

Steps towards a participatory digital library and data archive for archaeological information

Isto Huvila

Department of ALM, Uppsala University and Department of Information Studies, Åbo Akademi University

Abstract

The e-Science related literature sees digital libraries, digital archives and databanks as central collaborative infrastructures for scientific and scholarly research. The present paper discusses preliminary findings and observations from an action research study on a participatory digital library and data archive for archaeological and archaeology related data. The project aims to develop a combined digital repository and collaboratory for archaeological research with a special emphasis on participation and semantically rich relations between individual digital artifacts. The discussed digital library is based on semantic wiki technology and developed together with archaeologists from a small Nordic archaeology consultancy. The findings of the study indicate that the principal challenges associated with the life-cycle approach relate to the paradox of formality and informality of documentation and to the implications the system has on the customary procedures of archaeological information work.

Introduction

The concept of digital libraries has shifted from referring to simple repositories of access and storage to betoken central infrastructural nodes of information work. Borgman remarks that a digital library can be both a provider and a repository of exploitable information (C. Borgman, 1999). Chen perceives digital libraries as incorporating the entire life-cycle of information (Chen, 1998, 18). The e-Science related literature sees digital libraries, digital archives and databanks as central collaborative infrastructures for scientific and scholarly research. In science, the global digital libraries and databanks have proven to be successful in several disciplines, genomics being probably the most prominent example (C. L. Borgman, 2007). E-Science databanks and infrastructures have been developed also within arts and humanities, even though it has become apparent that arts and humanities specific approaches are needed (Blanke et al., 2009).

The notions of digital libraries and e-Science are both based on assumptions of 'globality'. The globality may be global in literal sense or globality within a specific domain or context. Chowdhury presents a vision of a "digital global library" (Chowdhury, 1999, 431-432) in a literal sense and many others present either explicit or implicit suggestions

that one of the benefits of digital libraries is that the information is available for anyone in more or less restricted sense (e.g. Huvila, 2007). Globality, accessibility and openness are not only practical, but also political notions related to the idea of information society (May, 2002) and therefore not necessarily quite straightforward to implement in practice. Experiences have shown that many of the most successful digital libraries have been in some sense local. The repositories have been specific to a geographical, organisation, disciplinary or cultural domain or context (Sure & Studer, 2005). Similarly, the critical issues such as long-term reliability and consistency (Frommholz et al., 2004) have proved to be easier to manage when the context is as limited as possible.

Many of the challenges on the way to the globality have been acknowledged for some time (Fox et al., 1995; Fox & Marchionini, 1998). The number of stakeholders causes difficulties (Arms, 2000, 19). As Blandford (Blandford et al., 2001) has noted, users become easily disoriented with seemingly consistent, but internally inconsistent universal systems, which do not correspond with their needs and expectations of working with the information. A poor match between consumer-users and digital libraries means that the user does not get information in exchange for his investment of time and effort, which reduces the significance of the collection (ref. e.g. Hong et al., 2001) and decreases the quality, scope and relevance of the digital library (Hill et al., 2000; Klas et al., 2006).

In an earlier study it was suggested that one of the reasons for the poor match between digital libraries and their users is that repositories tend to focus on specifics of individual information objects instead of their mutual relations (Huvila, 2007). The most significant problems reported by the informants of the study did not seem to relate to the individual digital artifacts or their description, but to the experienced tendency of the digital repositories to restrict the ways of conceptualising larger meaningful entities.

The present paper discusses preliminary findings and observations from an action research study on a participatory digital library and data archive for archaeological and archaeology related data. The digital library is meant to be global in the context of archaeology and related scholarly, scientific and professional disciplines. The project aims to develop a combined digital repository and collaboratory for archaeological research with a special emphasis on participation and semantically rich relations between individual digital artifacts. The emphasis is on facilitating information work on the entire life-cycle of information in order to address the earlier issues related to the discrepancies of information creation and use. The focus of the project is in exploring the technical and theoretical issues related to archaeological e-Science. The findings of the study indicate that the principal challenges associated with the life-cycle approach relate to the paradox of formality and informality of documentation and to the implications the system has on the customary procedures of archaeological information work.

Representation of knowledge in archaeology

Orlandi (1999) summarises the constituent premises of archaeological informatics while discussing the ideas of Guimier-Sorbets (1996) on the role of multimedia in the publication of archaeological materials:

1. Like all scientific disciplines, archaeology is concerned on the treatment of information: The relation of archaeology and informatics and multimedia may thus be divided into three phases: elaboration of the documentation, interpretation, and diffusion of the results.

