Event-based Archaeological Registration Principles

Peter Jensen
Aarhus University, Denmark. *farkpj@hum.au.dk*

Abstract:
Over the course of the last decade, Aarhus University has been developing a GIS data structure combined with a series of procedures for excavation documentation which facilitates the use of a more efficient digital approach to excavation practices. The new procedures are aimed at ensuring the proper relation of data, uniformity in classification and inclusion of interpretations. Moreover, new registration concepts and classes are introduced in order to properly structure and qualify excavation data for the digital platform. For example, the concept of drawings is replaced by Documentation Events and Data Collections, which introduce a historical dimension of data in recording practices, and therefore makes it possible to distinguish between original data, interpretations, and the various combinations thereof as well as preventing the accumulation of redundant and derived data.

Key Words: *Event-based, Data Management, Documentation Principles*

Event-based Archaeological Registration Principles

Over the course of the past decades, considerable advances in software and hardware capabilities for managing archaeological field data have spawned a steady migration from traditional analogue, paper-based recording principles towards an increasingly digital approach. In Danish archaeology the inherent problem remains the fact that very little effort has been made to truly embrace the possibilities of these new, digital media and to change the documentation principles and strategies accordingly. Instead, much of what we see are digital documentation strategies which are largely dominated by routines and concepts that refer to a more traditional paper-based registration and the framework and limitations which are thereby implied. This limits the use of new digital recording techniques, complex data types and advanced analysis capabilities.

It is only when it comes to the digitization of excavation plans that a considerable effort has been made to make use of these new technologies, resulting in an extensive amount of vector data. However, the implementation of digital data management systems for these GIS-data, and archaeological data in general, is still underdeveloped considering its potential.

Aarhus University has been developing a GIS data structure combined with a series of procedures for excavation documentation which facilitates the use of a more efficient digital approach to excavation practices. The new procedures aim at ensuring the proper relation of data, uniformity in classification and inclusion of interpretations. Moreover, new registration concepts and classes are introduced in order to properly structure and qualify excavation data for the digital platform. It is the aim of this paper to present the preliminary considerations and experiences of this methodological and theoretical framework.

Analogue vs. Digital Documentation

By evaluating excavations carried out by Aarhus University over the last ten years, some key issues and problems concerning the management of analogue and digital archaeological data types were identified and different solutions proposed. We found:

- Problems with managing vector and raster GIS data resulting in an excessive amount of GIS layers/tables, which were often organized in an intricate system of files and folders.

- Problems with managing derived data. It was difficult to keep track of the original data.

- Problems with missing metadata. It was difficult to get a grasp of the historical dimension of the data we had. How were data developed and created in subsequent interpretations? How valid were these interpretations, especially when combined with the data from historical excavations as well as that from modern re-excavations?

- Problems with using traditional lists of data (which are conceptually out-dated) while simultaneously aspiring to perform a digital registration, and lacking the proper reinforcement of cross references between different data types.

One of our first objectives was to deal with the concepts used in traditional recording principles which, without any further consideration, had simply been transferred to a digital equivalent. We wanted to address the inability to handle new data types within the boundaries of these traditional concepts. One of the least functional concepts in the digital realm turned out to be the use of 'drawings'.

The traditional end product of a paper-based documentation strategy would almost always be comprised of drawings, supplemented by several lists of archaeological data: features, structures, finds and photos. But suddenly we were generating new types of data that would be lost if reduced to a simple drawing.

Instead of hand-drawing in the field, we were introducing the use of digital photos in combination with measured points from a total station that were rectified using BASP Airphoto (Scollar 1998) and georeferenced in MapInfo or ESRI ArcGIS. At the same time, we began taking two series of photos: an observation series and an interpretation series in which the archaeologist had physically sketched their interpretation into the soil. The field interpretations depicted on the interpretation photos were then later on used for mosaicing and vectorization in the GIS software, resulting in vector drawings similar to a traditional drawing.

