CRM standard as a semantic framework that can harmonise data from many different institutions while providing a richer environment (when compared to its digital sources) in which to explore and research cultural heritage data. Together with the ResearchSpace project this initial work is set to expand to include broader and denser datasets and crucially provide greater interactivity to support sophisticated research methods. With improvements in semantic database technology, greater cooperation between the academy and the cultural heritage sector and the maturity of the CIDOC CRM ontology, a more effective realisation of the research paradigm can be established. This can align discovery, argument and belief generated by research projects with the information systems of cultural heritage institutions for the benefit of future research, improved institutional knowledge resources and a radical improvement of engagement with the public using rich contextual information. This dynamic improves the relevancy of both Humanities research and the cultural heritage sector as a whole.

The software already developed in the CLAROS and the ResearchSpace initiatives can be used as demonstrators of the potential of the CIDOC CRM and provide important reference points for engaging both academic research and cultural heritage communities. By showing how the boundaries of computer interpretation and data modelling are being pushed by semantic technology and intelligent ontology, a cross disciplinary dialogue can be established with a range of experts including archaeologists, anthropologists and art historians, to generate new ideas for the collaborative exploitation of this technology.

The Portable Antiquities Scheme (PAS) database (http://finds.org.uk) now holds details for over 900,000 archaeological objects found by the general public within England and Wales. These objects are richly documented through over 200 metadata fields, which provide a glorious opportunity for the creation of linked archaeological data. This paper will discuss how the author has been inspired by various projects around the Globe in the last two years to enable data to be mapped to the CIDOC-CRM and linked to a wide array of ontologies and resources and to supply our data to high profile academic projects such as Pelagios (http://pelagios-project.blogspot.co.uk/). This paper will discuss the ontological choices made, the pros and cons of the chosen and the methods employed to enable their integration (for example the use of OpenRefine). It will then move in to discuss exemplars within the field that have proved to be an inspiration for the author. Firstly, lessons have been learnt from the Mellon funded ResearchSpace project (http://researchspace.org) with Portable Antiquities data being mapped in a similar fashion to British Museum CIDOC-CRM rdf. Secondly, the author was inspired by assembled intellectual geniuses that came together for the Linked Ancient World Data Institute in 2012 (ISAW, New York University) and 2013 (Drew University, New Jersey) to discuss linking Classical world resources together to provide a richer foundation for academic research. This paper will show how the PAS data has been linked to a wide variety of resources such as Pleiades (http://pleiades.stoa.org), the Ordnance Survey (http://data.ordnancesurvey.co.uk), the British Museum (http://collection.britishmuseum.org), dbpedia (http://dbpedia.org) and Nomisma (http://nomisma.org). This paper will also examine the technological choices that have been made for delivering and producing this RDF dataset (for example via Apache SOLR indexes and the use of XSLT) to a wider audience, how Linked Data is consumed (through the use of SPARQL and opensource software packages such as EasyRDF) within the PAS website for enriched pages and the software used for implementing a triple store. It will attempt to demonstrate that this process was not too difficult and provide a real world example of how to emulate it for other small archaeological finds datasets.

The interoperability of the ArSol database (Soil Archives): reflections and feedback experiences on the use of the CIDOC-CRM ontology and the integration of thesauri

Le Goff Emeline 1*, Marlet Olivier 2*, Rodier Xavier 2*, Husi Philippe 2*
1 : UMR 7324 CITERES (CITERES) CNRS/Université François Rabelais - Tours MSH Val de Loire Université François Rabelais BP60449 37204 Tours cedex 03 http://citeres.univ-tours.fr
2 : UMR7324 CITERES (UMR7324 CITERES-LAT) Université François Rabelais - Tours-CNRS * : Corresponding author

The ArSol system (Soil Archives) was developed by the «Archéologie et Territoires» Laboratory (CNRS - Tours University) for processing archaeological data. It can be used for all stratigraphic excavations and has the dual purpose of data management and research. Since its elaboration in 1990, access to data constituted one of the priorities. The software used for the DataBase Management System (DBMS) is 4th Dimension TM (4D). It was chosen because it enables a client/server and multi-platform application and can be accessed via the internet.

