

Meeting Programme

Friday 30th November

Time	Speaker	Affiliation	Title
13:00	Welcome Address		
13:15	Dr. Karsten Tolle & Dr. David Wigg-Wolf	University of Frankfurt / Römisch-Germanische Kommission	Data quality at database and higher levels - Our work with the numismatic database AFE
13:45	Vladimir Stissi & Jitte Waagen	University of Amsterdam	Fighting aging... of data. The example of (Greek) pottery databases
14:15	Victor Klinkenberg	Leiden University	A cautious step into 3D spatial analysis
14:45	TEA / COFFEE BREAK		
15:15	Stan Roosen & Steven Soetens	Free University Amsterdam	A 3D model of the medieval urban subsurface of Vlaardingen (Zuid Holland)
15:45	Ferry van den Oever	Saricon	Archaeogeophysics, 'How deep can you beep?' and isn't 2.5D good enough?
16:15	Serge van Gessel	TNO Netherlands Organisation for Applied Research, Geoinformatics division	An expert in 3D subsurface mapping, Serge will be the discussant for the 'Z' theme
16:45	TEA / COFFEE BREAK		
17:00	CAA NL/FL AGM		General meeting of CAA-NL/FL Chapter. Election of officers and voting on the Constitution
18:00	END OF DAY 1		

Saturday 1st December

Time	Speaker	Affiliation	Title
09:00	Tobias Kohr	i3mainz / Fachhochschule Mainz	Distributed Geodatabases in Archaeological Joint Research Projects
09:30	Milco Wansleebe	Leiden University	Linked data: provenance metadata becomes 'ordinary' data
10:00	Loup Bernard	Strasbourg University	ArkeoGIS, merging French and German archaeological and environmental databases
10:30	TEA / COFFEE BREAK		
11:00	Georg Hohmann	Germanisches Nationalmuseum, Nürnberg	Ontology-based documentation of cultural heritage: the Semantic Research Environment "WissKI"
11:30	Dr. Felix Schaefer	Deutsches Archäologisches Institut, Berlin	The new research data centre IANUS - approaches to more data quality
12:00	Dr. Kim Cohen	Utrecht University	GIS reconstruction of the palaeogeography of the Holocene Rhine-Meuse delta
12:30	LUNCH BREAK		
14:00	Chris van der Meijden	LMU Ludwig Maximilian University, München	Ossobook - Spicing archaeo related sciences with archaeo-informatics
14:30	Dr. Guus Lange	Rijksdienst voor het Cultureel Erfgoed RCE, Amersfoort	Knowing knowledge
15:00	TEA / COFFEE BREAK		
15:30	Dr. Matthias Lang	University of Göttingen	ArchGate - an integrated Database-GIS-solution for archaeological fieldwork
16:00	Dr. Martijn van Leusen	University of Groningen	Handling Uncertainties in Legacy Site Data: the strange case of the Hidden Landscapes database
16:30	Closing discussion		
17:00	END OF CONFERENCE		

Abstracts, Friday 30 November 2012

13.15 Karsten Tolle / David Wigg-Wolf - Data quality at database and higher levels – Our work with the numismatic database AFE

Antike Fundmünzen in Europa (AFE) is an existing database of finds of ancient coins, predominantly from Germany, hosted by the Römisch-Germanische Kommission (RGK) with technical support provided by the Database and Information Systems (DBIS) group of the University of Frankfurt am Main. During the last year we have re-engineered the original Access database, and transformed it into an online MySQL based database. We are currently working on issues such as uncertainty: how do we model it in order to preserve existing data and to avoid ambiguity? In the long run 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). 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. The talk will present AFE and the lessons we have learned.

13.45 Vladimir Stissi and Jitte Waagen - Fighting aging... of data. The example of (Greek) pottery databases

Greek pottery, both in museum collections and as found in field work projects, has been entered in electronic databases from a relatively early date. The Beazley Archive, which went electronic in 1979 and online in 1998 (<http://www.beazley.ox.ac.uk/pottery/default.htm>), was one of the earliest large artefact databases on the web, and is still intensively used even if a bit outdated in many ways. Museums often do better, although some online datasets have been disappearing from view. Early electronic field work databases have rarely been put online in the first place, and many stored in obsolete formats are now in serious risk of disappearing. But is that really the main problem? Working with electronic pottery datasets produced in the late 1980s and early 1990s, I have noted that the data themselves are often as least as problematic as their transmission, if not more so. One reason for this is that we now know a lot more about many categories of material, but more importantly, supposedly 'traditional' and 'standardized' ways of classifying and analyzing pots have changed considerably over the last 20 years or so. In other words, early databases contain fields and categories we wouldn't use anymore, and lack some we would like to have. Perhaps it is only natural that data become outdated, but this is obviously also a form of aging we would like to avoid. When designing new pottery databases (like the one for the New Perspectives on Ancient Pottery (NPAP) research project developed at the Amsterdam Archaeological Centre in 2008-2011) this poses two problems/challenges: is it possible to make a more durable framework? And how can we save as much existing content as possible from becoming obsolete? Or: is it useful at all to revive old datasets? We have not any clear answers yet, but in this paper will explore some perspectives and some problems.

