Being Formal and Flexible: Semantic Wiki as an Archaeological e-Science Infrastructure

Isto Huvila
Uppsala University, Sweden. *isto.huvila@abm.uu.se*

Abstract:
The multiple challenges of representing archaeological information both using relational databases and semantic web technologies have been acknowledged in the literature. The present paper discusses findings and observations from an action research study on developing an integrated semantic digital data archive and collaboration platform for archaeological and archaeology related research using a semantic wiki based approach. The observations and findings from the project demonstrate that the discussed approach provides means to address some of the problems related to pre-coordinated formal representation of archaeological knowledge. At the same, the study stresses the importance of a full understanding of the implications of the both old and new systems of knowledge representation. Otherwise the new systems may introduce implicit infrastructural bias comparable to the ones addressed by the novel approach.

Key Words: *Semantic Web, Documentation, Linked Data, Semantic Wikis, Information Management*

Introduction

Digital data archives are a central component of collaborative e-Science infrastructures in scientific and scholarly research (Borgman 2007). Archaeologists have acknowledged the need for standardised but at the same time flexible and contextually adaptable repositories of archaeological data (Lock 2003). Even if there has been a drive to develop universal solutions, the experiences have shown that many of the most successful digital repositories have been in some sense 'local' (e.g. Sure and Studer 2005, Shaw et al. 2009). Both technologies and individual repositories have their origins in certain areas and philosophies of knowledge that makes them particularly successful in those specific areas. For instance, a database of facts is 'local' and useful in contexts based on formal and atomistic modes of knowledge, but at the same time, such a database might be less useful for an individual who leans on the hermeneutic tradition of knowledge (Shaw et al. 2009).

Many of the challenges related to universality have been recognised for some time (Fox and Marchionini 1998). One of the central problems of many universal knowledge systems is the implicitness (Olson and Schlegl 1999) and incoherence of their underlying theories of knowledge. As Blandford et al. (2001) remark, users become easily disoriented with seemingly consistent, but internally inconsistent universal systems, which do not correspond with their own conceptions and experiences of how information should be organised (Huvila 2006). A poor match between digital repositories and their users means that a user does not get relevant information in exchange for his investment of time and effort. This reduces the perceived significance of collections (e.g. Hong et al. 2001) and decreases the experienced quality, scope and relevance of the repositories (Klas et al. 2006).

The present paper discusses findings and observations from an action research study with an aim of developing an integrated semantic digital data archive and collaboration platform for archaeological and archaeology related research. The purpose of the discussed project was to develop both a general framework and a functional prototype of a collaboratory. A special emphasis was placed on fostering participation of all stakeholders of archaeological data and knowledge. In order to enhance the usability of the system throughout the continuum of archaeological information, the management of information processes was based on an evolutionary approach and the use of adaptable, but at the same time semantically rich relations between individual information objects. The platform was based on semantic wiki technology (Kotelnikov et al. 2007) and developed together with archaeologists working for a private Finnish archaeology consultancy, Muuritutkimus ky.

Archaeological Information Process

The complexities of the management of archaeological documentation and information have been acknowledged for a long time (Reilly and Rahtz 1992). The introduction of computers in archaeological work has facilitated the processing of information, and new documentation instruments have enabled archaeologists to capture more precise data than before. The new technologies have not, however, provided any obvious universal solutions to meet the fundamental challenges of managing archaeological information. On the contrary, the introduction of technologies has increased the amount of available data and underlined further the necessity of its effective management (Lock 2003). At the same time, the societal expectations of increased cooperation together with the apparent opportunities and benefits of larger integrated data spaces and e-Science infrastructures have spurred archaeologists to address the challenges relating to the effective management of archaeological information based on the primary documentation data.

