CHNT 18 Proceedings

A case study in archaeological documentation with ontological modeling

Luciana BORDONI[1] | Michela COSTANTINI[2] | Alessandra CURCIO[2]

[1] Enea-Utict | [2] Università di Roma "La Sapienza"

Abstract: Documentation of archaeological and cultural heritage sites is at the heart of the archaeological process and an important component in cultural heritage research and presentation; it is an essential step without which interpretation and analysis are not possible. It is what makes archaeology and cultural heritage "scientific". Archaeological knowledge is an important part of human knowledge and it is necessary in many applications, such as knowledge-based systems for archaeology. A formal, clear and declarative description for communicating, sharing and reusing archaeological knowledge among humans and software entities is of crucial importance. Formal ontologies have been viewed as a promising means to tackle this problem. An ontology is an explicit formal declaration of how to represent object concepts and other classes assumed to exist in some area of interest (a domain) and the relationships between them. In this sense an ontology is a specification of a conceptualization. A domain-specific ontology of archaeology is an explicit and formal specification of a shared conceptualization of archaeology; it captures the semantics of archaeological knowledge. In this work we present the use of an ontological approach to model and make available archaeological documentation of the Roman City of Ardea. The methodology used is summarized in the following steps: acquisition of the knowledge domain and organization of the ontological model. This ontological model consists of three major parts: archaeological categories, their relationships and axioms. It has been implemented in Protègè, the open source platform developed by Standford Center for Biomedical Informatics Research at School of Medicine of Stanford University. It allows to organize and catalogue information, it gives also the possibility to question the ontology using a query editor. In this paper we will present, on the basis of a case study, how ontology-based approach can be used to bring benefits to the archaeological documentation.

Keywords: Archaeology, knowledge acquisition, ontology, Protègè

Introduction

Research on the Semantic Web in cultural heritage is of primary importance today, especially if it is carried out with a transversal and interdisciplinary perspective that combines the more technological elements with domain issues and the evaluation of the impact that these solutions should have on the activities and practices of different communities, from professionals to the general public. The problem of representation of cultural heritage data, information and knowledge in contexts of semantic integration of heterogeneous metadata schemata is a fundamental concern, and it is also the area in which notable results of general value have been obtained (AROYO et al. 2007). Since the beginning of the 2000s, Semantic Web technologies and their potentials for the integration and exploitation of digital cultural heritage information have received increasing attention, and today they represent an exciting and dynamic field of interdisciplinary research (DigiCULT 2003). It is a common opinion that the diversity and epistemological richness of cultural heritage provides an excellent field for the deployment and experimentation of Semantic Web-based systems.

Semantic Web applications in the archaeological domain are still limited in number if compared to the experiences carried out in the broader scenario of cultural heritage. Nonetheless, several potential benefits are generally acknowledged concerning the Semantic Web and archaeological research (STAR 2010). Archaeological data are of a complex nature, and their inherent uncertainty makes the definition of models of representation an extremely difficult task. This characteristic shows its influence in a number of aspects ranging from the representation itself to the modalities through which data are combined, retrieved, and displayed. The creation of models of data, information and knowledge representation constitutes a traditional concern of cultural heritage documentation: classification systems, list of terms, catalogs, thesauri, and more generally schemata of domain concepts and their relationships have been created and used long before the introduction of computers in the sector; better, they have often represented crucial aspects in the theoretical debate. Therefore, it is natural that in the context of Semantic Web applications for cultural heritage, the implementation of these resources in machine processable formats constitutes a key area of research and experimentation (May 2006). On the other hand, the possibilities connected to the creation of formalized frameworks and domain models of knowledge using knowledge representation languages (such as the ones connected to the Semantic Web) offer the possibility to improve research on the documentation and analysis methods of cultural heritage (REMUNA 2006). In the following we present the use of an ontological approach to modeling and make available archaeological documentation for a case study: the Roman City of Ardea.

The paper is organised as follows. After this introduction, Section 2 provides the methodology used. Section 3 presents the ontological model and how it has been implemented in Protègè. Conclusions and perspectives are made in section 4.

Methodological approach

Historical Background

The City of Ardea, included in the map IGM 158 NW IV, stands on a rocky outcrop, in front of western offshoots of the Alban Hills and dominates the surrounding area. Its territory is included in the Agro Romano and extends south of Pomezia, with the view to the sides of the Castelli Romani and the Tyrrhenian Sea, bordered on the south by the cities of Anzio and Aprilia (fig.1).

