

## **S21 Linked pasts: Connecting islands of content**

*Leif Isaksen, Keith May*

While ever more archaeological and historical content is available online, direct connectivity between independent resources remains comparatively rare. Semantic Web and Linked Data approaches are just some of the possible mechanisms which can facilitate interconnections and this session will be dedicated to presenting concrete examples of any activities which promote cross-navigations, discovery and integration of heterogeneous content. Topics for papers may include, but are not restricted to:

- Interface development and user support for ingestion, annotation and consumption
- Management, publication and sustainability of Linked Data resources
- Building cross and inter-domain Linked Data communities
- Processes for establishing usage conventions of specific terms, vocabularies and ontologies
- Alignment processes for overlapping vocabularies
- Engage non-technical users with adopting semantic technologies
- Licensing and acknowledgment in distributed systems (especially those across multiple legal jurisdictions)
- Incorporation within other software paradigms: TEI, GIS, plain text, imaging software, VR, etc.
- Access implications of integrating open and private content
- Mapping the Field—what components are now properly in place? What remains to be done?

Papers should try to provide evidence of proposed approaches in use across multiple systems wherever possible. Purely theoretical papers and those dealing solely with a single data system are explicitly out of scope for this session. Papers which address both social and technical issues, or bridge between archaeology and other disciplines are especially welcome.

### **S21-01 An ontology for a numismatic island with bridges to others**

*Karsten Tolle, David Wigg-Wolf, Ethan Gruber*

Nomisma.org is a collaborative project to provide stable digital representations of numismatic concepts according to the principles of Linked Open Data. These take the form of http URIs that also provide access to machine-readable information about those concepts, along with links to other resources. We have also constructed an ontology for representing concepts in our thesaurus, and this has been applied to digital representations of physical specimens, enabling linking between specimens in Nomisma-defined numismatic concepts. In our presentation we will describe the processes by which we designed this Ontology with a view to allowing the highest possible flexibility and therefore reducing the barriers to using it. It must be stressed that designing the Ontology was a long-lasting process and is still ongoing. It was often challenging to combine existing requirements and to solve misunderstandings between different parties. We will further present how the Nomisma.org Ontology is used from three different viewpoints. For each viewpoint we will also demonstrate how the numismatic data are already linked to other data sources, thesauri, gazetteers or systems, such as Zenon (<http://zenon.dainst.org/>), Geonames, or others. The goal here is to enable and show how this can be used to combine different archaeological areas with our numismatic data. The back-end system of Nomisma.org: We provide information how we handle data, imports and maintenance issues. Providers of numismatic datasets: Nomisma.org provides the ability to others to publish their RDF datasets (based on the Ontology and with additional modelling requirements) via the Nomisma.org site. For the maintenance of datasets we use Vocabulary of Interlinked Datasets (VoID). We will furthering addition present how the Ontology is currently used by actors outside Nomisma.org (Online Coins of the Roman Empire, Antike Fundmuenzen Europa, Portable Antiquities Scheme, and others) in order to connect numismatic data between different sources.

## **S21-02 A Linked (Open) Data hub at the Norwegian Directorate for Cultural Heritage: A case study**

*Carl Anders Olsson*

The Norwegian Directorate for Cultural Heritage has since 2014 been working on establishing a centralized Linked (Open) Data hub for its own heritage related digital information. This RDF-hub, which contains content from seven separate databases, has a web based search interface on its top. There is also an associated SPARQL-endpoint which offers both the public and third party developers' access to the open part of this data. The technical infrastructure is built using a standard RDF approach using predominantly open source tools. This proposed presentation aims at giving an overview of the key components of this new infrastructure from both a technical and content wise point of view. It will also address issues concerning the further development of the data hub. Key concepts are issues relating to:

- The benefits or drawbacks of mapping parts of this data to other heritage vocabularies such as EDM, CIDOC CRM etc.
- Challenges with the ambition to share as much of the data as possible with both other governmental agencies and with the public
- New or other uses of the aggregated data in relation to reporting, planning or research
- Potential third party interest and use of the openly shared material in mobile or web based dissemination or in other revenue generating activities

## **S21-03 Linking periods: Modeling and utilizing spatio-temporal concepts in the chronOntology project**

*Sebastian Cuy, Wolfgang Schmidle, Florian Thiery*

In the last years several web services emerged that manage and make accessible place thesauri for the archaeologies and historical sciences. By making use of semantic technologies these applications are able to act as linked data hubs thereby making possible the linkage of multiple datasets of varying thematic focus and of different structural properties. Another common denominator of archaeological data resources, besides geo-spatial properties, is the temporal classification of research objects. One of the applications that try to assume a role similar to that of gazetteers but for temporal concepts and cultural periods is developed in the chronOntology project.