The first obvious challenge relates to the complexity of the different types of data and information. A related challenge is that particular types of data and data sources tend to necessitate varying levels and qualities of precision and accuracy. The outcome of the analysis of archaeological data, a qualitative understanding of the past human activities is based on an elaborate brew of qualitative and quantitative data from the literature and archaeological stratum. The body of relevant information consists, for instance, of measurements, scientific analysis results, objects, samples, drawings and photographs (Greene 1998). The different types of materials are seldom comparable *per se* and because of the varying technical formats of representation, they are not easily managed in a single system without multiple parallel data structures and conceptual separation of the data models for the each type of the data (Signore 2009). The codification and representation of different types of archaeological data is an intricate problem (Orlandi 1993). The challenge of precision and accuracy is amplified by the use of computers. Among archaeologists, there seems to be a certain intrinsic preference for high levels of precision in the documentation process, even if an extremely high precision would not be strictly necessary (Huvila 2007). By allowing artificially high levels of precision, computers have been observed to feed this propensity and to support a false sense of accuracy even with technically inaccurate data (Kantner 2000).

The second challenge relates to the incompleteness of available archaeological information and the ambiguous relation of primary data and the consequentially emerging archaeological knowledge (Thomas 2006). In spite of the increased precision of measurements and the growing amount of available data, the documented and documentable data represent always an unknown sample of the original data. Another challenge is that even the theoretical 'complete corpus of original data' consists of miscellaneous remnants of past human and natural processes and is capable of providing only indirect evidence of the past human activity.

Thirdly, the present archaeological information practices lead to major discontinuities in the information process (Greene 1998) and abrupt breaks in the *information continuum* (Oliver 2010). Typical documentation practices of many field projects are based on an arrangement that one person or a small group of people is responsible for mapping and taking measurements using a total station, GPS or comparable equipment. Separate group of people excavate the site, collect finds and document their observations on standardised forms. Further, different people might be responsible for taking photographs and processing finds. Even if the roles would overlap, as they tend to do, the data ends up in separate silos. In many cases only the field director and deputy directors have a direct responsibility and possibility of developing a general idea of the excavation, but in practice, there is often very little time and possibilities for an extended synthesis of the data during the field season. As a result, there is often a major gap between the capture of data and the analysis and reporting of the findings.

The main part of the analysis, organisation and compilation of the data is conducted in the post-excavation phase by individual specialists or teams. Part of the work may by conducted by people who might have not participated the excavation at all or if they have been part of the field team, they might have been on site only during a part of the season. These separate silos are brought together in a report that again tends to reproduce the distinction of the different types of data. The contents of personal field notebooks and the typically short and cautious summary of the investigation process and principal findings incorporated is the final report are the only documents with an explicit aim of synthesising different observations. Finally, if the results are published (most of the data, especially from rescue excavations is never analysed or published properly (Greene 1998, 83-84)), even then the different groups of materials and aspects of excavation are very often reported separately. Thomas (2006, 30) has argued the obvious that these discontinuities are a major problem and essentially lead to a suboptimal documentation of archaeological sites and investigation processes.

The outcome of the discontinuities in the documentation and reporting process is a broken information continuum with an increasing likelihood of the emergence of inferior inferences. The different actors involved in the interpretative process from field archaeologists to the public audience are not properly connected and the different phases of interpretation are not converging as well as they could. The fundamental problem with the broken continuum is that the gaps are highly difficult to bridge in the future. Because archaeological fieldwork is interpretive in all of its stages (Thomas 2006, 30) a discontinuity in one of the stages is difficult to remedy afterwards.

Semantic Wikis

"Wiki" is a form for collaborative writing. The term, a Hawaiian word for "fast" describes the underpinnings of the approach as a technique for easy and effortless editing of web documents using a simplified markup language. A semantic wiki can be a traditional wiki system augmented with semantic web technologies with an aim of improving the possibilities for automatic processing of information (Berners-Lee et al. 2001) stored in the wiki. Another point of departure is to see semantic wikis in a more abstract sense as a combination of wiki and semantic web approaches that incorporate the ideals of easy collaborative information creation, enhanced searchability and machine readability.