2. The third phase, diffusion, has to be based on the notion of publishing cumulative information.

3. The second notion begets a need to make effective typologies and to standardise archives and the archival processes.

4. Empowering the communication of ideas and information between researchers is constituent to the archaeological work.

According to the premises, first, archaeological information systems need to focus and accommodate the information work processes related to archaeological work instead of emphasising atomic data objects. The systems should grasp not only documentation, but also the following steps of interpretation and diffusion of the results. Further, the systems should support accumulation of information, not only data. In order to be cumulative, a relevant degree of standardisation is needed. However, standardisation should be seen as instrumental to the accumulation of information and communication of ideas instead of being an aim of its own right.

According to Huvila (2006), the knowledge organisation challenges of the practises of archaeological information organisation may be classified to two major categories, the 1) structure and 2) dynamics related issues. The first category relates to the technical issue of how a data structure is capable of representing the complexity of archaeological information (the critical success factor of *fit* archaeological information work, Huvila, 2006). The second category is related to the question of how the data structures support its dynamics (*sustainability*, Huvila, 2006). Knowledge organisation schemes and data structuring approaches discussed in the literature and used in practice encompass different approaches to address the challenges related to these categories.

Most of the current data management systems used in archaeological documentation and information processing, are based on the relational data model (Codd, 1970 ref. e.g. Drap & Long, 2001; Hynst et al., 2001 and Papalexopoulos et al., 2001). The relational model is technically very efficient. A consequent reason for its popularity in archaeology is that it is the regular data model used in the majority of commercial and open source database management systems (ref. Bell & Eiteljorg, 2006). The major problems with the model are that it is not optimal for processing complex and heterogeneous data, and the attempts to process complex data lead to highly complex and difficult to manage relational structures.

The problems of the relational model have lead to proposals of using more tractable data models. Hyperlinking represents an exemplary alternative approach, which has been suggested as a substitute to the relational model (e.g. Agnello et al., 2003). XML, Semantic Web and ontology related technologies promise flexibility, man and machine readability and extensibility (Niccolucci & Cantone; Schloen, 2001; Niccolucci, 2002; Barchesi, 2004; Bell & Eiteljorg, n.d.; ref. also Cantone, 2002 and Ross, 2003). Neither hyperlinking nor ontology based approaches are entirely unproblematic. Veltman summarises the essential cultural heritage related problems of the present ontological approaches advocated by the Semantic Web movement to the issues relating to the 1) management of different world-views, 2) evolution of the definitions and meanings, 3) distinction between the words and concepts, 4) handling of the new classes of relations and 5) dynamism of the models of knowledge organisation (Veltman, 2004).

The first problem of the digital data processing relates to the use of absolute estimations instead of subjective interpretations, which are prevalent in the real-life contexts (ref. Gabucci, 2005). The poor fit of formal data structures in archaeology, is a result of the nature of the archaeological knowledge, which is based mostly on hermeneutical interpretations instead of ontological representations of truth (cf. Bénéel et al. (2001)). In an attempt to implement a support mechanism for a more advanced degree of subjectivity, Niccolucci et al. have demonstrated the possibilities of using fuzzy logic to represent confidence and reliabilities (ref. Niccolucci et al., 2001; Hermon & Niccolucci, 2003; Hermon et al., 2004 and D'Andrea 2004). Besides the subjectivity, similarly pressing problems relate to the representation of the complete dimensionality of the archaeological space (ref. Barceló & Vicente, 2004), the persisting issue of the huge amount and fast accumulation of data, and the linking and organising of all related information in meaningful entities.

The Semantic Web approach addresses the notion of multiplicity of knowledge claims by multiple coinciding ontologies (i.e. 'multiple overlapping truths') (Maedche et al., 2003). From the ontology point of view, the overlap, evolution and different versions are a problem, which requires specific management (Noy & Klein, 2004). From the hermeneutically aligned knowledge and information point of view, on the other hand, the approach of seeing the overlap as a problem, is a problem itself, because according to the viewpoint, knowledge is perceived to be overlapping *per se*. There is a clear difference between expressing a claim as an interpretation or as a parallel truth.

A further approach to counter the rigidity of the ontologies and taxonomies is based on participatory description and management of information. Folksonomies are based on collective tagging of resources and statistical clustering of the tags. Kansa (2006) suggests a folksonomy based approach for archaeological documentation. Zhou and Bénéel (2008) have implemented such a system for studying Greek vases. Tags are potentially useful, but they do not *per se* negate the need of formal taxonomies or controlled vocabularies. From the management point of view, the most acute concern with the approach is that the folksonomies are equally difficult to control or manage as the knowledge and information are. Social tagging may empower usability and make information more findable (Morville, 2005; Golder & Huberman, 2006), but it does not contribute specifically to making it more manageable.