14.15 Victor Klinkenberg - A cautious step into 3D spatial analysis

Spatial analysis has been an integral part of most archaeological research in the past decades. Recently the use of GIS has been extended with the possibility to visualize three dimensional GIS data. Most formal quantitative methods of spatial analysis and statistics have not altered much by the introduction of 3D into GIS. Other methods, such as thematic mapping, seem to have gained importance with the addition of the third dimension. These methods are of a more informal nature and their inferential potential relies heavily on a human observer of the method's output. Therefore these methods increase in value with an improved graphic validity. Using my research on a Palaeolithic find distribution and a Neolithic cemetery I will argue that the advent of 3D in archaeological analyses is a definite step forward, albeit a step that should be taken cautiously.

15.15 Stan Roosen / Steven Soetens - A 3D model of the medieval urban subsurface of Vlaardingen (Zuid Holland)

The subsurface of the city centre mound of Vlaardingen holds a rich and very well preserved archaeological archive of organic remains. This preservation is puzzling since the site is elevated and located on top of a fossil fluvial area. For that reason, the geoarchaeology of the subsurface was studied by various disciplines. More relevant for the 3D modeling this was done by a series of corings and soundings that include archaeological information, but also lithostratigraphy, TGA and grain size analysis. The present paper demonstrates the translation issues from raw coring data to a 3D subsurface model, by comparing different software packages, by investigating data integrity and the methodological problems encountered. The main issues concern: automatisisation of creating 3D elements from raw data; interpolation techniques to stratigraphical modeled layers; and creation of modeled profiles. This project is the result of collaboration between the VU University Amsterdam, the City of Vlaardingen and the University of Amsterdam.

15.45 Ferry van den Oever - Archaeogeophysics, 'How deep can you beep?' Isn't 2.5D good enough?

Near-surface geophysics, or non-destructive techniques, are slowly but steadily becoming part of the standard toolbox for prospection. These techniques are to be used in addition to other (destructive) techniques. Geophysical datasets and the 3rd dimension: How to achieve a real 3D-model and is there an advantage in volumetric analyses? Who is using real 3D-models? Isn't 2.5D enough? Is there a need for 3D-models in commercial archaeology? Determining which geophysical technique where and how to be used is of course very important. Apart from gathering data the correct way, how do we handle this data? The datasets itself are becoming larger and larger. Is working in the cloud part of the solution? What about archiving (meta-)data? How to go about with data fusion? It's about time to set up Dutch 'best practice' guidelines for geophysics!

16.15 Serge van Gessel – Discussant for the 'Z' session

Abstracts, Saturday 1 December 2012

09.00 Tobias Kohr - Distributed Geodatabases in Archaeological Joint Research Projects

The paper introduces two current research projects conducted at the i3mainz with different archaeological background but common technical solutions. Within the HiGeoMes (Historische Geographie Obermesopotamiens) project conducted by a bi-national group from France and Germany, textual and archaeological data need to be integrated in a spatio-temporal context. As part of the “data curation” initiative of the Geocycles research centre at the University of Mainz the connection of archaeological data with geological information from the Eifel area is established. Both projects focus on the collection of archaeological sites within different regional scale. HiGeoMes is based on existing data sets of the ancient Near East which quality and quantity was enhanced by investigating additional spatial and bibliographic sources. The Eifel-data are initially collected from bibliographic research and administrative documents facilitated by geodata available online. The distributed environment with contributors from several places and different scientific disciplines were a common challenge. By implementing web-clients based on FOSSGIS technology, data acquisition as well as dissemination is obtained by the concept of service oriented architecture (SOA). While most of the projects effort concentrated on the development of the clients, several issues were tackled concerning data quality, 'legacy'-data and spatial as well as temporal accuracy. The paper will exemplify encountered problems and solutions. We will especially argue for using OGC-compliant web services for preserving and describing data quality.