Even if the conceptual premises and technical features of individual semantic wiki software packages have substantial differences (Buffa et al. 2008), the most of the systems tend to be based on two alternative approaches. A

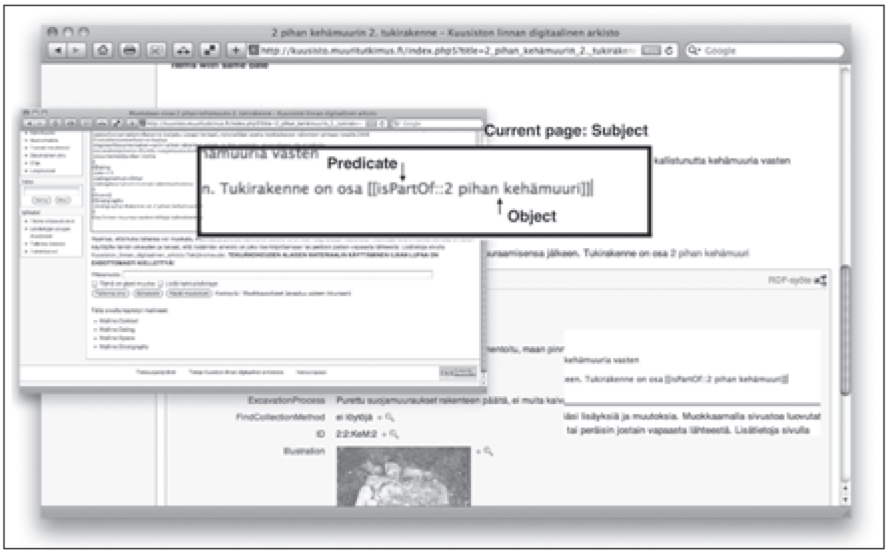

Figure 1. Semantic Mediawiki markup.

part of the systems are semantic repositories augmented with various degrees of wiki functionality. Semantic markup (e.g. RDF triplets) and wiki-data are usually stored separately and the principal emphasis is on an effective management and processing of semantic information. KiWi (Schaffert et al. 2009) and Metaweb (used in Freebase, http://www.freebase.com) are examples of such strongly typed semantic wiki systems. Others, including Sweetwiki (Buffa et al. 2008) and Semantic Mediawiki (http://www.semantic-mediawiki.org) are essentially ordinary wiki systems with added semantic markup functionality. In these systems, the semantic information is typically inserted directly in the wiki-text and the focus of the repositories is in the semantification of free-form information or the combination of strongly and weakly typed data.

Semantic Mediawiki, the platform adopted for the project discussed in the present text, is an extension to Mediawiki, a popular wiki software package used, for instance, to run Wikipedia. Semantic Mediawiki is used on a large number of Mediawiki based web sites in scientific, professional and leisure contexts (examples at http://smw.referata.com/wiki/Special:BrowseData/Sites). The system uses an RDF-based semantic markup. Formal description of a wiki page is done by adding RDF-triplets in the wiki-text using an extension to the standard Mediawiki markup language (Fig. 1). In addition to the basic functionality, a series of extensions to Semantic Mediawiki allow, for instance, the use of a variety non-textual data types and the management and editing of RDF-data using forms (Semantic Forms) and a graphical user interface (http://www.projecthalo.com).

Methods and Layout of the Study

The work with the predecessor of the discussed system started in 2009 after an earlier shorter pilot project that focussed to the development of a site and monument specific integrated information system of two Finnish medieval and early modern castles (Huvila 2008). The earlier focus on castle research was extended in the present project to cover issues related to generic archaeological field documentation.

The project was conducted in cooperation with a Finnish archaeological consultancy Muuritutkimus ky. The company has been interested in developing an integrated system for the management of excavation and conservation data. The work was conducted by the author in collaboration with two archaeologists employed by the consultancy. One of the informants was working as a field director and the CEO of the company and the other one as a researcher with a special competence in archaeological field documentation and measurement.

Features and requirements were discussed in meetings, implemented and tested in field trials in actual locations with real data. Part of the software development and the evaluation of appropriate laptop and tablet-pc computers was conducted directly in the field. The meetings were started with a round up of new ideas and thoughts, which had emerged after the previous meeting, and were continued with a presentation of the new features of the system, documentation of spontaneous comments made by the archaeologists and a round of specific questions and issues relating to the development work and requiring a direct consultation of the archaeologists. Finally, the next stages of the development work were agreed upon 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 the documentation of archaeological features on an actual site (about 30 minutes). A short informal meeting and approximately two additional development-trial cycles per day followed the trial.