The territory of Ardea was populated in the Paleolithic and have been found tombs of the Copper Age, dating from the early second millennium BC. In the Iron Age the settlement was made up of three villages of huts on the three plateaus on which still stands the city (Civitavecchia, Acropolis, and Casalazzara), where they were found traces of the holes of the huts and a necropolis at "Monte della Noce", on the plateau of Civitavecchia, with graves childhood and a princely tomb female eighth century BC. Ardea, was born as agglomeration agricultural and developed through trade, favored by the city's location and by the inclusion of a port-channel at the mouth of the river Interlocking (Castrum Inui). From eighth to sixth centuries BC was one of the most important centers of southern Lazio, with a rich crafts.

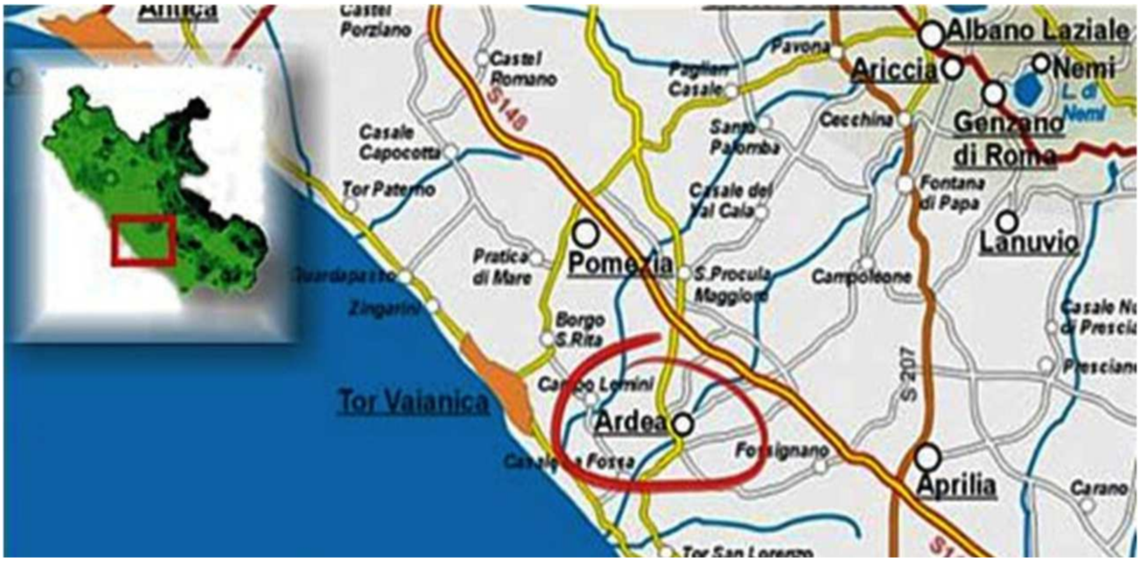

Fig.1 – Localization of Ardea

The city reached its period of maximum development in the seventh century BC. It was renowned for the production of weapons and ornaments. On several occasions the people of Ardea were allies or enemies of Rome, in the context of the events of the Latin League. During the fifth century BC the city was dominated by the struggle against the Volscians and in the fourth century the Gauls, after having sacked Rome, turned against Ardea and the besieged, without success; were indeed the Ardeati, led by Furio Camillo, in exile in the city, which rejected the siege and marched to Rome and they freed from the Gallic. In the second Roman-Carthaginian treaty of 348 BC, Ardea is again named among the allied cities of the Romans. At this time dates the rebuilding of the city walls. During the Second Punic War, Ardea was one of the twelve colonies who refused military aid to the Romans. Between the third and second centuries BC Ardea declined, due to the economic crisis of the centers of Lazio, whose resources were dried up in the Punic wars and the subsequent war against the Samnites. In the Imperial Age few remnants of settlement dating back to the fifth century survived, while the villas were built along the coast. The city of Ardea began to grow only from the ninth century, thanks to the progressive depopulation of the *domus cultae* (small agricultural centers founded by the Popes in the countryside for the cultivation and reclamation) and the needs of defense against the Saracens. In 1130 the antipope Anacletus II attributed the *civitas* Ardeae to the Benedictine monks of the Basilica of St. Paul Outside the Walls in Rome. Then the feudal control of the city was the subject of bitter disputes between the noble families of Rome. In 1419 Pope Martin V gave the city to the Colonna's family. The estate then passed to other papal families: from Borgia returned to the Colonna, until in 1564 it was sold to Giuliano Cesarini. In this period the city was a farming village, following the fortunes of the families who ruled from time to time. In 1816, the city became a part of Genzano of Rome. From 1932 the area has been the subject of work of land reclamation which followed the repopulation of the village and the surrounding countryside. The village was re-founded and became part of the municipality of Pomezia until its inception in 1970 when Ardea becomes Municipality.