Fox (2005) and his research group have focussed on a theoretically founded methodology for constructing a generic digital library framework, which may be used as a basis for domain specific digital libraries. The so called 5S methodology is based on an idea of first providing a highly general data structure. The new items and collections are imported with their respective schemas and directly mapped to the existing framework. When needed, the existing schema is augmented with new additional objects (i.e. entities). The proposed approach has many strengths. The inherent problem with such an approach is, however, the increasing complexity of the growing system when new collections are added. Simultaneously, the number of objects within each dimension increases and due to partial matches between them (Raghavan et al., 2005). For instance, objects VESSEL and POTTERY discussed by Raghavan et al. (2005), are not necessarily exclusive as object types. They merely originate from different collections.

Methods and layout of the study

In order to address and test the issues of technology and knowledge organisation system adaptation, a semantic wiki based archival system framework was developed to house the materials of three different cultural historical sites located in Southwestern Finland. The work was conducted in cooperation with a small local archaeological contractor interested in developing the management of excavation and conservation data. The work was conducted by the author in collaboration with two archaeologists. The archaeologists worked at a small private archaeological company. One of the informants was working as a field director and the CEO of the company and the other one as a researcher with special competence in archaeological field documentation and measurement. Features and requirements were discussed in meetings, implemented later and tested in field in actual locations with real data. Part of the software development and evaluation of appropriate laptop and tablet-pc computers was conducted directly on field. The meetings were started by a general round of ideas and thoughts after the last meeting and proceeded with presentation of new features in the system, spontaneous comments by archaeologists, specific questions and issues in the development work requiring consultation of the archaeologists. Finally, the next stages of development were agreed and a new meeting was booked. The author took notes during the meetings and used these documents during the following phase of the research. During the field trials, the process was accelerated. A meeting held in morning was followed by system development and elaboration phase (about an hour) and field trials comprising documentation of archaeological features on site (about 30 minutes). The trials were followed by a short informal meeting and approximately two new additional development-trial cycles per day.

The development process was based on action research (Greenwood & Levin, 2000; Kemmis & McTaggart, 2000) as a principal methodology of investigation. According to the action research approach, the problems (research question) addressed during the research were the misfit of formal data and qualitative interpretations, and the difficulty of representing the full dimensionality of archaeological data. After a brief initial analysis of the situation, the problems were decided to be addressed by building a functioning prototype of a digital library with appropriate basic functionality and comprehensive possibilities to develop both the technical and content aspect together with the actual users of the archive. The present paper discusses the first observations and findings of the process. A full evaluation of the project is yet to be conducted.

The semantic wiki based approach was chosen to meet the structural and dynamics related challenges of fit and sustainability (Huvila, 2007). The principal strength of wiki based approaches is in collaboration and automatic versioning system, which keeps all changes recorded and allows infinite rollbacks to earlier versions. The approach promised also to deliver structural and representational flexibility and formality when it was needed, it provided means to work with multiple media forms and adaptability for implementing sustainability strategies. Essentially the approach was a compromise of flexibility and formality, which was estimated to accommodate the needs of specific branches of archaeological and related research (Huvila, 2006) and to provide enough structure for maintaining a degree of findability and searchability (e.g. Sure & Studer, 2005). Further, the platform allowed continuous edits, accumulation of material and inclusion of interpretative material and making

the material available on the web for the purposes of diffusion of the information (Orlandi, 1999). In addition, the platform was successfully used in an earlier digital archive project (Huvila, 2008).

The digital library was based on a system of semantically typed, described and interrelated records. A single object (a page in wiki terminology) may represent a single actual physical (stored outside the system) or digital (stored inside the system) document, a concept, a type or a part of them. Thus an object (wikipedia) can be a resource or a description about an object (cf. Kotelnikov et al. 2007). Objects can be or represent archaeological objects, sites, features, structures, documents, books or any other entity of information. Each object is supposed to have one or several *types* including technical types of Semantic Mediawiki system (such as Page, String, Number or Date, see Krötzsch, Markus et al. 2007), internal types used to distinguish several archival and functional categories of records (e.g. different collection related and archaeological entities, types or documents, example records, keywords and digital reconstructions) and types based on CIDOC-CRM classes, e.g. “E31 Document” for documents and “E38 Image” for images (Crofts et al. 2007). Besides by typing, relationships between different objects are expressed by using CIDOC-CRM Properties (e.g. Crofts et al. 2007), and when needed, by using other preferably standardised types of semantic linking. Besides having a *type*, each object can have an unlimited number of *properties*, which can be shared by different objects. The property *colour* can be used to denote colour of objects expressed using a variety of systems (such as Munsell or Pantone for different kinds of objects) documented within the repository. Lastly, objects can belong to *categories* (e.g. Kuusisto Castle,) and dynamic categories called *concepts* (e.g. “Iron objects from Kuusisto Castle”: all objects belonging to category Kuusisto Castle and having a material property referring to iron).