This approach facilitates new possibilities of reinterpretation later on, due to the availability of an un-interpreted photo series, combined with an interpreted photos series as well as a derived vector drawing.

Despite its digital aspect, the end product of our efforts remained, in effect, a drawing. Nonetheless, the process produced photo representations of the observed and interpreted structures, which in fact were of much higher quality as a source of research material than the reduced vector drawing. This complex data type, comprised of several individual pieces of different types of data, could hardly be contained within the concept of a 'drawing' any longer. At times the lack of proper management of cross-references, files and metadata led to a somewhat chaotic situation.

We wanted to eliminate the archaic concept of hand drawings and solely depend on different digital approaches to create the end-goal, which was not limited to a vector drawing, but rather was a combination of raster and vector data. We chose to introduce the concept of a Documentation Unit which represents physical areas or boundaries within the excavation of an extent similar to that of a traditional drawing, but comprised of photos, measurements, rectification, georeferencing and digitization elements.

The Documentation Units were, however, a difficult data type to handle. They produced and contained vast amounts of data – much of it redundant copies as well as data which were derived from other data. When dealing with excavations which lasted several seasons it became very clear that combining previous year's data with new data could pose a problem. Some parts had already been documented in the previous seasons, but became reinterpreted as new sections of profile walls and trenches were taken to a slightly different depth. Combining and merging the diverse vector data from different seasons and documentation situations resulted in either another derived dataset or an inappropriate editing of the existing data in which we lost track of the original interpretation and source of the vector drawing. Instead we turned to the dual concepts of Documentation Events and Data Collections.

These concepts would introduce a historical dimension of data in the recording practices and made it possible to distinguish between original observation data, interpretations and various combinations thereof, as well as preventing the vast amounts of redundant and derived data.

Documentation Events and Data Collections

The philosophy behind this documentation principle is that no piece of archaeological data is ever edited, modified or deleted once the Documentation Event has taken place.

Each series of photos is a Documentation Event, just as is every series of measurements, interpretations, descriptions, sampling and digitizing sessions etc.

By using Documentation Events a time-stamping of all data and spatial objects is inherited, and it is therefore possible to distinguish between primary excavation data and derived data. It is even possible to perform temporal queries in order to reconstruct any historical research situation. For example, it would be possible to visualize how the archaeological interpretation appeared on a specific date before other events, such as further excavations and reinterpretations had taken place.

The Documentation Events facilitate more dynamic interpretations and reinterpretations because additions and modifications of data do not compromise the original records, but are added as new Documentation Events – keeping the original data intact. An excellent illustration of this would be a case in which repeated excavations at the same archaeological site lead to reinterpretation and new registrations of the same archaeological structures. The old data do not lose their significance. They represent the original observations that were made by the original excavations, which can not be repeated or redone and, therefore, are still valid. To facilitate the management of all these events, they are all related to a Data Collection.

The Data Collections are made up of, for instance, the individual trenches, parts of the excavation field and the larger profile walls. They can even represent different documentation levels of the same geographic area, relating all the Documentation Events to a structure that is conceptually easier to handle. Furthermore, the Data Collection concept is a way of forcing the archaeologist to plan ahead and consider the goals of the excavation very explicitly, as the Data Collections will not only handle all other data, but also function as a 'gateway' to the later visualization of data. From the outset, the archaeologist must make plans for what the end goal of the documentation will be and what will be important to visualize.

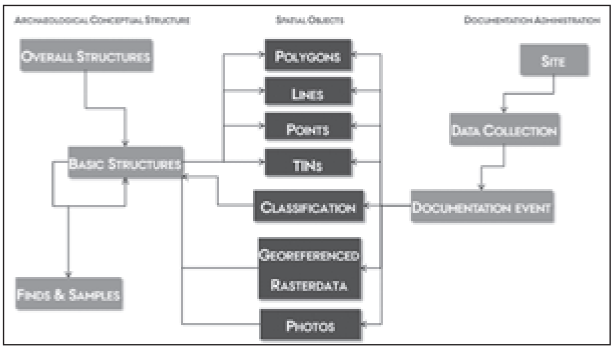

Figure 1. The data structure. The devision into three columns allow three different entry points of data exploration.