The development process was based on action research (Greenwood and Levin 2000; Kemmis and McTaggart 2000) as a principal methodology of investigation. According to the action research approach, the problems (research question) addressed during the research were 1) the misfit of formal data and qualitative interpretations, 2) the difficulty of representing the complexity of archaeological data, and 3) the discontinuity of archaeological information process from the field to the archive and publication. After a brief initial analysis of the situation, the problems were decided to be addressed by building a functioning prototype of a digital collaboratory with an appropriate basic functionality and comprehensive possibilities to develop both the technical and content aspects together with the actual users of the system.

A semantic wiki-based technological approach was chosen because of its anticipated capability to meet the structural and work process related challenges posed by the two critical success factors of archaeological information work, fit and sustainability (Huvila 2007). The principal strength of wiki based approaches is their focus on collaboration and the use of automatic versioning systems designed to keep all changes recorded and to allow infinite rollbacks to the earlier versions of the data. The approach promised also to deliver a necessary degree of structural and representational flexibility and formality whenever it was needed. It provided also means to work with multiple media forms and offered necessary adaptability for making the system relatively sustainable even in the future. Essentially, the approach was a compromise of flexibility and formality that was estimated to accommodate the most of the present and future requirements of the major stakeholders of the data (Huvila 2006) and to provide enough structure for maintaining a practicable degree of searchability and findability of the data (e.g. Sure and 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 choice of the semantic wiki approach was motivated by positive experiences from an earlier digital archive project (Huvila 2008). Finally, the system allowed archaeologists to work in a single system from the first observations to analysis, interpretation, reporting and finally, publication of the data with a capability of flexible linking of the data to other web-based scholarly and general resources.

The documentation framework was based on a model 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, type or a part of any of them. Therefore, an object (wiki page) can be a resource or a description of a resource (Kotelnikov et al. 2007). Objects can be or may represent archaeological objects, sites, features, structures, documents, books or any other entities 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 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). In addition to typing, relationships between different objects can be expressed by using CIDOC-CRM Properties (e.g. Crofts et al. 2007), and when needed, by using any 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 the colour of objects expressed using a relevant system of representation (such as Munsell or Pantone for different kinds of objects) documented within the repository. Lastly, objects can belong to *categories* (e.g. Kustö Castle) and dynamic categories called *concepts* (e.g. "Iron objects from Kustö Castle" i.e. all objects belonging to category Kustö Castle and having a material property referring to iron).

Semantic Wikis and Archaeological Documentation Process

The development process showed several benefits and advantages of the assumed approach and the chosen technical platform. The general observation was that the semantic wiki based approach provided a workable framework for addressing the major problems identified in the beginning of the action research process. The combination of semantic web and freeform description removed some of the major points of criticism of the relational database systems made by the proponents of the semantic web approaches (e.g. Barchesi 2004; Ross 2003) and the constituent critique of the archaeological application of semantic web itself (Veltman 2004). The main benefits identified during the process were that the approach permits:

1. Combination of formal semantics and freeform description,

2. Evolutionary enrichment of documentation throughout the *information continuum,*

3. Documentation without a monolithic conceptual model,

4. Enriching documentation with strong semantics as the conceptual understanding of the research object develops during the research project,

5. Supporting archaeological reasoning, management and dissemination of the information throughout the research process and beyond without a need to move data from one system to another for the purposes of publication and presentation.

First, a central advantage of the assumed semantic wiki based approach was the anticipated possibility to flexibly combine formal and informal data. In Semantic Mediawiki, the semantic markup is typed within ordinary wiki-text and the system allows flexible combination of formal data and free text, images, drawings, external files and, for instance, links to external resources (as illustrated in figure 1) Strongly typed descriptions and measurements can be easily augmented with preliminary remarks on their significance, possible interpretations and links to earlier data from other excavations and the literature.

The second identified benefit was that the approach permits an evolutionary enrichment of documentation during the entire information continuum that is similar to the evolution of texts in ordinary wiki systems, for instance, in Wikipedia. The first iterations of data were entered into the system in field and the same data could be elaborated throughout the analysis, interpretation and reporting process. Finally, the data could be linked to external systems. One example of this was tested successfully in a related project that aimed at integrating the same data resources to a popular presentation of an archaeological site in the virtual world of Second Life (www.secondlife.com). It was observed that the semantic wiki approach contributed to the process of gradual semantification of the data by facilitating the involvement of both more and less technically literate archaeologists (a functionality urged by Isaksen et al. 2010). The continuous linking and simultaneous elaboration of data was made possible by turning the documented entities into linked data already in the field. In practice, every documented entity had a persistent URI from the moment the entity was identified in the field.