Joining the trend of exchanges of data and interoperability of applications, ArSol will become a system in compliance with computing standards, both in its technical aspects (ontology) and semantic aspects (thesaurus). In this perspective, we are a partner of the MASA consortium (Memory of the archaeologists and the archaeological sites) which aims towards a unified access to digitized corpuses, and is interested in the CIDOC-CRM (Conceptual Reference Model of the International Committee of Documentation).
In this context, the complexity of ArSol then becomes a difficulty: as it was elaborated in order to take into account the needs for stratigraphic and artifact data as well as to the whole documentation constituting excavations archives, this system was constructed as an open system; it is flexible and above all not conditioned by the integration of predefined thesauri. Wishing not to modify the conception logic of the recording structure of ArSol which is based on the manual stratigraphic recording system in use since 1969 on excavations in Tours, it was decided to develop an additional module in order to satisfy interoperability requirements. The first step consists in the selection of the main information included in ArSol: features and artifacts. Secondly, a stage of mapping aims to establishing the correspondences between fields necessary for the use of the CIDOC-CRM standard and the available fields in our application. In 2014, the integration of the CRM in the database ArSol will thus be realized through a module allowing to export a selection of data towards a RDF version built according to the CIDOC-CRM and structuring as well the value of fields endowed with list of values. This last operation aims at insuring a compatibility with thesauri used by the community of the Social and Human Sciences - in particular PACTOLS. Joined into the ArSol-CRM module, these reference tables will constitute a footbridge to allow outer users to reach all the connected information, listed in ArSol. To summarise, this paper will present our reflections as for the construction of the ArSol-CRM ontology, and the methods used to manage semantic equivalences within the framework of the compatibility with thesauri. This experience raises numerous questions around the perspectives and around the sustainability of the development of this kind of specific application, in particular in the field of archaeology. It is part of a valuation practice of the data while respecting their integrity and without forgetting their intrinsic purpose which is to answer research questions at the scale of a site or of a territory.

Geosemantic Tools for Archaeological Research (GSTAR)

Cripps Paul 1,2, Tudhope Douglas 1*

1 : Hypermedia Research Unit, University of South Wales Pontypridd Wales, UK CF37 1DL http://hypermedia.research.southwales.ac.uk/
2 : Archaeogeomancy http://www.archaeogeomancy.net/
*: Corresponding author

Much work has been undertaken over the past decade relating to the application of semantic approaches to archaeological data resources, notably by English Heritage and the University of South Wales. These two organisations, over the course of a number of projects, developed an archaeological extension to the CIDOC CRM ontology through the Ontological Modelling Project (Cripps & May, 2010), then applied this to a number of archaeological resources through the subsequent STAR project (May, Binding and Tudhope, 2011), implementing tools to facilitate integration of other resources through the STELLAR project (May, Binding, Tudhope, & Jeffrey, 2012), and now, in partnership with the Bespoke HER User Group, RCAHMS, RCAHMW and Wessex Archaeology, are implementing SKOS based vocabularies and associated tools to enable the augmentation of these semantic resources through the SENESCHAL project. From the outset, it was observed that the spatial component of archaeological data would be a key element, archaeological data being inherently spatial in nature. To date, most current applications of spatial semantics in the heritage sector have focussed on place names and named locations for sites and monuments and object provenances. The GSTAR project aims to extend semantic approaches to archaeological data fully into the geospatial domain and is instead focussing on the detailed spatial data emerging from archaeological excavation and survey work and is investigating approaches for the creation, use, management and dissemination of such spatial data within a geosemantic framework, building on the CIDOC CRM, with particular reference to sharing and integration of disparate resources. This paper will present work to date in the first year of the GSTAR project. This has been centred on the identification of suitable platforms and methods for the integration of semantic and geospatial data including comparisons of different approaches emerging from the Semantic Web and Geospatial research communities. Testing and prototyping has been accomplished using sample data from the Archaeology Data Service, making use of available geospatial and (geo) semantic tools, both FOSS and commercial.
**Linked Open Pottery**

**Gruber Ethan 1, Smith Tyler Jo 2**

1: American Numismatic Society (ANS)
75 Varick Street, floor 11 New York, NY 10013
http://numismatics.org/

2: University of Virginia
P.O. Box 400182, Monroe Hall, Charlottesville, VA 22904-4182
http://www.virginia.edu

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 author has endeavored to apply these methodologies to the field of Greek pottery. While several thesauri have been published online, not all employ linked data, and none of them currently interact as «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 the Getty vocabularies (soon to be published as linked open data), British Museum, and other thesauri, supplemented by URIs from the Pleiades Gazetteer of Ancient Places (http://pleiades.stoa.org) for production places, Virtual International Authority File (http://viaf.org) identifiers for people, and DBpedia URIs whenever applicable.