Methodological steps

This case study was suggested by the historical and archaeological importance of the area combined with the ability to provide updated and useful data for tourism in defense of a territory subjected to an ever expanding urbanization. All the territory, which was originally based on an agricultural economy, is part of the "outer belt of Rome" and has undergone major changes due to private construction activity but also due to the construction of warehouses or to the presence of huge factories.

The first step focused on the study of the history of Ardea and topographical mapping. We looked at the cards made by IGM in 1957 in scale 1: 25,000 and at the Regional Technical Map of Lazio, on a larger scale (1:10,000). Each known archaeological tab, was inserted into the card catalog with information about the location, the type of discovery, the dating, and the references. All collected data were systematized through a filing which allowed us to structure the individual objects within the more articulated concept of "context" and to connect these contexts to the study of the buildings. This filing, far from being a mere container of knowledge, is found to be an essential tool to restore the network connections information necessary to reproduce the scope of membership of each well archaeological investigation. Knowing well what the images are an essential support to the word written to allow access to specific content by a public non-specialist, it was decided also to accompany the card tables photographic contexts in which they are represented by grouping the materials by site of origin and in the second instance by inventory number, so to facilitate future computerization of data. All data were compared with aero photos.

The second step concerned the process to model and make available archaeological documentation of Ardea. In particular, it was necessary to "simplify" the archaeological record in order to transform the data into simple concepts so that their performances could be arranged so as to be "understood" and processed by an electronic computer. An essential tool through which it was possible to gradually fill this distance has been the adoption of the method of representation of knowledge through ontologies more or less structured and/or complex. The notion of ontology is used with different meanings, and has a long history in philosophy and a newer part of Artificial Intelligence (AI) and an even more recent context of Information Systems (Guarino 1998). In a philosophical sense, it refers to an ontology as a "system of categories that give an account of a particular vision of the world" and, as such, this system doesn't depend on the particular language used to describe it. The prevalent use of the notion of ontology in Artificial Intelligence is to refer to "engineering artifacts" consisting of a specific vocabulary, used to describe a given reality, to which is added a set of explicit assumptions regarding the meaning denoted by the terms of vocabulary. Assumptions that, in general, are expressed in the form of theories logic, where the vocabulary terms are names of unary predicates or binary logic system (first order) chosen. Finally, ontology is defined as a structure consisting of concepts, properties of concepts, relationships and constraints between the concepts. Ontologies are defined independently of the current data reflect an understanding, common semantic domain of discourse and can be used to share and exchange information between sources also heterogeneous (Uschold & Gruninger 1996). Ultimately, they are declarative specifications of the concepts of base in a domain. In the simplest case an ontology describes a hierarchy of concepts related to reports of subsumption (class - subclass), while in the most of cases sophisticated adding appropriate axioms for expressing relationships between concepts and/or to constrain their interpretations . Ontologies can be represented with directed graphs and labeled, where the nodes correspond to the concepts and bows to the roles and/or relationships.

The semantic structuring obtainable with ontologies differs from the composition of superficial information and formatting which is managed. For example, relational databases or semi-structured storage systems, such as XML -DB. In fact, with the database of the entire semantic content is virtually captured in the application logic , while you are with ontologies able to provide a specific objective of the domain of discourse that is a "consensual agreement" on the concepts and relationships that characterize the way which is expressed in a given knowledge domain. In other words, you can separate the representation of the domain from the processing logic of the same.