In this project, funded by the German Research Foundation, the German Archaeological Institute (DAI) together with the i3mainz are developing a system for storing, managing, mapping and making accessible descriptions of temporal concepts. The core of this endeavor is a rich semantic modeling of various existing terminological systems for cultural periods using a data model based on the CIDOC- CRM and its extensions. The use of the rich ontological model provided by the CRM on the one hand permits representing the measurable temporal extent (with the possibility for fuzzy edges) while also making it possible to embed temporal concepts in a network of semantic relationships to other temporal concepts, connected historical regions and thematic contexts.

Besides documenting the general architecture and data model of the project the paper will present possibilities of querying different heterogeneous data resources collected throughout various digitization and digging activities inside the DAI with the help of concepts defined in chronOntology. We will also point out the potential and problems of reasoning over geographically and temporally connected datasets.

### **S21-04 The Matrix: Connecting time and space with archaeological research questions involving spatio-temporal phenomena and the conceptual relationships between them**

*Keith May*

The most common type of record in archaeological recording systems is the spatial dimension. When recording new layers, buildings, or any physical objects, we measure height, width, depth and for archaeological features we will describe shape in plan and section as well as attributes like profile, diameter and breaks of slope. Recording of temporal information about similar features is far less prevalent, but is still an important (perhaps more crucial) part of the record, particularly for objects where the dates of coins, brooches, or pottery and other 'finds' objects (with relative chronologies) are used for temporal reasoning or inferences about deposition dates and sequences across archaeological stratigraphy. Having divided the archaeology into various units for recording purposes, we use stratigraphy, and associated temporal logical relationships between the physical materials recorded as the 'reasoning glue', in the form of Phases and Periods, to connect all these different spatial and temporal phenomena back together again with various narratives to explain our conclusions. For 'single context recording' most archaeological temporal reasoning is based on the principles of stratigraphic superposition, the "Above and Below relationship" (Harris). But further principles of temporal reasoning are also available (Allen). The CIDOC CRM uses the Allen operators to describe not just superposition but a set of more complex temporal logical relationships that can pertain between archaeological data. This paper will give an insight into how conceptual reference modeling can be used to explore these issues and how associated semantic technologies can enable semantically enriched deductions about the spatio-temporal relationships which fundamentally link such archaeological data together. It will also consider where further work is needed to deal with not just spatial or temporal records but to reason about wholly spatio-temporal phenomena and how this can form the basis for new linkages between archaeological information across space-time.

[References]

Allen, James F. Maintaining knowledge about temporal intervals. In: Communications of the ACM. 26 November 1983. ACM Press. pp. 832–843, ISSN 0001-0782

Harris, E. 1979 Principles of Archaeological Stratigraphy. London & New York: Academic Press. ISBN 0-12-326651-3

### **S21-05 When data meets the enterprise. How we turned a merger of organisations into a confluence of information**

*Koen Van Daele, Maarten Vermeyen, Sophie Mortier, Leen Meganck*

In 2012 Flanders Heritage Agency was created as a central agency dealing with immovable cultural heritage - broadly defined as archaeology, built heritage and cultural landscapes - in Flanders. Prior to this, tasks of this agency were carried out by several independent agencies. The merger created a very heterogenous set of business processes, IT-components and systems. This, together with a new heritage legislation, prompted a re-evaluation of these systems and their business processes.

This paper will delve into our system architecture, built on a core separation of concerns between data driven applications and proces driven applications. We will explain how we came to implement this in a service oriented architecture . We will detail how and why we chose to go with REST services instead of SOAP services.

The resource oriented focus of REST services has served us well in creating inter-linking data sources that are firmly grounded in the World Wide Web and the HTTP protocol. We will demonstrate how we link these resources by using cool URI's. While a majority of our links are between the resources we create and maintain ourselves, we will also look at how we

interact with external resources and services when it comes to specific domains such as vocabularies and GIS.

Finally we will look at how we are further enhancing our data by more formally publishing it through the use of semantic technologies such as RDF. We aim to create truly linked open data in this way. We will look at some of the stumbling blocks we have encountered along the way. The most significant one to date being the clash between open data and privacy regulations and how to implement access control on linked data.

### **S21-06 Where is the House of the Dwarves? Enhancing granularity in the Pleiades Gazetteer: The examples of ancient Sicily and Pompeii**

*Valeria Vitale, Jeffrey Becker, Jonathan Prag*

Pleiades, the online gazetteer of ancient places, has, for several years, been a precious resource for classicists and historians. Assigning a stable URI to each ancient place has allowed a number of digital projects to build on top of this infrastructure, developing new tools and resources such as Pelagios.