09.30 Milco Wansleebe - Linked data: provenance metadata becomes ‘ordinary’ data

The data structure of the information stored in the Semantic web/Linked Open Data realm is so flexible that it is possible to include any statement about data quality very easily. Explicit information about the spatial accuracy or uncertainty in the typological assignment should preferably be added by the original researchers, but can be added to the web of data by others later, based on peer reviews or re-use experience. Legacy, national or alternative typologies can be mapped to a target typology once, so crosswalks are available to everyone. The context of data acquisition and analysis is directly clear, since it is automatically part of the relations between the events, actors and (data) documents. Currently experiments with Linked Data in Dutch archaeology have been limited, both in the number and in the scope of demonstrators. Technological problems do not seem to be the main reason for this. We lack user friendly information systems that allow archaeologists to interact with a substantial amount of Linked Data. Only then they could explore new ways to answer archaeological research questions. As long as the content is still very limited, the benefits will not be obvious and the additional effort archaeologists have to invest in publishing archaeological and provenance data as Linked Data is not easily made or justified in the current economic situation.

10.00 Loup Bernard - ArkeoGIS, merging French and German archaeological and environmental databases

French and German archaeologists and geographers work together for several years now, to put the 2.0 version of ArkeoGIS online. The webGIS is a free cooperative tool. Data preservation is guaranteed by the TGE ADONIS, who hosts the project and organizes interaction between different social sciences and humanities online programs. Gathering inventories and research databases, coupled with lists of palaeoenvironmental analysis grants us unprecedented access to our neighbor's data in French or German language, and helps our non-bilingual colleagues to apprehend it. ArkeoGIS is also an amazing tool to reveal the matters of soil occupation/anthropic impact and erosional behavior in the Holocene upper Rhine valley. Listing different bases online is the best way to actualize older databases, when put on a dynamic map, literature or interpretational differences between the datasets appear immediately and are easy to update with the export tool.

11.00 Georg Hohmann - Ontology-based documentation of cultural heritage: The Semantic Research Environment "WissKI"

The semantic research environment WissKI enables the easy use of Semantic Web technologies for documentation, research and online cooperation the easy way. It combines the benefits of common content management systems and semantic web frameworks and serves as platform for curating knowledge. A unique feature is the way WissKI deals with data: Each entry is treated as an individual according to the CIDOC Conceptual Reference Model (CRM) and stored in a graph-based storage. So, scientific knowledge is turned into a complex semantic network which can be queried, browsed and visualized. It allows to manage structured information using forms as well as and non-structured information using semi-automated text annotation. The possibility to integrate common authority files allows the automatic creation of selection lists which assist manual data input. Based on the Erlangen CRM, an OWL-DL implementation of the CIDOC CRM, the systems is highly customizable to fit specific requirements and still ensures the compatibility of the information that stored inside the system. WissKI also makes it very easy to exchange data according to the standards of the scientific community. It provides several interfaces and formats like OAI-PMH, SPARQL, LIDO or RDF out of the box and allows the definition of custom import and export formats. By using WissKI a single researcher, an institution or a project can instantly join the Semantic Web and Linked Data communities and benefit from the advantages of this approach.

11.30 Dr. Felix Schaefer - The new research data center IANUS - approaches to more data quality

The paper will give some insights about IANUS, a new research center in Germany for digital data from archaeology and classical studies. This project is funded by the DFG and currently is under construction. Since September 2011 its tasks and duties are being defined and financial as legal frameworks are discussed. Once established, IANUS will be comparable to eDNA in the Netherlands and ADS in the UK. Primary goals will be the long-term preservation, the dissemination and the aggregation of digital data. For all these aspects the issues of data management, data quality, data storage and re-use are crucial questions. Although still in the planning phase, the paper will present some current ideas concerning the outlined aspects of managing data quality. One focus will be the new IT-guidelines which comprise both accepted standards and best-practice examples and which IANUS is going to host and to promote within the German community. Hopefully these will help data producers to improve the data handling and to enhance the consciousness for the data quality within research projects. Another issue will be the quality of deposited data that needs to be archived. Especially the integrity and homogeneity of the documentation and metadata submitted along with files themselves it is one key element in order to make digital information reusable in future times by new, unfamiliar users.