Third, the approach is not based on a single monolithic data model that would have forced archaeologists to accommodate their observations to follow a rigid predetermined framework as in a typical database system based on a strongly typed model. Template system allows the creation and combination of data models and descriptions within the system. In addition, it was possible to make references to external ontologies, but at the same time, to give each archaeologist a freedom of adopting the most appropriate descriptors according to their own experience and expertise.

Fourth, the documentation could begin with minimal information without a need to document more or more accurately that was possible by the time of the observation. In course of the fieldwork and later on, the first observations could be elaborated with measurements, new photographs, details, literature, and alternative interpretations.

Finally, the trials of linking the semantic wiki to a Second Life based virtual presentation and external web sites showed how the same system could be used to manage popular information and to function as a data source for meaningful external applications. The ease of linking data between independent systems facilitates also the management and interlinking of data across multiple sites and multiple research teams.

In spite of the general satisfaction with the results of the discussed version of the system, a number of challenges were identified during the action research cycles. A mainly technical challenge relates to the adopted technical platform (Semantic Mediawiki) and its encyclopaedic design philosophy, which make it less optimal for the practical archaeological field documentation work. Other challenges were of more fundamentally theoretical kind:

1. The number and meaning of links.
2. Contradictory claims in semantics.
3. Manageability.
4. What a wiki page is actually representing?
5. URL-based naming of pages (in SMW) i.e. how to create linked data during the documentation process when the relations of entities are evolving (in contrast to earlier projects working with semantification of pre-existing repositories (e.g. Tudhope et al. 2011)).
6. How to tweak wiki-workflow to match even closer the archaeological workflow.
7. The fundamental discrepancy between the "semantics" in archaeological sense and "semantics" in a "semantic" web sense.

First, there has to be a consensus of linking and the meaning of the links in the system. In a networked system, links are necessary for indicating relations between entities. Semantic linking of two distinct entities incorporates a making of a statement about the meaning of that particular link. In order to incorporate certain viewpoints (as preferred orders of data in the form of particular types of links) there has to be a consensus of what links are supposed to be created in the first place and what is the preferential order of the different types of links. In the documentation of archaeological stratigraphy it is necessary to use typed links to adjacent strata to indicate assumptions about stratigraphical sequences. In contrast to the stratigraphical relations, the documentation of the most of the other relations between archaeological entities is not steered by equally strong conventions. Further, it is not similarly obvious that all types of objects or materials, or their assumed places of origins, are equally important to link to each other. The support for arbitrary linking has many advantages, but in order to improve the searchability of the documentation, it is necessary to maintain an elementary consensus of common descriptors and linking practices in terms of what everyone are supposed to link and where.

Secondly, because the approach allows an arbitrary number of parallel and conflicting interpretations of same archaeological entities, there has to be an agreement of how the conflicts are documented and how it is possible to choose and indicate preferences between competing renderings of the same evidence. Possible solutions that were implemented in course of the project were based on signing interpretations, public voting and implementation of a specific type of 'consensual and alternative interpretation' relations.

The third challenge identified with the discussed approach relates to manageability of the system. In comparison to traditional strongly typed database applications and semantic systems, the openness and flexibility of a semantic wiki puts pressure on the post-coordination of data and continuous evaluation and integration of the content. In practice, it seems that the field director or a specific archaeological information manager needs to supervise and facilitate the documentation work throughout the process. Even if the approach is likely to increase the need for explicit coordination and interpretation in field, it reduces the time required for the development of documentation instruments (e.g. forms) before field season and is likely to reduce the frequency of problems related to a misfit of the pre-coordinated instruments and the actual characteristics of the investigated site. Further, the closer in-field coordination reduces the time needed for post-excavation management and re-organisation of documentation data, encourages discussion within the field team, and potentially de-centralises and shifts interpretation of the site from the post-excavation phase to the field.