The paper will discuss the challenges in data normalization between vocabularies, the creation of an ontology in terms that pottery specialists understand, employing SKOS, OWL, Dublin Core, etc. when possible, and the development of an architecture for maintaining, editing, and publishing linked open data. This architecture is based primarily on Nomisma.org, utilizing XForms for editing linked data in web forms, allowing for sophisticated validation scenarios, SPARQL/Update, and other types of REST interactions.

Only by building a bridge between existing thesauri can large-scale aggregation of pottery data be possible. The paper will conclude with several test-case scenarios, showing quantitative and geographic visualizations made possible by the normalization discussed above: the statistical distribution of pottery shapes by vase painters or production places or the distribution of the vessels created by specific potters over time and space, rendered in the form of maps and timelines.

**Uncertainty handling for ancient coinage**

**Tolle Karsten 1, Wigg-Wolf David 2**

1: Databases and Information Systems (DBIS) Goethe University - Frankfurt am Main http://www.dbis.cs.uni-frankfurt.de/

2: Römisch-Germanische Kommission (RGK) Deutsches Archäologisches Institut - Frankfurt am Main http://www.dainst.org/de/department/rgk

*: Corresponding author

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 artifacts, then this context information helps to infer the dating of the latter.

However, due to corrosion and decomposition coins may be bad preserved and even experts can sometimes hardly identify the exact coin type of a given coin. In our database Antike Fundmünzen Europa (AFE), where we record finds of ancient coins within Germany, we want to preserve as much information as possible. This also includes containments of possible coin types, or marking attributes of the coin as uncertain if the exact value can not be assured. Together with other context information and reasoning techniques one then can narrow existing possibilities.

As long as data remains 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 want 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). Following the idea of linked open data, this means that ontologies are needed. One of the cornerstones we
envisage is the use of Nomisma.org IDs in order to provide a common reference for the identification of entities and concepts. Uncertainty is handled only to a very limited extent in Nomisma.org, which makes sense since uncertainty is not something specific for coinage and should be modeled on an upper level. On a very generic level there is SKOS. SKOS provides un-sharp relations and therefore also deals with uncertainty. However, SKOS was build for the mapping between different data models. It does not help us to model the uncertainty within one model. 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. Currently we are exploring the Uncertainty Ontology of the W3C and how it could be adopted for AFE. Within the paper we will provide a brief overview of existing solutions for modelling uncertainty and incomplete data. However, the focus of the paper consists of our use case descriptions, how we plan to model them and our lessons learned.

Some Issues on LOD in Cultural Heritage: the Case of Historical Place Names

Signore Oreste, Fresta Giuseppe, Martinelli Massimo
CNR-ISTI
via G. Moruzzi, 1 ? 56124 Pisa
http://www.isti.cnr.it/

The context is essential to fully understand the meaning of any artifact. This implies a highly interdisciplinary approach where the scholar uses his/her knowledge to link different information items. However, the scholars’ knowledge remains implicit and often unexpressed. The huge amount of information available both on the web and in the archives makes mandatory to link such dispersed information items to get new knowledge, without imposing a single rigid schema. The Semantic Web technologies and the Linked Open Data (LOD) approach promise to offer a way out.

In the presented case study the issue of historical place names has been considered. Starting from the experiences that lead to TGN, an ontological representation of the pilot study data has been drawn. However, to really conform to the LOD principles, a further refinement lead to a more complex ontology, where the role of Events has been introduced. Finally, to really share the knowledge on a world wide scale, a «reconciliation» with the CIDOC CRM ontology has been performed, enriching the reference ontology with appropriate properties and classes. The paper shows as Semantic Web technologies support a LOD approach, and why reusing well established and shared ontologies, possibly adding some extensions, is mandatory to achieve the goal of sharing and enhancing knowledge.

Periods, Organized (PeriodO): a Linked Data gazetteer to bridge the gap between concept and usage in archaeological periodization

Rabinowitz Adam 1, Shaw Ryan 2, Kansa Eric 3, 4
1 : The University of Texas at Austin (UT Austin)
Department of Classics 2210 Speedway, Stop C3400 Austin, Texas 78712-1738
www.utexas.edu
2 : The University of North Carolina at Chapel Hill (UNC)
http://unc.edu/
3 : Open Context
http://opencontext.org/
4 : The University of California at Berkeley (UC Berkeley)
http://www.berkeley.edu