The ontological model and its implementation

Protégé (http://protege.stanford.edu/) is an open source tool, developed at Stanford University, which provides an interactive graphical editor to design ontologies and for acquisition knowledge; it makes possible to create and edit reusable ontologies and it is a valuable aid for system developers and domain experts to create knowledge-based systems (NOY et al. 2000). Protégé ontologies can be exported in several formats: RDF (S), XML Schema and OWL; it is based on Java and it is extensible through many APIs and plug-ins. Also it has numerous plug-and-play environments that enable rapid application development. The Protégé platform provides two ways to create ontologies: the Protégé-Frames editor, which allows to build and populate ontologies that are based on "frame", according to the OKBC protocol, and the Protégé-OWL editor, which allows to build ontologies for the Semantic Web, in particular according to the OWL language. We used the second method and we also selected OWL format files, so the ontology contains only the empty class “owl:Thing” that is the set that contains all individuals and it is for this reason that all classes are subclasses of “owl: Thing”. OWL ontology includes descriptions of classes, properties and their instances. We defined classes that represent concepts in the domain taken into consideration, with slots that represented the properties and relations between the concepts. In the ontological model a set of instances constitute the base of knowledge; a class also can have subclasses that represent concepts more specific. Protégé allows to define classes of an ontology and their hierarchy by displaying a tree to represent it. We can change this ontological tree, accessing information quickly. We created the tree and we decided the relationship between the concepts that are part of it. It is managed through a panel that is defined “Taxonomy ontology” and which displays the attributes associated with the selected concept tree. The tree was structured in two thematic macro-areas (classes) “Beni_culturali_mobili” and “Beni_culturali_immobili” that have different sub-classes; it has not only allowed a greater connection between the parts of which make up the ontology, but also to develop linkages between conceptual levels. We set restrictions on subclasses, so that, for example, the subclass "temple" belongs only to the super-class "Bene_immobile"; then we created the relationships between the different categories through specific "properties", the key-elements of the ontology. We defined the “object property”, relations between individuals belonging to classes (generally between classes) and the “data type property” that are the relations that link the attributes of individuals to types allowed. Through them we can have a perfect circularity of the information, so from any point we decide to start the search; it is always possible to reach all the others. This feature is the main novelty of this research because it amplifies the chance to question ontology and to compare different types of information. After the tree-structured ontology, any archaeological resource was filed and entered into the software Protégé. The purpose of this ontology is to allow to search archaeological objects through relations e.g. their materials, their dating, their location, etc. As shown in the figure 2, the user can query "cinerario_fittile" and obtain information for all the archaeological evidences found in "località_Salzara", next to "cinerario_fittile". In this application domain, relations between objects, materials, locations, etc. are very important because they constitute large part of the context, which gives meaning to the study. Also it is the opportunity to interact with a web GIS with all of the standard GIS functionality to browse and query the maps, as the zoom functions, the general reference map, the queries of features or attributes. Moreover, the presence of the on-line guide makes the software available as a service to a user facing differentiated: archaeologists and who usually operates in this context but also it is aimed at tourists and other passionate.