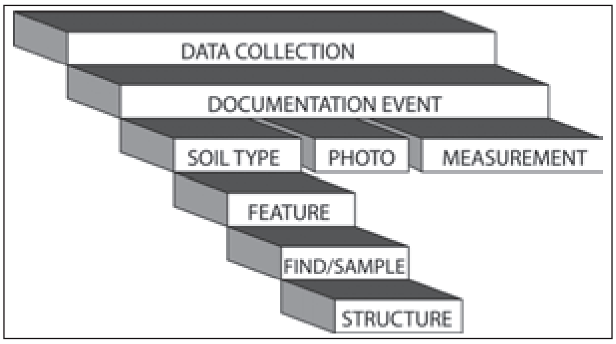

Figure 2. The ladder of documentation.

The usage of event-based documentation principles is not a revolutionizing thought in archaeology. However, it is new to the realm of field archaeology in Denmark. The concepts and theories of temporal databases were introduced to Danish archaeology as early as the 1980s and 90s by the development of the artefact database of the Danish National Museum (GENREG), which is based on exactly the same type of time-stamping and non-destructive update operations on which our documentation principles rely (Eaglestone et al. 1996). The GENREG system accounts for the history of an artefact or object as it is discovered, preserved, described, interpreted, stored, and displayed, much in the same fashion that our documentation principles show the reinterpretations of an archaeological structure or feature over the course of time. Our data structure can be visualized as three individual columns (Fig. 1).

- On the right hand side we have the administrative parts of the event-based documentation: Data Collections and the Documentation Events.

- On the left hand site we see the conceptual archaeological structures: finds, artefacts and samples, the basic structures and the overall structures.

- Between the two columns we have all the visual data: features, polygons, polylines, points, tins, rectified and georeferenced photos.

Different end products are attained by approaching the data from one of the columns. It is possible to explore data from an administrative point of view, in which the Data Collections illustrate the division of the physical excavation into practical objects of investigation and the usage of Documentation Events results in what could best be described as a log or diary, listing the time and place of all data generating actions which take place during and after the excavation. It is also possible to approach data from an interpretative, archaeological point of view in which the individual as well as the combined archaeological structures are visualized and described. Finally, it is possible to approach the data from a completely visual point of view, querying the spatial or graphical data directly to create illustrations that combine vector and raster representations.

The Ladder of Documentation

Another important issue was to find a way to reinforce the relationships between the different data types or classes. It is self-evident that cross-referencing is the precondition for data coherency and usability; e.g. finds are found within something and photos show something, features cut something, deposits overlay something and so on.

The archaeological interpretation in the new

Required	D#	Data Collection
	D#B#	Documentation Event
Optional	FT#	Soil Types
	L#	Deposits
	A#	Basic Structures (features)
	OA#	Overall Structures
	X#	Finds and Samples
	F#	Photos
	O#	Rectified Photos
	I#	Illustrations
	Etc….	

Table 1. Expandable system of classes, identified by prefix.

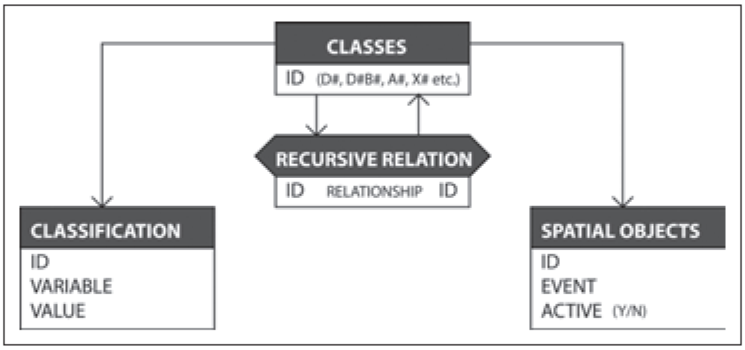

Figure 3. The database model.

procedures is based on a hierarchy of classes; from interpreting the basic archaeological structures (postholes etc.) and their interrelations across the combined, overall structures (buildings etc.), to the final meta-constructs, which are the structures which hold the abstract archaeological interpretations. In order to ensure every cross-reference, one must start at the topmost 'step' and go down the ladder when documenting, and for each step define the relation backwards, up the ladder (Fig. 2).