As the web of data built around shared concepts of place and person continues to expand, the need for a resource to anchor shared concepts of time becomes ever more critical. While there are generally accepted standards for the expression of calendar dates, so far the period concepts used to describe archaeological data have proven resistant to integration in a Linked Data environment. A proliferation of heterogeneous, collection-specific thesauri have made the interoperability of period designations equally difficult for datasets using the CIDOC-CRM, despite the
thorough specification of period entities in the ontology. The publication of various period thesauri as Linked Data, as envisioned by the SENESCHAL project and planned by the Getty Art and Architecture Thesaurus, is a productive development. There are still serious barriers, however, to using these thesauri to foster chronological interoperability. The expression of spatial and temporal coordinates in a quantitative, rather than a qualitative, form is necessary for the reconciliation of period concepts in different data sources. But existing thesauri are either too geographically specific for general use (e.g. the English Heritage thesaurus, which applies only to the UK) or too qualitative and vague to provide resolvable spatio-temporal coordinates (e.g. the Getty AAT). We argue that the barriers to the use of periods in interoperable systems stem from tensions between the way periods are conceptualized as unitary temporal entities in thesauri and the idiosyncratic, unstructured, and contextual way they are used in practice. A solution to the problem lies in the creation of a gazetteer not of agreed-upon, centralized, and abstracted period concepts, but of the individual assertions made by authoritative scholarly sources about the temporal and spatial boundaries of periods in particular contexts. This approach promotes scholarly agency in creating and using period concepts, while still providing the formalism needed for computer-aided reasoning and inference on period assertions. The Periods, Organized (PeriodO) project seeks to develop such a gazetteer. Rather than attempting to impose consensus about the nature of an underlying period concept -- the single Platonic form of the «Archaic period», for example -- PeriodO will formally model existing period assertions with spatio-temporal coordinates from authoritative sources, both online and in print. These assertions -- the multiple Platonic shadows of the «Archaic period» -- will form the core of an expanding gazetteer that will facilitate the reconciliation, visualization, and exploration of diverse uses of period expressions across linked datasets and texts.

This paper describes plans for the creation of the PeriodO gazetteer, including the formats it will employ to represent spatio-temporal coordinates and the Linked Data standards to which it will adhere. It discusses visualization strategies and the relevance of emerging tools and standards such as the Topotime specification under development by Elijah Meeks and Karl Grossner or the recent addition of spatiotemporal concepts to the CIDOC-CRM. Finally, it offers a consideration of the potential role of the PeriodO gazetteer in research, teaching, and Linked Data, as well as an assessment of the obstacles to its implementation.

A metadata schema for cultural heritage documentation data retrieval through publication- Using STARC metadata schema to handle 3D Cultural Heritage Documentation (The case of recording sites in Israel)

Kyriaki Yiakoupi 1, Hermon Sorin 2*
1 : The Cyprus Institute (STARC)
http://cvi.ac.cy/starc.html
2 : The Cyprus Institute (CyI)
www.cyi.ac.cy
* : Corresponding author

This paper will present the Science and Technology in Archaeology Research Centre (STARC) metadata schema, explaining step by step the data capture technologies and procedure of a case in Israel. In addition, there will be a presentation of the publication and data interoperability of the content online, related to the STARC metadata schema. A number of European significance tangible and intangible heritage such as museum objects, buildings, archaeological sites and excavation areas are already located in virtual libraries and digital collections. What is more, the existing material is consisted of 3D surveys, reconstruction techniques, images, texts, videos and documentation of various steps. Such a scientific content necessitated the establishment of virtual museums, digital libraries and scientific repositories. Today, the creation of digital collections is well understood by the majority of the scientific and education community. Consequently, this gave rise to the creation of different separable repositories in which each provider could manage and organise their digital content separately, trying to compromise it to its content standards. Nonetheless, in order to be able to retrieve information from all these repositories it is important to have a well-structured and standardised metadata schema. Further to this paper there will be
a detailed examination on how an adhesive schema can act as the backbone structure for management, preservation and the archival of the digital assets.

The topic of this research is related to knowledge management and communication through new ways of reasoning with information technologies and metadata. Long term preservation and access to 3D digital content required the utmost of our concern to the development of standardize metadata and ontologies. Thus, an appropriate and careful structured classification of entities can ensure this.

Finally, the paper will discuss the features and dynamics of the STARC schema through a case of recording sites in Israel, as well as explain how this standardization of entities can give a full record of the digital acquisition tasks.

An Ontology for 3D Visualization in Cultural Heritage

Vitale Valeria
King's College London, Department of Digital Humanities

The use of 3D computer graphics and modelling techniques in the study of the ancient world has been mainly limited to the display of traditional research. Often, their value has been assessed merely on aesthetic quality.