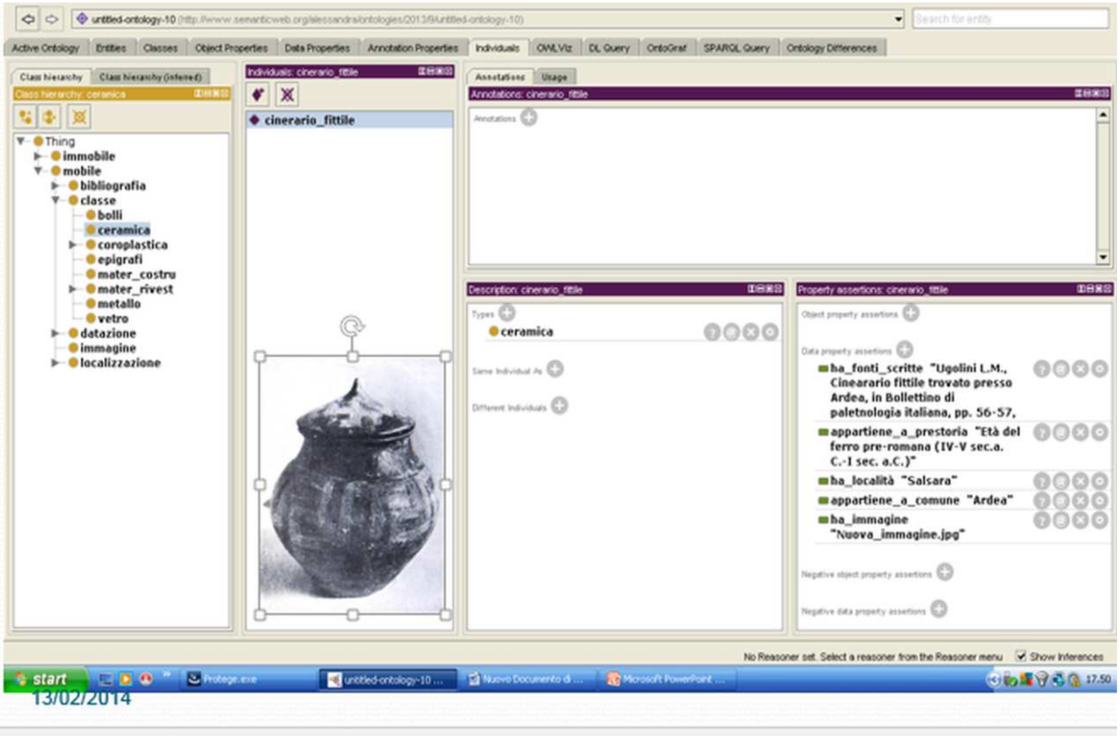

Fig. 2 – A form in Protégé

Conclusions and perspectives

There is no doubt today that the trajectories along which the Web is evolving and the richness and heterogeneity of the Archaeology domain make interdisciplinary research combining the two areas an exciting challenge, as well as a fertile ground for original contributions to the study of the Information Society. The Semantic Web scenario together with the new approaches in the deployment of Web applications by combining existing services and data, emerged as the technological reference frameworks for the research in the context of cultural heritage. The analysis of real applications and systems is of primary relevance, since it dramatically improves the evaluation of the feasibility of the new approaches. For this reason, a key direction and contribution of our research has been the joint discussion of theoretical and methodological aspects, together with the discussion of more practical issues at the application level. The goal of this application is to shown as ontology mapping is the process of reconstructing data semantics of a dataset to understand the structural aspects of the data. These discussions have greatly benefited from the design of a well-delimited and real case study and the development of a prototype application. A central aspect this research investigated has been the issue of representation, in terms of the approaches for modelling domain data, information and knowledge. This activity confirmed that the role of domain experts (archaeologists) in defining appropriate mappings of legacy metadata schemata to the model is fundamental, but there are still a few contributions that provide a description and a discussion of this activity in all its complexity. Nevertheless, this approach is fundamental in order to make it possible for cultural heritage professionals to evaluate, understand and effectively use the model. From a more general perspective, the selection and integration of other relevant data sources will constitute a central activity for future development. However an interdisciplinary approach is today more than necessary since domain and technological knowledge is the key combination for designing and deploying effective solutions.

References

AROYO, L., HYVÖNEN, E., AND VAN OSSENBRUGGEN, J. EDITORS (2007). Proceedings of Workshop 9: Cultural Heritage on the Semantic Web, Busan, Korea. 6th International Semantic Web Conference and 2nd Asian Semantic Web Conference. Accessible at: http://www.cs.vu.nl/~laroyo/CH-SW/ISWC-wp9-proceedings.pdf

AUER, P., CAVALLINI, F., GIFFI, E., LATTANZI, M. (1998). Strutturazione dei dati delle schede di catalogo, Normativa per la strutturazione e il trasferimento dei dati, ICCD.

CANTONE, F. (2005). "Shared Technologies in archeologia: nuove prospettive di gestione e condivisione di dati in rete", Archeologia e Calcolatori 16: 271-290.

CRESCENZI, L. (1989). Scavi ad Ardea, in Archeologia laziale, (Quinto Incontro di studio del Comitato per l'archeologia laziale), Roma. 38-47.

CRESCENZI, L., TORTORICI, E. (1988). Ardea: resti di capanne nell'area del tempio di Colle della Noce, Quaderni della Soprintendenza Archeologica del Lazio. Roma.. 29-32.

DI MARIO, F. (2007). Ardea, la terra dei Rutuli, tra mito e archeologia : alle radici della romanità : nuovi dati dai recenti scavi archeologici, a cura di DI MARIO F., Roma.