These documentation principles require strict procedures in which one always refers 'back' on the ladder of documentation in order to be certain that every archaeological construct or class is related to something else. The cornerstone is an array of different data types or classes. Due to the dynamic data structure it is possible to extend or reduce the amount of different classes to facilitate efficient documentation specifically aimed at the overall strategy of the excavation. At the same time, the strict hierarchical structure is not a limiting factor since it is possible for any type of class to relate to any other.

In Practice

The practical realization of the data structure and documentation principles is based on *one* single GIS layer of features in which polygons, polylines and points are all layered on top of each other. Each feature is tagged with three pieces of information which relate the spatial object to the classes in the data structure and facilitate queries from each of the columns. Each feature has a unique identifier, a Documentation Event that 'created' this specific feature and information about whether or not this particular spatial object is part of the current archaeological interpretation. This solution serves as a very efficient approach to prevent redundant information or control derived data. From a data management point of view, the problems of the inherent anarchy of continuously growing amounts of different GIS layers and tables seen in the traditional GIS-solutions are completely eliminated as one needs only to handle a single GIS layer. The database model itself is to some extent based on the object-oriented data models of the GUD, IDEA and ArchaeoInfo projects (Andresen and Madsen 1996a; 1996b; Madsen 2003) in which the dynamic classification and description of data is central, as is the notion of being able to relate everything to anything in a very simple structure (Fig. 3).

Furthermore, spatial data is integrated within the data structure and is not kept as external files, which facilitates more complex descriptions of the spatial data, the elimination of redundant data and a faster and more efficient search and display of features, finds,

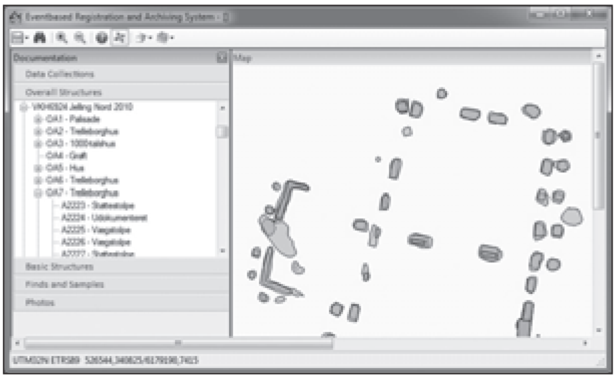

Figure 4. ERAS screenshot showing several documentation events on top of each other.

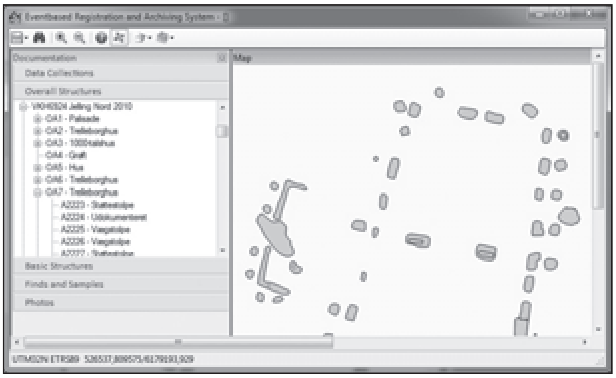

Figure 5. ERAS screenshot visualizing the final interpretation based on queries by event.