Towards a digital library

The project was based on the earlier observations by Orlandi (1999) and Huvila (2007). One of the initial concerns of the project was to find out a method and technology that would support flexible collaborative addition and modification of descriptors for each object without limiting the number of individual descriptors. As expected on the basis of earlier studies (Huvila, 2007), the principal issues were not related to the issues of how to describe an object, how to make the descriptions consistent and how to manage these descriptions. The functionality of the chosen platform, the semantic wiki supported relatively well simultaneous use of both highly unconstrained and very formal descriptions.

The second major concern was to implement a system with similarly flexible tools for grouping and classifying the objects that would address the concern for more focus on communicating and documenting where an object belongs rather than what it is and how it should be described (Huvila, 2007). In many archaeological information systems the tendency is to focus on meticulous description of individual objects as a part of a hierarchical structure of sites, archaeological contexts and layers its relevance resides only in the analytical hierarchy, not in any meaningful human context. The semantic wiki based approach provides means for both free form and structured description of individual objects (i.e. pages) and their relations in form of semantic types, properties and concepts. Objects can be related to each other by referring to their common properties like in relational data

structures. Another possibility is to explicitly create organisation in form of networks and hierarchies by using *concepts* for classifying objects.

The proposed wiki based system addresses a number of problems related to the flexibility and rigidity of relational and ontology based data structures. In theory, it escapes also the use of implicit universal ontology found in the 5S system (Fox, 2005). Semantic wiki does not, however, escape the issues related to increasing complexity of the system of knowledge. It gives freedom to renegotiate and complement information and tools to manage the structure, but leaves the actual decisions to the users of the system. In comparison to the entirely uncontrolled folksonomy based approaches, the system uses a combination of controlled and uncontrolled descriptions.

Finding an appropriate (digital) structural framework to represent the archaeological data is a question of conforming the structures of human information with an appropriate set of computer based structures. The four factors, which empower the computerisation of archaeological data identified Richards and Ryan identified in 1985 (for Richards & Ryan, 1985, 16 ref. Lock, 2003, 4) are that

1. duplication should be avoided in the selection of attributes,
2. attributes need to be separated of the attribute states,
3. the deliberate human selection needs to be identified on an appropriate level of precision. Only reasonable attributes should be documented, and
4. that the frame of reference of the study affects the required information.

In spite of their age, the factors have retained their basic validity. The benefits and challenges of the semantic wiki approach concerning each factor are summarised in table 2. The consequence of that the wiki philosophy relies on human effort is that also the semantic wiki based archaeological archive requires special care to be taken so that duplication is avoided and the attributes (properties) and attribute states (property values) are kept separate (Table 2). It is important to check the list of existing properties and categories before adding a new one. The benefit of a wiki-based approach is that it is easy to make cross-references and to merge and link properties when duplicates are found. Similarly the properties and concepts can be documented within the system, on one hand, in order to avoid misconceptions and, on the other hand, to avoid the need to browse external references. The functioning of the wiki based approach to knowledge organisation functions depends on multiple factors. The existence of a critical mass of participants is important as well as the consistence of their work. Both successful and unsuccessful wikis exist. However, both kinds of examples of traditional databases exist as well.

According to the experiences so far, the most problematic issue related to the chosen wiki based approach is the way how each individual object (i.e. page) is identified by a unique title used both as a label of the object and as its reference in the links and URL-addresses. The problem is that the name of an archaeological object depends on how it is interpreted. A shard of pottery may be labelled as representing a vessel, pottery shard, lamp or something else. Similarly archaeological sites and excavated features may have different names depending on the context of interpretation. When something is found during an excavation, an archaeologist seldom knows definitely what the object or feature is and whether it is a part of another object or a feature. Both the use of arbitrary reference numbers and contextual labels based on the original interpretation were tested. The latter approach proved to be somewhat more useful, but not without its inherent problems possibly

	Benefits	Challenges
1) Duplication should be avoided	Tools for managing attributes exist; easy to manage existing duplicates	No automatic enforcement; human effort needed
2) Attributes separate of attribute states	Semantic wiki provides a framework for managing attributes	Requires planning (applies to all data management systems)
3) Human selection	Can be precised for each case individually; no global constraints	Need to establish a common frame of reference (in form of data policy)
4) Frame of reference	Flexibility; several frames can coexist	Need to establish a common frame of reference (in form of data policy)

Table 2: Benefits and issues related to semantic wikis and archaeological data

leading to a need to relabel or move large numbers of objects or risking to compromise the intelligibility of the labels.