Fourthly, a wiki system raises a conceptual challenge of determining what a page (the central metaphor of a typical wiki system) can represent. The Wiki object model (WOM) based semantic wikis (e.g. SeedWiki, Buffa et al. 2008) have a specific ontology for determining the semantics of different components *within* the system. In the absence of a specific WOM, it is necessary to use other means to define and communicate the meaning of a page (similarly to the meaning of links and linking as discussed earlier) to all users of the system. A page can itself *be* an entity of data, it can be used as a representation of a physical entity (e.g. archaeological object or stratum) or it can contain data about other digital entities (e.g. digital photographs). In addition, a page can represent an abstract entity such as 'observation' (an event that connects simultaneously or quasi-simultaneously identified and documented entities) or 'interpretation' of a group of entities.

The fifth challenge relates to the naming of the pages, and consequently, to the form of the persistent URIs used to create linked data (e.g. Tudhope et al. 2011). The earlier work on archaeological semantics has focused on the semantification of existing collections of data with a pre-existing structure. Assigning persistent URIs on the fly during the documentation work means that the naming decisions cannot be based on an existing general view of the site and a consensual structure of the relations between different entities. In the present case, even the identity (e.g. whether an entity is a stone structure or a group of large stone objects) of the observed entities is subject to change during the later phases of the excavation and analysis process. On the other hand, many wikis, including Semantic Mediawiki, and their markup are based on a heavy use of meaningful page names. Stripping the URIs of any human readable information about the entity and its type complicates the work with the pages and entities. One possibility is to resort to a frequent refactoring of the names of the entities. Refactoring is not, however, a problem-free solution. The changing URIs makes the system difficult to understand for its users and the number of necessary redirects from previous URIs to the new ones obfuscates the structure of the data.

The sixth challenge relates to the dissimilarities of the workflow in wiki-based systems and in archaeological fieldwork. Even if it is basically desirable to configure the documentation system to follow the typical procedures of the archaeological fieldwork, it is useful to take a critical stance to the prevailing documentation practices. Documentation system can be also seen as an agent for positive change. As observed by Lock (2003, xiii) and Orlandi (1993), the information work of archaeologists is not profoundly 'computer minded'. It is closely configured to accommodate the present digital and analogue systems of organising knowledge and information. The major challenge of developing new approaches is the question of how to develop information work and digital systems together in order to establish an appropriate match of tools and practices. The observations made during the present project suggest that the typical approach assumed by information systems, (e.g. wikis and database management systems) to start a workflow with the creation of a document or a database record is different from the actual workflow of archaeological documentation work. In the present study, instead of starting with the identification and documentation of individual elements in the stratum, the archaeologists tended to begin with a general glance of the entire stratum. An iterative process of elaborative classification and identification of smaller entities followed the first overview. The observation progresses from the general to the specific aspects and parts of the observed area and only after a proper identification of potentially pertinent features, it proceeds to their naming. To accommodate this reverse (from an information system point of view) strategy of observation and naming, a specific entity called 'Observation' was implemented in the documentation system to represent the initial observation from which the individual entities are derived.

The last challenge identified in the project was that the proposed approach is not a solution to the problem of discrepancy of formal and informal data in archaeology. Archaeological reasoning incorporates a wide range of different types of semantics but the idea of semantics in the semantic wiki context is based on the same formal model of semantics than the 'semantics' of the semantic web. The semantic wiki based approach helps to accommodate competing systems of semantics, but essentially, it offers tools to explicitly address only the formal type of semantics. As Almeida et al. (2011) argue in the context of the semantic web as a whole, it would be important to admit the limits of the semantic web technologies in being capable of representing linguistic, philosophical and, for instance, different types of, often hermeneutic, forms of archaeological semantics. In an ideal case, a semantic wiki approach should attempt to capture and provide tools for managing also these other types of semantics in addition to the formal one.

In practice, however, in addition to the problems related to the understanding of 'semantics' in the major semantic web technologies, it would be also important to acknowledge that the wiki approach itself imposes a certain additional idea of meaning that is based on a network of links between 'pages'. Understanding this category of representation and its outcomes are equally important as the understanding of the consequences of the semantic underpinnings of more traditional computer systems emphasised by Orlandi (2004). Like all infrastructures in e-Science and beyond, the infrastructure itself presents a series of affordances and constraints to the evolution of the scientific and scholarly processes. Besides criticising the limitations of old systems, it is equally important to acknowledge and be aware of the new constraints and bias.