DI MARIO, F. (2009). Ardea, l'area archeologica in località Le Salzare-Fossa dell'Incastro, in Lazio e Sabina, scoperte, scavi e ricerche (Atti del Convegno Quinto incontro di Studi sul Lazio e la Sabina, Roma 3-5 dicembre 2007), Roma. 331-346.

DIGICULT PROJECT (2003). "Towards a SemanticWeb for Heritage Resources", DigiCULT Thematic Issue 3. Accessible at: http://www.digicult.info/downloads/ti3_high.pdf

FRUTAZ , A. P. (1962). Mappa della campagna romana, in Le Piante di Roma, a cura di FRUTAZ A. P. Roma.

FRUTAZ, A. P. (1972). Le carte del Lazio, a cura di Frutaz A. P. Roma.

GIZZI, T. (1979). "Stazioni di superficie del Pontiniano nel territorio di Ardea", Documenta Albana 2.1: 19-42.

GRUBER, T.R. (1995). "Toward principles for the design of ontologies used for knowledge sharing", International Journal of Human and Computer Studies 43 (5-6): 907-928.

GRUPPO ARDEATINO DI PROMOZIONE CULTURALE (1983). Ardea: la storia, i monumenti, il territorio : guida storico turistica realizzata con il contributo della provincia di Roma, a cura del GRUPPO ARDEATINO DI PROMOZIONE CULTURALE, Roma.

GUAGLIUMI, A. (1981). "Risultato di un'indagine conoscitiva dell'AC Ardeatino-Laurentino. Necropoli dei Rutuli presso Ardea", Antiqua Rivista dell'Archeoclub d'Italia, 18: 70-73.

Guarino, N. (1998). Formal Ontology in Information Systems. In N. Guarino (a cura di), Proceedings of FOIS'98. Amsterdam: IOS Press. 3-15.

GUIDI, A. (1983). “Materiali provenienti da recenti ricognizioni in Ardea”, Immagini di una ricerca (Catalogo della Mostra) Roma 86-89.

MAY, K.(2006). “Integrating Cultural and Scientific Heritage: Archaeological Ontological Modelling for the Field and the Lab”, CIDOC CRM Sig Workshop, Heraklion. http://cidoc.ics.forth.gr/workshops/heraklion_october_2006/may.pdf

NOY, N. F.; SINTEK, M.; DECKER, S.; CRUBEZY, M.; FERGERSON, R. W.; AND MUSEN, M. A. (2001). “Creating semantic web contents with protege-2000”, IEEE Intelligent Systems 16(2):60–71.

PIEROBON BENOIT, R., PROTO, F., AIELLO, A., BRANDI, S., FURNARI, M. (2009). “Concettualizzazione e contestualizzazione dei beni culturali archeologici”, Archeologia e Calcolatori 16: 321-339,

REMUNA PROJECT (2006). REMUNA: Rete dei Musei napoletani. Istituto di Cibernetica CNR-E.Caianiello.

SIGNORE, O. (2009). “Representing Knowledge in Archaeology: From Cataloguing Cards to Semantic Web”, Archeologia e Calcolatori, 20: 118-128.

SIGNORE, O., MISSIKOF, O. (2005). “La Gestione della Conoscenza, in Archeologia: Modelli, Linguaggi e Strumenti di Modellazione Concettuale Dall'XML al Semantic Web", Archeologia e Calcolatori, 16: 291-319.

STAR PROJECT (2010). Semantic Technologies for Archaeological Resources. Accessible at: http://hypermedia.research.southwales.ac.uk/kos/star/

USCHOLD, M. & GRUNINGER, M. (1996). Ontologies: Principles, Methods and Application. Edinburgh: University of Edinburgh.

Imprint:

Proceedings of the 18th International Conference on Cultural Heritage and New Technologies 2013 (CHNT 18, 2013)

Vienna 2014

http://www.chnt.at/proceedings-chnt-18/

ISBN 978-3-200-03676-5

Editor/Publisher**:** Museen der Stadt Wien – Stadtarchäologie

Editorial Team: Wolfgang Börner, Susanne Uhlirz

The editor’s office is not responsible for the linguistic correctness of the manuscripts.

Authors are responsible for the contents and copyrights of the illustrations/photographs.