According to the experiences from the project so far, it seems to be reasonable to expect that the chosen approach addresses the first two factors (Table 2). It is the last two, however, with which the semantic wikis have their most significant benefits. They are more contextual notions in comparison to the first ones. In a wiki, it is possible to implement the level of precision and the impending information needs according to the present standards and future expectations on the basis of the available resources without compromising significant future changes. The contents of a semantic wiki are flexible to update and restructure. The challenge is to enforce a minimum level of standardisation in documentation for the sake of findability and retrievability of information. In practice, this needs to be managed in form of a common policy. The benefit of using a semantic wiki is that the system supports flexible documentation of such policies and conventions within the system itself.

The notion of precision appeared as twofold during the development process. Archaeologists tend to prefer high levels of precision *per se* during the documentation process, even if a high precision would not be strictly necessary (Huvila, 2007). By allowing artificially high levels of precision, computers seem to amplify this tendency and do simultaneously unwittingly support a sense of spurious accuracy even when the initial data is far from being precise (Kantner, 2000). Therefore the fit of the actual precision and the documented precision needs to be considered and documented meticulously before actually deciding which data to input at any particular moment. Semantic wiki allows variation in precision on any level from individual objects to larger categories. Simultaneously as it is important to include any relevant data, which is likely to be lost is not documented, it is possible to document less relevant features or something that is documented elsewhere in lower detail. Levels of precision can be also combined and individual formal properties complemented with draft notes and preliminary commentaries. Instead of forcing to be comprehensive beyond the basic immediate needs, the approach makes possible to focus on providing the

often lacking context and relations (i.e. the source and viewpoint, Huvila, 2007) of the data to other data.

The most significant issue related to the practical development of the new system was that it presented essentially a quite new approach to archaeological information work. Archaeologists are accustomed to document their work in an itemised manner in the databases and to hold the general picture surprisingly separate from the details of the documentation. The semantic wiki presents another kind of approach where individual objects and details are processed together with their inter-relations and the general understanding of the archaeological site.

In spite of the constituency of the information and communication technology, the computers themselves do not beget independently any constraints or affordances associated to the systems of knowledge organisation. In spite of the significance, which computers have in provoking changes, the observations made during the present study underline the fact that digital information processing in archaeology is directly related to the general tribulation of formalism in the humanities (Orlandi, 1999, 1997). As suggested by Lock and Orlandi (Lock, 2003, xiii; Orlandi, 1993), the archaeologists' information work is not profoundly 'computer minded', but still closely configured to accommodate the present digital and analogue systems of organising knowledge and information. The challenge is how to reconfigure information work by introducing digital systems in the traditionally implicit information processing and by convincing the archaeologists that not all of the traditional work procedures are contributing to the documentation in the new system.

Conclusions and future work

The first observations and findings from the development of a semantic wiki based archaeological digital library system confirm many of the anticipated challenges. Semantic wiki approach proved to be effective in shifting the organisatory foundation of the digital library from the notion of organising and annotating individual documents to networks formed by archaeological material and other sources of information. However, even though the change was made to make the system follow closer the actual patterns of archaeological information work, the change of paradigm proved to be difficult to internalise by the participating archaeologists.

There are still several aspects requiring special attention in the digital library. User acceptance is a major issue like the need for training. Other aspect requiring further testing are the sustainability of the networked data model and whether the naming of the objects functions in the long run. The system has already shown its benefits in bringing together different kinds of information from different sources and providing a collaborative environment for information work in archaeology. In this sense it has the technical functionality to become an e-Science collaboratory for archaeological work.

References

- Agnello, F., Corsale, R., Franco, V., Lo Brutto, M., Midulla, P., Orlando, P., et al. (2003). Cultural Heritage and Information, An investigation into a Dedicated Hypertext. *The International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences*, XXXIV(5/W12), 7-12.