Conclusions

The findings of the present study demonstrate that the semantic wiki based approach provides means to address some of the problems related to pre-coordinated formal representation of archaeological knowledge. It is, however, important to acknowledge that even a semantic wiki presents a pre-coordinated bias to the documentation process by steering the ways how observations are documented and structured. Semantic wiki makes it possible to combine formal semantics with freeform description and allows evolutionary enrichment of documentation throughout the continuum of archaeological information that starts already before the actual documentation takes place and continues from fieldwork to analysis, research, publication and applications in other scholarly fields and in the public presentation of archaeological information. Unlike in traditional systems, the documentation does not need to start with a monolithic predetermined conceptual model that forces archaeologists to accommodate their findings into a rigid interpretational framework. At the same time, a semantic wiki provides tools for enriching documentation with strong semantics as the conceptual understanding of the research object develops during the research project. Finally, an advantage of the approach is that the system may be used to support archaeological reasoning, management and dissemination of the information throughout the research process and beyond without a need to move data from one system to another for the purposes of publication and presentation.

Acknowledgements

I would like to express my gratitude to my colleagues working at the project, Ph.D., adjunct professor Kari Uotila (Muuritutkimus ky) and MA, archaeologist Markus Kivistö for a fruitful cooperation in the presented project, and for Finnish Board of Education for financing the project "Archaeology in the virtual worlds".

Bibliography

Almeida, M., Souza, R., and Fonseca, F. 2011. "Semantics in the semantic web: A critical evaluation." *Knowledge Organization* 38 (3):187–203.

Barchesi, L., Ceccarelli, C. 2004. "Linguaggi dichiarativi per la ricerca archeologica." *Archeologia e calcolatori* 15:95–113.

Berners-Lee, T., Hendler, J., and Lassila, O. 2001. "The semantic web." *Scientific American* 284 (5):28–37.

Blandford, A., Stelmaszewska, H., and 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*, 179–188. New York: ACM Press.

Borgman, C. L. 2007. *Scholarship in the Digital Age: Information, Infrastructure, and the Internet.* Cambridge: MIT Press.

Buffa, M., Gandon, F., Ereteo, G., Sander, P., and Faron, C. 2008. "Sweetwiki: A semantic wiki." *Web Semantics: Science, Services and Agents on the World Wide Web* 6 (1):84 – 97.

Crofts, N., Doerr, M., Gill, T., Stead, S., and Stiff, M. eds. 2007. *Definition of the CIDOC conceptual reference model*. ICOM/CIDOC Documentation Standards Group, Continued by the CIDOC CRM Special Interest Group. http://cidoc.ics.forth.gr/docs/cidoc_crm_version_4.2.2.pdf.

Fox, E. A., and Marchionini, G. 1998. "Toward a worldwide digital library." *Communications of the ACM* 41 (4):29–32.

Greene, K. 1998. *Archaeology: An introduction*. London: Routledge.

Greenwood, D. J., and Levin, M. 2000. "Reconstructing the relationships between universities and society through action research." In *Handbook of qualitative research*, edited by N. K. Denzin, and Y. S. Lincoln, 85–106. Thousand Oaks: Sage.

Hong, W., Thong, J. Y. L., Wong, W.-M., and Tam, K. Y. 2001. "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*." PhD diss., Åbo Akademi University. Åbo: Åbo Akademi University Press.

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*. University of Osijek and Rutgers University.

Huvila, I. 2008. "Participatory archive: towards decentralised curation, radical user orientation and broader contextualisation of records management." *Archival Science* 8 (1):15–36.

Isaksen, L., Martinez, K., Gibbins, N., Earl, G., and Keay S. 2010. "Interoperate with whom? Formality, archaeology and the semantic web." In *Web Science Conference 2010, April 26-27, 2010, Raleigh, NC.*

Kantner, J. 2000. "Realism vs. reality: creating virtual reconstructions of prehistoric architecture." In *Virtual Reality in Archaeology. Proceedings of the Computer Applications and Quantitative Methods in Archaeology (CAA) 1998, Barcelona*, edited by J. A. Barceló, M. Forte, and D. H. Sanders, 47–56. Oxford: Archaeopress.