We believe that Pleiades could stimulate and facilitate new and interesting applications by assigning URIs to geographical units smaller than cities. The aim of this paper is to support this idea discussing two examples:

The digital epigraphic project iSicily. This project involves the identification, locating and addition to Pleiades of various Sicilian contrade (an administrative unit that was common in rural southern Italy) and other sublocations that have been recorded in archaeological reports and previous bibliography as findspots of antiquities, or positions of ancient monuments (some of which no longer exist). The availability of these URIs allows a strong and informative synergy between academic research on those antiquities and museum metadata, expressed in linked data, showing relationships and suggesting potential patterns and future lines of enquiry.

The city of Pompeii. This project involves minting specific URIs for each Pompeian building. The Campanian city offers a unique case study due to the amount of information, bibliography, and often confusion, which orbits around many of the single buildings.

Assigning a URI to each of them will help grouping and disambiguating the names and the interpretations (sometimes dramatically different) assigned to the same building during the last 250 years. Secondly, it will facilitate dialogue between several existing and future digital projects about Pompeian buildings. Lastly, it will link the information about Pompeian artefacts stored in databases, digital repositories or museum archives with the exact building where the artefact was found, and not exclusively with the generic provenance "Pompeii", thus offering an immediate basic level of contextualisation and highlighting connections with other artefacts related to the same building.

### **S21-07 LOD for Numismatic LAM Integration**

*Ethan Gruber*

The American Numismatic Society (ANS), founded in 1858, is a research institute focusing coins from all eras and regions. It owns one of the largest collections of coins in the world, one of the largest numismatic libraries, is a publisher of monographs and journals, and maintains an archive of research notes from scholars associated with the Society. The ANS has been involved in the publication of numismatic databases and dissemination of such materials following Linked Open Data (LOD) methodologies since 2011; aspects of these digital projects (from Nomisma.org to Online Coins of the Roman Empire, <http://numismatics.org/ocre/>) have been detailed at previous CAA conferences.

While these other projects have focused on implementing LOD techniques in the publication of coin hoard or typological databases, this paper focuses on applying open standards from

across the Library, Archive, and Museum domains to thoroughly integrate the ANS's numismatic collection, library, archive, scholarly publications, and typological and hoard databases. We have begun a new project to digitize nearly 100 monographs into TEI, inserting links to people or places defined on Nomisma or the Pleiades Gazetteer of Ancient Places, citations to books or archival materials held by the ANS, coins in our or other museum collections, and references to hoards or coin types published online. These digital monographs, in essence, become research gateways into similar topics in the larger ancient world linked data cloud. Furthermore, these TEI documents may be deconstructed into RDF. Passages about the Macedonian city of Amphipolis may be made available to researchers through the Pelagios Project. Similarly, a user viewing a particular coin in our collection database may be read a paragraph about the coin, extracted from a TEI document. Our ultimate goal is to create an improved research experience for our users, allowing them to traverse seamlessly from one service to another, whether they begin their search within the ANS project network or arrive from external sources, like Pelagios.

### **S21-08 Pelagios Commons: Decentralizing the Web of historical data**

*Leif Isaksen, Simon Rainer, Pau de Soto Cañamares, Elton T. E. Barker*

Pelagios is an international initiative concerned with the development of Linked Open Data (LOD) methods, tools and services so as to better interconnect the vast and ever-growing range of historical resources online. In particular it associates place references within those resources to online gazetteers that offer URI-based identifiers for such places. Some of its major outputs have been the development of Recogito, a tool for semantically annotating place references in images and texts, and Peripleo, a service for visualizing and exploring the graph of data that these annotations form.

In parallel with these developments a community of practitioners has started to form with interests in a range of related activities: the annotation of curated or third-party content; the production of specialist gazetteers; the integration of place annotations with those of people, periods and things; and the visualization and analysis of graph-based data, to name but a few. Since its early stages Pelagios has made concerted efforts to consult and support such stakeholders, but as it has grown new opportunities and challenges have emerged. In particular we have established that within a heritage context, LOD's principal advantage is its ability to relate independently maintained projects without requiring centralization. But what are the social ramifications of such an approach? In a world in which funding, academic legitimacy, intellectual property, and even conference presentations assume the authority of individuals and institutions, can LOD communities ever scale effectively?

This paper reports on early developments within Pelagios Commons, a new phase of Pelagios which focuses explicitly on addressing technical and social decentralization within Web-based projects of this nature. It will present our experiences in establishing Special Interest Groups, and the different challenges faced in devolving LOD architectures. It will also seek to foster discussion and critique from those planning or implementing similar community-driven projects.