- Arms, W. Y. (2000). *Digital Libraries*. MIT Press.
- Barceló, J. A., & Vicente, O. (2004). Some problems in archaeological excavation 3d modelling. In M. der Stadt Wien Referat Kulturelles Erbe Stadtarchäologie Wien (Ed.), *Enter the past, the e-way into the four dimensions of cultural heritage. caa 2003 computer applications and quantitative methods in archaeology, proceedings of the 31st conference, vienna, austria, april 2003* (Vol. 1227, p. 400-404). Archaeopress.
- Barchesi, L., C. Ceccarelli. (2004). Linguaggi dichiarativi per la ricerca archeologica. *Archeologia e calcolatori*, 15, 95-113.
- Bell, T., & Eiteljorg, H. I. (n.d.). *XML: Panacea or panettone?* (Abstract of a paper presented in the 34th Annual Meeting and Conference of Computer Applications and Quantitative Methods in Archaeology CAA2006, Fargo, April 18-21, 2006)
- Bell, T., & Eiteljorg, H. I. (2006, Winter). Still more on XML - finding a common ground. *CSA Newsletter*, 18(3). Available from <http://csanet.org/newsletter/winter06/nlw0601.html>
- Bénel, A., Egyed-Zsigmond, E., Prié, Y., Calabretto, S., Mille, A., Iacovella, A., et al. (2001). Truth in the digital library: From ontological to hermeneutical systems. In *Proceedings of the fifth european conference on research and advanced technology for digital libraries, darmstadt, september 4-9, 2001* (Vol. 2163, p. 366-377). Berlin: Springer.
- Blandford, A., Stelmaszewska, H., & Bryan-Kinns, N. (2001). Use of multiple digital libraries: a case study. In *Jcdl '01: Proceedings of the 1st acm/ieee-cs joint conference on digital libraries* (pp. 179-188). New York, NY, USA: ACM Press.
- Blanke, T., Hedges, M., & Dunn, S. (2009). Arts and humanities e-science—current practices and future challenges. *Future Generation Computer Systems*, 25(4), 474-480. Available from <http://www.sciencedirect.com/science/article/B6V06-4TPF47W-4/2/563da1634375a83a038504025c37020d>
- Borgman, C. (1999). What are Digital Libraries? Competing Visions. *Information Processing and Management*, 35(3), 227-243.
- Borgman, C. L. (2007). *Scholarship in the Digital Age: Information, Infrastructure, and the Internet*. Cambridge, MA: MIT Press.
- Cantone, F. (2002). 3D standards for scientific communication. In G. Burenhult & J. Arvidsson (Eds.), *Archaeological informatics: Pushing the envelope caa 2001. computer applications and quantitative methods in archaeology. proceedings of the 29th conference, gotland, april 2001* (Vol. 1016, p. 163-172). Archaeopress.
- Chen, S.-S. (1998). *Digital libraries : the life cycle of information*. Better Earth.
- Chowdhury, G. G. (1999). *Introduction to modern information retrieval*. London: Library Association.
- Codd, E. F. (1970, June). A relational model of data for large shared data banks. *Communications of the ACM*, 13(6), 377-387. Available from <http://www.acm.org/classics/nov95/>
- Crofts, N., Doerr, M., Gill, T., Stead, S., & Stiff, M. (Eds.). (2007). *Definition of the cidoc conceptual reference model*. ICOM/CIDOC Documentation Standards Group, continued by the CIDOC CRM Special Interest Group. (Retrieved from http://cidoc.ics.forth.gr/docs/cidoc_crm_version_4.2.2.pdf)