Kemmis, S., and McTaggart, R. 2000. "Participatory action research." In *Handbook of qualitative research*, edited by N. K. Denzin and Y. S. Lincoln, 567–605. Thousand Oaks: Sage.

Klas, C.-P., Fuhr, N., Kriewel, S., Albrechtsen, H., Tsakonas, G., Kapidakis, S., Papatheodorou, C., Hansen, P., Kovacs, L., Micsik, A., and Jacob, E. 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*, 308–309. New York, NY: ACM Press.

Kotelnikov, M., Polonsky, A., Kiesel, M., Völkel, M., Haller, H., Sogrin, M., Lannerö, P., and Davis B. 2007. "Interactive semantic wikis. Deliverable D1.1, NEPOMUK Consortium, Kaiserslautern." http://nepomuk.semanticdesktop.org/xwiki/bin / download/Main1/D1%2D1/D1.1_v10_NEPOMUK_Interactive_Semantic_Wikis.pdf.

Krötzsch, M. 2007. *Special:types*. Semantic Mediawiki. http://semantic-mediawiki.org/index.php/Special:Types.

Lock, G. 2003. *Using computers in archaeology: Towards virtual pasts*. London: Routledge.

Oliver, G. 2010. "Transcending silos, developing synergies: libraries and archives." In *Information research 15 (4). special supplement: Proceedings of the Seventh International Conference on Conceptions of Library and Information Science – Unity in diversity – Part 2*.

Olson, H. A., and Schlegl., R. 1999. "Bias in subject access: A content analysis of the critical literature." In *Information Science: Where Has It Been, Where Is It Going? Proceedings of the 27th Annual Conference of the Canadian Association for Information Science*, edited by J. Turner, 236–247. Montreal: CAIS/ ACSI.

Orlandi, T. 1993. "Sulla codifica delle fonti archeologiche." *Archeologia e calcolatori* 4:27–38.

Orlandi, T. 1999. "Multimedialità e archeologia." *Archeologia e calcolatori* 10:145–157.

Orlandi, T. 2004. "Archeologia teorica e informatica archeologica. Un rapporto difficile." *Archeologia e calcolatori* 15:41–50.

Reilly, P., and Rahtz, S. 1992. *Archaeology and the information age*. London /New York: Routledge.

Ross, S. 2003. "Position paper on integrity and authenticity of digital cultural heritage objects." In *Towards a semantic web for heritage resources*, edited by DigiCULT, vol. 3 of *DigiCULT Thematic Issue*, 7–11. Salzburg: DigiCULT.

Schaffert, S., Eder, J., Grünwald, S., Kurz, T., and Radulescu, M. 2009. "Kiwi – a platform for semantic social software (demonstration)." In *The semantic web: Research and applications*. *Lecture Notes in Computer Science* 5554, edited by L. Aroyo, P. Traverso, F. Ciravegna, P. Cimiano, T. Heath, E. Hyvönen, R. Mizoguchi, E. Oren, M. Sabou, and E. Simperl, 888–892. Berlin/Heidelberg: Springer.

Shaw, R., Corns, A., and McAuley, J. 2009. "Archiving archaeological spatial data: Standards and metadata." In *37th Annual Computer Applications and Quantitative Methods in Archaeology (CAA) Conference, Williamsburg, Virginia*.

Signore, O. 2009. "Representing knowledge in archaeology: from cataloguing cards to semantic web." *Archeologia e Calcolatori* 20:111–128.

Sure, Y., and Studer, R. 2005. "Semantic Web technologies for digital libraries." *Library Management* 26 (4/5):190–195.

Thomas, J. 2006. "The great dark book: Archaeology, experience, and interpretation." In *A companion to archaeology*, edited by J. Bintliff, 21–36. Malden/ Oxford: Blackwell Publishing.

Tudhope, D., Binding, C., Jeffrey, S., May, K., and Vlachidis, A. 2011. "A STELLAR role for knowledge organization systems in digital archaeology." *Bulletin of ASIS&T* 37 (4):15–18.

Veltman, K. H. 2004. "Towards a semantic web for culture." *Journal of Digital Information* 4 (4).