12.00 Dr Kim M Cohen - GIS reconstruction of the palaeogeography of the Holocene Rhine-Meuse delta

Since 1999 Utrecht University maintains a reconstruction of the Holocene channel belt network of the Rhine-Meuse delta, stored in GIS. New channels have formed and older channels have been abandoned through natural sedimentary processes such as avulsion and transgression. The resulting Holocene delta substrate can be seen as a spaghetti of channel belts, composed mainly of sand, that dissect flood basin sequences of clay and peat. These features are mapped at very high resolution, and for the great majority of the channel belts direct age control on their beginning and abandonment is available. Utrecht University keeps a series of databases of primary raw data for this type of research, for example for borehole descriptions and ¹⁴C datings obtained on relevant sediments. The GIS of the Rhine-Meuse delta channel belt fragments is our database storing the mapping and dating of recognized channel belt features and is our tool for iterating to the optimal palaeogeographical reconstruction for any given time during the Holocene. I will discuss the functional setup of the GIS for delta network reconstruction and the philosophy behind it. I will give a few examples of the way the database is presently used in archaeological, geological and physical geographical applications. The reconstruction is a widely used resource in the 'Malta' archaeological practice in the Netherlands today, but the original digital map is now over 10 years old and many details have significantly changed (improved, we think). I will highlight the progress in recent years, including our extensions into the Niederrhein area in Germany.

14.00 Chris van der Meijden - Ossobook - Spicing archaeo related sciences with archaeo informatics

To successfully establish the new scientific branch of archaeo informatics the main problems are based on standardization problems, understanding of advanced informatics (i.e. data mining) within archaeo sciences and setting up data communication infrastructures. Our experiences are based on the development of OSSOBOOK, an intermittently synchronized database system that allows any authorized user to record data on bones offline at the site and later synchronizing this new data with a central data collection. Powerful data mining and similarity search tools have been integrated. The actual development steps are establishing a standardized minimal electronic finding description and the implementation of an enhanced database connection interface for data mining communication techniques to set up an archaeo data network.

14.30 Guus Lange - Knowing knowledge

The Semantic Web or Linked Open Data are sometimes together called Web 3.0 as the successor to the Web 2.0 of the social media. A very successful social medium is of course FaceBook, but Web 2.0 really promised to be the new means for online collaboration. As such, this development has not been hugely successful up to now. One of the reasons might be that in the digital world, people and computer programs use different knowledge schemes while these differences are not immediately obvious. When I call a sherd 'Dressel 20', do you know exactly what I mean? When we do data retrieval we find terms, but not the concepts behind them. In digital communication the contextual information is lacking that usually makes up for this lack, at least partly, in real world communication. When reading a report one can assess the quality and its place in the scientific tradition by glancing through it quite fast. For a computer search program this is simply not possible at all. Other measures should be taken to make sure we retrieve the right information. Therefore The SemWeb approaches these problems by including semantics into the data files themselves so that both computers and people do understand each other by sharing the same concepts without necessarily talking the same 'language'. Thesauri-management is the key application here. The Rijksdienst voor het Cultureel Erfgoed develops an infrastructure for the heritage field and their partners from outside to use to exchange trustful and meaningful information. In my presentation I will discuss this infrastructure and will show the first results.

15.30 Dr Matthias Lang - ArchGate - an integrated Database-GIS-solution for archaeological fieldwork

This paper will discuss the ArchGate-project. ArchGate is an integrated Database-GIS-solution for archaeological fieldwork developed by the institute for archaeology at the University of Göttingen and the department for Geo Computer Science at the Bochum University of applied science. The software integrates all necessary modules for archaeological fieldwork in one common database and makes all the data available in a shared interface. ArchGate is exclusively based on open-source-software and standardized data-formats. Our paper will discuss the possibilities and the problems of using standardized data-formats in archaeological databases. We will focus on the use of standardized SKOS-XML vocabularies.

16.00 Dr. Martijn van Leusen - Handling Uncertainties in Legacy Site Data: the strange case of the Hidden Landscapes database

All landscape archaeology projects attempt to incorporate legacy archaeological data. Until the early 1980s this would have been almost exclusively site-oriented data, in large part produced from older 'paper' records sometimes going back to the later decades of the 19th century; up until the late 1990s, when GPS and GIS penetrated the discipline, field walking surveys often produced exclusively site data. A typical problem for those building regional archaeological data management systems today is therefore to find ways of storing legacy site data of widely varying quality; a further problem for those wanting to use such data for analytical purposes is to assess its quality and to store quality metadata in such a way that they can be usefully included in queries. Uncertainties can, and will, arise about all site parameters as supplied by the source; more-over, we do not necessarily agree today about the character and significance of these uncertainties. This papers presents an outline of the solution adopted by the author's research group at the department of Mediterranean Archaeology of the Groningen Institute of Archaeology, using as an example its MS-Access database of sites compiled for the 'Hidden Landscapes' project 2005-2010.