- D'Andrea, A. (2004). L'entropia dell'archeologia computazionale ovvero dall'ordine al disordine. *Archeologia e calcolatori*, 15, 219-238.
- Drap, P., & Long, L. (2001). Towards a digital excavation data management system: the "grand ribaud f" e-truscan deep-water wreck. In *Vast '01: Proceedings of the 2001 conference on virtual reality, archeology, and cultural heritage* (pp. 17-26). New York, NY, USA: ACM Press.
- Fox, E. A. (2005). *Digital libraries: Archaeology, automation, and enhancements* (Tech. Rep.). Invited talk. The International Advanced Digital Library Conference (IADLC), Nagoya University, Japan, August 25-26.
- Fox, E. A., Akscyn, R. M., Furuta, R. K., & Leggett, J. J. (1995). Digital libraries. *Communications of the ACM*, 38(4), 22-28.
- Fox, E. A., & Marchionini, G. (1998). Toward a worldwide digital library. *Communications of the ACM*, 41(4), 29-32.
- Frommholz, I., Knezevic, P., Mehta, B., Niederée, C., Risse, T., & Thiel, U. (2004). Supporting Information Access in Next Generation Digital Library Architectures. In *Proceedings of the sixth thematic workshop of the eu network of excellence delos, cagliari, italy* (pp. 49-60).
- Gabucci, A. (2005). *Informatica applicata all'archeologia*. Roma: Carocci.
- Golder, S., & Huberman, B. A. (2006). Usage patterns of collaborative tagging systems. *Journal of Information Science*, 32(2), 198-208.
- Greenwood, D. J., & Levin, M. (2000). Handbook of qualitative research. In N. K. Denzin & Y. S. Lincoln (Eds.), (pp. 85-106). Thousand Oaks, CA: Sage.
- Guimier-Sorbets, A.-M. (1996). Le traitement de l'information en archéologie: archivage, publication et diffusion. *A&C*, 7, 985-996.
- Hermon, S., & Niccolucci, F. (2003). A fuzzy logic approach to typology in archaeological research. In *Caa 2002 the digital heritage of archaeology. computer applications and quantitative methods in archaeology. proceedings of the 30th conference, heraklion, crete, april 2002* (p. 307-310). Archive of Monuments and Publications, Hellenic Ministry of Culture.
- Hermon, S., Niccolucci, F., Alhaique, F., Iovino, M.-R., & Leontini, V. (2004). Archaeological typologies - an archaeological fuzzy reality. In M. der Stadt Wien Referat Kulturelles Erbe Stadtarchäologie Wien (Ed.), *Enter the past, the e-way into the four dimensions of cultural heritage. caa 2003 computer applications and quantitative methods in archaeology, proceedings of the 31st conference, vienna, austria, april 2003* (Vol. 1227, p. 30-34). Archaeopress.
- Hill, L. L., Carver, L., Larsgaard, M., Dolin, R., Smith, T. R., Frew, J., et al. (2000). Alexandria digital library: user evaluation studies and system design. *Journal of the American Society for Information Science*, 51(3), 246-259. Available from [http://dx.doi.org/10.1002/\(SICI\)1097-4571\(2000\)51:3<246::AID-ASI4>3.0.CO;2-6](http://dx.doi.org/10.1002/(SICI)1097-4571(2000)51:3<246::AID-ASI4>3.0.CO;2-6)
- Hong, W., Thong, J. Y., Wong, W.-M., & Tam, K.-Y. (2001, Winter). Determinants of user acceptance of digital libraries: An empirical examination of individual differences and system characteristics. *Journal of Management Information Systems*, 18(3), 97-124.
- Huvila, I. (2006). *The ecology of information work – a case study of bridging archaeological work and virtual reality based knowledge organisation*. Åbo: Åbo Akademi University Press. (Diss. Åbo Akademi University)

- Huvila, I. (2007). Don't ask me 'what', ask 'to what it is related' in a digital library. In *Proceedings of the libraries in the digital age (lida) 2007 conference*.
- Huvila, I. (2008). Participatory archive: towards decentralised curation, radical user orientation and broader contextualisation of records management. *Archival Science*, 8(1), 15–36.
- Hynst, S., Gervautz, M., Grabner, M., & Schindler, K. (2001). A work-flow and data model for reconstruction, management, and visualization of archaeological sites. In *Vast '01: Proceedings of the 2001 conference on virtual reality, archeology, and cultural heritage* (pp. 43–52). New York, NY, USA: ACM Press.
- Kansa, E. (2006). *Data integration with ArchaeoML and tagging*. (Abstract of a paper presented in the 34th Annual Meeting and Conference of Computer Applications and Quantitative Methods in Archaeology CAA2006, Fargo, April 18–21, 2006)
- Kantner, J. (2000). Realism vs. reality: creating virtual reconstructions of prehistoric architecture. In J. A. Barceló, M. Forte, & D. H. Sanders (Eds.), *Virtual reality in archaeology. proceedings of the computer applications and quantitative methods in archaeology (caa) 1998, barcelona* (Vol. 843, p. 47–56). Oxford: Archaeopress.
- Kemmis, S., & McTaggart, R. (2000). Handbook of qualitative research. In N. K. Denzin & Y. S. Lincoln (Eds.), (pp. 567–605). Thousand Oaks, CA: Sage.
- Klas, C.-P., Fuhr, N., Kriewel, S., Albrechtsen, H., Tsakonias, G., Kapidakis, S., et al. (2006). An experimental framework for comparative digital library evaluation: the logging scheme. In *Jcdl '06: Proceedings of the 6th acm/ieee-cs joint conference on digital libraries* (pp. 308–309). New York, NY, USA: ACM Press.
- Kotelnikov, M., Polonsky, A., Kiesel, M., Völkel, M., Haller, H., Sogrin, M., et al. (2007, January). *Interactive semantic wikis* (Deliverable No. D1.1). Kaiserslautern: NEPOMUK Consortium. (Available at http://nepomuk.semanticdesktop.org/xwiki/bin/download/Main1/D1%2D1/D1.1_v10_NEPOMUK_Interactive_Semantic_Wikis.pdf)
- Krötzsch, Markus et al. (2007). *Special:types*. Semantic Mediawiki. (Retrieved from <http://semantic-mediawiki.org/index.php/Special:Types>)
- Lock, G. (2003). *Using computers in archaeology: Towards virtual pasts*. Routledge.
- Maedche, A., Motik, B., & Stojanovic, L. (2003). Managing multiple and distributed ontologies on the Semantic Web. *The VLDB Journal*, 12(4), 286–302.
- May, C. (2002). *The Information Society: A Sceptical View*. Cambridge: Polity Press.
- Morville, P. (2005). *Presentation "Ambient Findability" in the "Panel on the intersection of IA (information architecture) and KM (knowledge management) at the Second International Conference on Knowledge Management 2005: Nurturing Culture, Innovation and Technology. Charlotte, NC. 27th October 2005*.
- Nicolucci, F. (2002). Xml and the future of humanities computing. *SIGAPP Applied Computing Review*, 10(1), 43–47.
- Nicolucci, F., & Cantone, F. (2003). Legend and virtual reconstruction: Porsenna's mausoleum in x3d. In *Caa 2002 the digital heritage of archaeology. computer applications and quantitative methods in archaeology. proceedings of the 30th conference, heraklion, crete, april 2002* (p. 57–62). Archive of Monuments and Publications, Hellenic Ministry of Culture.