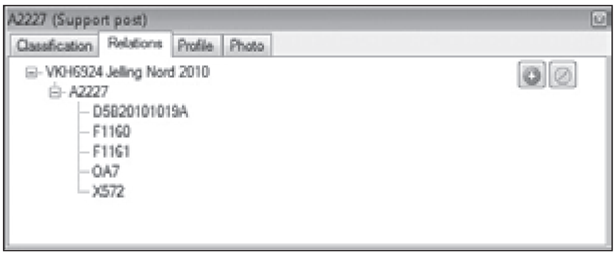

Figure 6. ERAS screenshot illustrating the recursive classification and relations.

structures, topographical data and historical maps.

At the core of the data structure lies the classes-object, which holds any class within the documentation. The hierarchies of Data Collections and Documentation Events, as well as the relations between all other classes, are created by a recursive relation on the classes-object. In addition, a class is connected to feature data and the archaeological classification and description. As a very usable remnant from the traditional documentation principles, different types of classes are identified by their prefix. This facilitates querying and performing thematic mapping on the basis of the single GIS layer of features (e.g. mapping all trench borders by querying on features that have the prefix D# and mapping all basic structures by querying on features that have the prefix A# and so on).

So far, the event-based archaeological registration principles have delivered very promising results, but it must also be emphasized that the procedures are under continuous development, practical testing and evaluation.

ERAS – Event-based Registration and Archiving System

An on-going project is the development of the database user interface, which facilitates both graphical and textual querying of data. In its present form the ERAS (Event-based Registration and Archiving System) consists of a .NET application running as a client on Microsoft Windows systems. One of the most important requirements for the system from the very beginning was the ability to easily migrate historic GIS data, mainly in the form of MapInfo tables and ESRI Shapefiles. Import functions combine GIS data with Microsoft Excel Spreadsheet and Access Database data. Data is up-qualified to fit within the framework of the event-based principles and finally stored as a single XML document per excavation or site. The XML documents are either stored locally on the client computer or retrieved from a central web server through a simple user login form.

Apart from the inability to store, for example, raster data internally in the XML documents, the advantages and possibilities of organizing data by the use of Documentation Events have already become evident. All excavation data is consolidated and easily accessible.

The screenshots in figures 4 and 5 show examples from the excavations carried out by the Danish National Museum in the Jelling Project 2009-2011 (http://jelling.natmus.dk/jellingprojektet/language/uk/). We see postholes in a Viking house documented by several Documentation Events on top of each other (Fig. 4), which by querying can be reduced to visualize the final interpretation of the Overall Structure (Fig. 5). In figure 6 the recursive relation between different classes in ERAS is illustrated.

The continuous development of both documentation principles and application will deliver an approach that is able to include and exploit many more of the possibilities of the digital media. We should be able to introduce new concepts and new data types such as 3D point clouds, laser scans and different types of raster data, the end goal being the increase of overall archaeological data quality combined with the appropriate metadata to describe the history of data collection, generation and interpretation.

Acknowledgements

The methodological and theoretical framework presented, as well as the practical implementation and testing is the product of the cooperative efforts of Casper Skaaning Andersen, Liv Stidsing Reher-Langberg and Mads Holst from the Department of Prehistoric Archaeology at Aarhus University.

Bibliography

Andresen, J., and Madsen, T. 1996a. “IDEA - the Integrated Database for Excavation Analysis.” In *Interfacing the Past. Computer Applications and Quantitative Methods in Archaeology CAA95,* edited by H. Kamermans, and K. Fennema, 3-14. Leiden: Analecta Praehistorica Leidensia 28.

Andresen, J., and Madsen, T. 1996b. “Dynamic classification and description in the IDEA.” In *III International Symposium on Computing and Archaeology,* edited by P. Moscati *Archeologia e Calcolatori 7*:591-602.

Booch, G. 2007. *Object-Oriented Analysis and Design with Applications*. Boston: Addison-Wesley.