Behind every scholarly 3D visualisation is a thorough study of excavation records, iconographic documentation, literary sources, artistic canons. However, this research is not always detectable in the final outcome, and 3D visualisations do not seem able to meet the standards of scientific method (reproducibility) and academic publishing (references and peer-review).

The London Charter presents methodological guidelines for recording data, but it does not offer a formal framework in which to place this information. Each modeller is left to simply follow their own style. Time and resource constrains not only affect the accuracy and availability of the documentation, but also make it very unlikely that a researcher, or even a team, develop more than one visualisation of the same cultural heritage place/object, perpetrating the naive idea that only one visualisation is possible or correct. The growing compatibility between 3D content and web browsers allows the application of RDF technology to 3D visualisations for cultural heritage to try overcoming these issues. Dividing a 3D model into smaller elements, assigning a Unique Resource Identifier (URI) to each of them and applying a dedicated ontology would make possible to connect the 3D elements with each other and with external resources. More specifically, an ontology for 3D visualisation in cultural heritage could, in the first place, define and describe the components of the 3D model and their relationships. This would help rebuilding data and metadata if the visual component was not readable anymore, enhancing accessibility, sustainability and longevity of the information. Through a dedicated ontology, a researcher could also assess the degree of speculation involved in the creation of each 3D element and its relationship with sources and referents, thus presenting 3D visualisation as a scientific hypothesis and not an «exact reconstruction». Moreover, the use of a specific ontology would: constrain and standardise the documentation, making it synthetic instead of verbose speed up the recording process thus reducing time/cost and making the documentation more likely to be retained in a project, allow 3D visualisations to join and enrich the growing network of linked digital resources to study the past, making 3D visualisations human- and machine-searchable, connecting them with the literary and historical sources that mention the visualised artefact or building, allow and encourage comparison of different visualisations and interpretations of cultural heritage, as the same resource (historical, archaeological, literary etc...) will be connected to all the related visualisations that share the same vocabulary, allow citations, re-use and peer-review of 3D visualisations, as every 3D element (and its author) will be always identifiable and linkable through the URI, contribute to transform 3D visualisation from a univocal display of traditional research to a collaborative virtual environment that can be shared and implemented by different scholars.
New Reflectance Transformation Imaging (RTI) tools for context metadata generation

Schroer Carla 1, Mudge Mark 1*
1: Cultural Heritage Imaging (CHI)
2325 3rd Street, Suite 323 San Francisco, CA 94017
http://culturalheritageimaging.org
* : Corresponding author

This talk will explore the necessity for transparent evaluation of digital tools and methods used for the scientific documentary imaging of «real world» subjects. The goal is to produce a digital stand-in (digital surrogate) that can be used for subsequent scientific or scholarly examinations. The presentation will show how a «digital lab notebook» enables this transparency using the example of the methodology and software for Reflectance Transformation Imaging (RTI).

A «digital lab notebook» associated with a digital representation provides transparency, enabling people to assess its reliability and have confidence they can rely on it for their own research purposes. For RTI, methodology developed at Cultural Heritage Imaging includes a workflow and software tools that keeps track of the information about the capture and processing of the images through to the final digital representation. Our team is developing new methodology and tools by adding management of the process history data using the CIDOC CRMdig ontology and producing a digital lab notebook consisting of Research Description Framework (RDF) files. These files enable the discovery of the relationships among data and lead to the development of knowledge. This presentation will discuss a new direction for RTI practice including the tools under construction. New capture and validation tools for RTI image data will demonstrate this new approach.

The full pipeline of tools for the creation and processing of RTI data contains several parts, and users have many options. The talk will show what is built, what is planned, and what remains, in order to have a complete pipeline that includes managed process history data (context metadata) that meets the CRMdig and Linked Open Data requirements.

The talk will place this work in the context of key factors that determine whether an imaging technology can practically build a digital lab notebook and the implications for the scientific and transparency needs of the cultural heritage community.

Linking potter, pots and places: a LOD approach to samian ware

Thiery Florian 1*
1 : University of Applied Sciences Mainz
www.fh-mainz.de
* : Corresponding author

The Roman imperial period is known for the introduction of a red glossy pottery, the so called samian ware. It is crucial for archaeologists to date Roman sites in the Northern provinces. The research in samian ware pottery advanced in recent years. The spatial distribution and chronological pattern as well as the identification of kilnsites of individual potters are key research questions. Most recent studies are based on a dataset collected by several scholars in the past decades [1]. It comprises detailed and reviewed data of potters, pots and related places. The character of the data makes it reasonable to approach it utilizing the concept of the Linked Open Data (LOD) in order to increase its potential for research.