- Niccolucci, F., d'Andrea, A., & Crescioli, M. (2001). Archaeological applications of fuzzy databases. In Z. Stančič & T. Veljanovski (Eds.), *Computing archaeology for understanding the past, caa 2000, computer applications and quantitative methods in archaeology, proceedings of the 28th conference, ljubljana, april 2000* (Vol. 931, p. 107-115). Oxford: Archaeopress.
- Noy, N. F., & Klein, M. (2004). Ontology evolution: Not the same as schema evolution. *Knowledge and Information Systems*, 6(4), 428-440. Available from <http://dx.doi.org/10.1007/s10115-003-0137-2>
- Orlandi, T. (1993). Sulla codifica delle fonti archeologiche. *Archeologia e calcolatori*, 4, 27-38. Available from <http://rmcisadu.let.uniroma1.it/orlandi/pubbli/archeo.html>
- Orlandi, T. (1997). Informatica, formalizzazione e discipline umanistiche. In T. Orlandi (Ed.), *Discipline umanistiche e informatica. il problema della formalizzazione* (p. 7-17). Roma: Accademia Nazionale dei Lincei. Available from <http://rmcisadu.let.uniroma1.it/orlandi/formaliz.html>
- Orlandi, T. (1999). Multimedialità e archeologia. *Archeologia e calcolatori*, 10, 145-157. Available from <http://rmcisadu.let.uniroma1.it/orlandi/pubbli/arcmul.htm>
- Papalexopoulos, D., Kalafati, E., & Papadopoulos, S. (2001). Building memory. In *Vast '01: Proceedings of the 2001 conference on virtual reality, archeology, and cultural heritage* (pp. 27-32). New York, NY, USA: ACM Press.
- Raghavan, A., Vemuri, N., Shen, R., Gonçalves, M., Fan, W., & Fox, E. (2005). Incremental, Semi-automatic, Mapping-Based Integration of Heterogeneous Collections into Archaeological Digital Libraries: Megiddo Case Study. In *Proceedings of the european conference on digital libraries, ecdl 2005, vienna, sept. 18* (Vol. 23).
- Richards, J. D., & Ryan, N. S. (1985). *Data processing in archaeology*. Cambridge: Cambridge University Press.
- Ross, S. (2003, May). Towards a semantic web for heritage resources. In DigiCULT (Ed.), (Vol. 3, p. 7-11). Salzburg: DigiCULT. Available from http://www.digicult.info/downloads/thematic_issue_3_low.pdf
- Schloen, J. (2001, May). Archaeological data models and web publication using xml. *Computers and the Humanities*, 35(2), 123-152.
- Sure, Y., & Studer, R. (2005). Semantic Web technologies for digital libraries. *Library Management*, 26(4/5), 190-195.
- Veltman, K. H. (2004, March). Towards a semantic web for culture. *Journal of Digital Information*, 4(4). Available from <http://jodi.tamu.edu/Articles/v04/i04/Veltman/>
- Zhou, X. (2008, September). Student archival research activity: An exploratory study. *American Archivist*, 71(2), 476-498. Available from <http://archivists.metapress.com/content/n426270367qk3111>