Castleford, J. 1992. “Archaeology, GIS and the time dimension: an overview.” In *Computer Applications and Quantitative Methods in Archaeology 1991,* edited by G. Lock, and J. Moffett, 95-106. Oxford: ArchaeoPress.

Constandinidis, D. 2005. “Time to Look for a Temporal GIS.” In *The World is in your eyes: Computer Applications and Quantitative Methods in Archaeology 2005,* edited by A. Figueiredo, and G. Leite Velho. Tomar: CAAPortugal 2007.

Dennis, A., Haley Wixom, B. H., and Tegarden, D. 2005. *Systems Analysis and Design: An Object-Oriented Approach with UML.* John Wiley & Sons.

Eaglestone, B., Hilton, R., and Rold, L., 1996. “GENREG: A historical data model based on event graphs.” *Lecture Notes in Computer Science* 1134:254-263.

Hadzilacos, T., and Stoumbou, P. M. 1996. “Conceptual Data Modelling for Prehistoric Excavation Documentation.” In *Interfacing the Past. Computer Applications and Quantitative Methods in Archaeology CAA95,* edited by H. Kamermans, and K. Fennema, 21-30. Leiden: Analecta Praehistorica Leidensia 28.

Hodder, I. 1997. “Always momentary, fluid and flexible: Towards a reflexive excavation methodology.” *Antiquity* 71:691-700.

Hodder, I. 1999. *The Archaeological Process: An Introduction.* Oxford: Blackwell.

Jones, A. 2002. *Archaeological Theory and Scientific Practice*. Cambridge: Cambridge University Press.

Karlsson, H. 2001. "Time for an Archaeological 'Time-Out'?" In *It's About Time: The Concept of Time in Archaeology*, edited by H. Karlson, 45-59. Göteborg: Bricoleur Press.

Koussoulakou, A., and Stylanidis, E. 1999. "The use of GIS for the Visual Exploration of Archaeological Spatiotemporal Data." *Cartography and GIS* 26 (2):152-60.

Koussoulakou, A., Patias, P., and Stylianidis, E. 2000. "Documentation and Visual Exploration of Archaeological Data in Space and Time." *Proceedings of the Third International Conference on Ancient Helike and Aigialeia*. Diakopto, Achaia.

Light, R. B., Roberts, D. A., and Steward, J. D. 1986. *Museum Documentation Systems: Developments and Applications*. London: Butterworths.

Lucas, G. 2001. *Critical approaches to fieldwork: contemporary and historical archaeological practice*. London: Routledge.

Madsen, T. 2003. "ArchaeoInfo. Object-oriented information system for archaeological excavations." Paper presented at the *Computer Applications and Quantitative Methods in Archaeology. CAA2003 8-12 April, Vienna. Leidensia* 28:3-14.

Olivier, L. C. 2001. "Duration, Memory and the Nature of the Archaeological Record." In *It's About Time: The Concept of Time in Archaeology*, edited by H. Karlson, 61-70. Göteborg: Bricoleur Press.

Reed, P. A. 1995. *CIDOC Relational Data Model - A Guide*. International Documentation Committee of the International Council of Museums.

Rold, L. 1992. "Syntheses in object oriented analysis." In *Proceedings of Computer Application and Quantitative Methods in Archaeology, CAA92*, edited by J. Andresen, T. Madsen, and I. Scollar. Aarhus: Aarhus University Press.

Rold, L. 1995. "Informationsteknologi på Nationalmuseet." *Fortid og Nutid* 279-302.

Scollar, I. 1998. "AirPhoto - A WinNT/Win95 program for the geometric processing of archaeological air photos." *AARGnews* 16:36-38.

Sellis, T., Koubarakis, M., Frank, A., Grumbach, S., Gueting, R. H., Jensen, C., Lorentzos, N., Manolopoulos, Y., Nardelli, E., Pernici, B., Schek, H. J., Scholl, M., Theodoulidis, B., and Tryfona, N. (eds.). 2003. *Spatiotemporal Databases: the Chorochronos Approach*. Berlin: Springer Verlag.