The goal of the study presented in this poster is to migrate a subset of these data into an appropriate structure and to make it accessible through URIs. Furthermore the place resources should be linked to existing LOD projects. Finally the query of the links between potters, pots and places was to render possible through a web based interface.

The interoperability and possibility for data sharing is achieved by implementing suitable XML schemes (e.g. GML, MIDAS). Findspots and kilnsites are mapped to Pleiades places to benefit from the Pelagios project [2]. Typically Pleiades places are designed to support geographical annotation of textual sources or archaeological objects. A major challenge of the study was to link the ‘samian places’ to Pleiades places, because each follow their own spatial concept. Another aspect was the chronological information entailed in the dataset. Simulations of changes
in the relative chronology of the material revealed the potential of the linked data approach by providing immediate access to all implications on related datasets. The results yield that research on samian ware can benefit from implementing the concept of LOD. Enhanced perspectives arise when extending the linked data to other resources like dies and coins. When the data is published it could contribute to the international crosslinking of archaeological information. Moreover, the data is ready for semantically modelling the relations in an ontology to foster the disclosure of knowledge hidden in the data. The poster presents the results of a M.Sc. Thesis delivered at the University of Applied Sciences Mainz.


A Living Archive for Çatalhöyük

Grossner Karl 1*, Hodder Ian 2, Meeks Elijah 3, Engel Claudia 3

1: Stanford University
Stanford 560 Escondido Mall, Stanford, CA 94305. (650) 463-6137
http://digitalhumanities.stanford.edu/
2: Stanford University
archaeology.stanford.edu
3: Stanford University
* Corresponding author

Çatalhöyük is the site of a long-running archaeological project excavating and analyzing Neolithic period settlement ruins in central Turkey. Its large international team of researchers has been led since 1993 by Stanford’s Ian Hodder. In June 2012, Çatalhöyük was designated a World Heritage Site by UNESCO, in recognition of its global importance. When ongoing excavation activity ceases in a few years, the Çatalhöyük project will need to archive its large volume of digital products in one or more repositories for permanent storage and public access. Throughout Hodder’s tenure, Çatalhöyük data has been made available publicly under a Creative Commons license. Efforts are now under way to make data better organized and more easily accessible, such that it may be reinterpreted by future scholars and joined with data from other projects for comparative studies. This paper reports on progress to date towards realization of an envisioned «living archive,» in a pilot project at Stanford directed by Hodder. A living archive in this sense is one having its core contents «frozen» upon completion of excavation activities, but made highly accessible for future analysis and interpretation, and annotation by its publication as Linked Open Data (LOD), presented in a sophisticated interface allowing complex queries and faceted browsing on spatial, temporal and thematic dimensions. In this way, both core and interpretive data layers of the Çatalhöyük project would live on for future studies and perhaps the interpretive exhibits of digital museums. The tabular data of Çatalhöyük has been developed over a 22-year period with a single-context recording methodology, and stored centrally in a relational database system. Excavation and analysis are undertaken by researchers in 23 specialisms, whose data is stored in a dozen sets of team-specific tables. Team data all refer to object finds and samples from excavation units, a term corresponding fairly closely with the more commonly used contexts. Units are meaningfully grouped in spaces (within and outside of buildings), features, and clusters. The first step in the Çatalhöyük Living Archive project has been the transformation of a significant portion of the data from a relational model to experimental graph representations. Foremost is a data model and instance data described by an RDFS/OWL ontology, within a Sesame triple store. The graph structure facilitates the «reconstitution» of excavation contexts including units, spaces, features, building, clusters, and areas, permitting researchers to more easily identify patterns of co-location?for example, of neonatal burials and hearths, or of zoomorphic figurines and animal bones. The graph structure also enables locating objects according to complex attribute criteria, grouping them, and asserting alternate interpretive categories, such as adornments, tools, or ritual objects. Some other key elements of the Çatalhöyük Living Archive project include: 1) flexible spatial

88
and temporal views on the data; 2) an application programming interface (API) facilitating integration in comparative studies; 3) textual analysis of diary entries and research reports, and 4) an interactive interpretive application demonstrating empirical support for increasing levels of entanglement? the dependence and dependency relations between people and things (Hodder 2012; 2011).