



HAL
open science

Adding Semantic Extension to Wikis for Enhancing Cultural Heritage Applications

Eric Leclercq, Marinette Savonnet

► **To cite this version:**

Eric Leclercq, Marinette Savonnet. Adding Semantic Extension to Wikis for Enhancing Cultural Heritage Applications. International Conference on Digital Information and Communication Technology and its Applications, 2011, France. pp.348-361. hal-00934969

HAL Id: hal-00934969

<https://u-bourgogne.hal.science/hal-00934969v1>

Submitted on 22 Jan 2014

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

Adding Semantic Extension to Wikis for Enhancing Cultural Heritage Applications

Éric Leclercq and Marinette Savonnet

LE2I Laboratory UMR CNRS 5158
Université de Bourgogne
9, Avenue Alain Savary
21078 Dijon, France
{Eric.Leclercq,Marinette.Savonnet}@u-bourgogne.fr
<http://le2i.cnrs.fr>

Abstract. Wikis are appropriate systems for community-authored content. In the past few years, they show that are particularly suitable for collaborative works in cultural heritage. In this paper, we highlight how wikis can be relevant solutions for building cooperative applications in domains characterized by a rapid evolution of knowledge. We will point out the capabilities of semantic extension to provide better quality of content, to improve searching, to support complex queries and finally to carry out different type of users. We describe the CARE project and explain the conceptual modeling approach. We detail the architecture of WikiBridge, a semantic wiki which allows simple, n-ary and recursive annotations as well as consistency checking. A specific section is dedicated to the ontology design which is the compulsory foundational knowledge for the application.

Key words: Semantic Wiki, Ontology Engineering, Cultural Heritage application

1 Introduction

Collaborative platforms that manage scientific knowledge are essential tools for scientists to help them to formalize their ideas, to develop theories collaboratively, to publish results (research articles, technical reports, data sets, etc.) and to produce knowledge for different kind of users. Moreover, a collaborative platform should be able to integrate other services such as visualization tools, or spatial analysis tools.

Wiki solutions meet the requirements of a web platform with collaborative capabilities. Easy setup and rich editing support are primary reasons for the widespread adoption of wikis. Users can enter text and others types of data (pictures, video) and connect content through hyperlinks. Most of wikis also provides a versioning system to track content changes and a full-text search engine for querying wiki pages.

The narrative structure is one advantage of wiki documents centric approach, compared to a database centric approach. In a database centric approach, the database schema is built upon entities identified in the first step of analysis, and thus based on an instant knowledge. In domains characterized by a rapid evolution of knowledge, such as biology or archaeology, a static database schema is not suitable and can be proscribed by the cost of evolution. Nevertheless, a mere document management system is not sufficient to catch interdependent structures of knowledge. For example, domain specialists often need to comment on primary data. Adding semantic annotation capabilities to documents allows different levels of interpretation and can sustain: 1) knowledge evolution by keeping track of the successive annotations; 2) better quality in the query evaluation process; and 3) amenable result displayed according to user skills. Annotations can be defined at a coarse grained level (whole document) or at a fine grained level (i.e. attached to a piece of text). An ontology must be associated to the annotation system to provide a semantics for annotation terms according to domain knowledge.

Semantic wiki solutions meet the requirements of annotation system and knowledge description. Adding semantics yields two dimensions of enhancements to a wiki: 1) adding a more formal structure to the wiki; 2) exporting, integrating and reusing information by the adoption of standard semantic technologies. Compared to a traditional database, a semantic wiki allows: 1) to expand the structure of documents content; 2) to enable a data model emergence from the usage; and 3) to support collaborative, distributed workflows and processes. Semantic wiki thus seems to combine the best from two worlds: structure from databases as well as expandability and collaboration capabilities from wiki systems.

The rest of the paper is organized as follow: section 2 gives an overview of the CARE project, section 3 describes the requirements and WikiBridge architecture, section 4 describes semantic tools for archaeology, and section 5 discusses related works. Finally, section 6 concludes the paper.

2 Project overview

The aims of the international project CARE (*Corpus Architecturae Religiosae Europaeae*) is the setting up of a corpus describing Christian edifices in Europe (<http://care.u-bourgogne.fr>). Italy, Spain, Czech Republic, Poland, Slovakia, France and Croatia have been included in the project four years ago. Each edifice is described in a document that focuses on the description of states of evolutions from the 4th century to the 11th century. The French corpus focuses on the 7th and 8th centuries with very rich decades in terms of number of monuments. The French part of the project is supported by a French ANR funding (ANR-07-CORP-011).

Representing and managing knowledge in cultural heritage require a deep understanding of specific concepts. Building collaborative platform brings out some challenging characteristics: 1) complexity of data (heterogeneous, incom-

plete, uncertain, inconsistent, spatial, temporal); 2) domain knowledge barrier; 3) evolving knowledge; and 4) skills of actors.

2.1 The CARE community

From an organizational perspective, the CARE project takes the form of an expert network collecting and providing information on edifices, analyzing historical sources, filling documents and collaborating in the exploitation of the corpus through smaller research groups. Furthermore, the project involves undergraduate students that help in collecting information but lack expertise required to interpret data. In France, more than sixty researchers are collecting and analyzing data concerning approximately 2700 monuments. Two key characteristics describe the CARE community:

multi-disciplinarity: The data collecting process involves archaeologists, historians, art historians, topographers, draftsmen. It is designed as a collaborative process which merges information from various disciplines;

inter-disciplinarity: The interpretation of data brings together all the actors which also enrich their respective practices by the confrontation of methods or problems.

2.2 Conceptual modeling for the foundational knowledge

Linster in [11] shows that the interaction among domain experts, knowledge engineers and tools creates the knowledge. He has shown that the process of elaborating a knowledge-based system is a constructive model-building process that includes: a discussion process between knowledge engineers and domain experts as well as the construction of a conceptual model (i.e. a general and abstract framework). Thus, the knowledge engineering activity encompasses the design of two kinds of models: model to make sense and model to implement systems.

We have applied Linster's guidelines to the CARE project in order to initiate a foundational knowledge from the corpus of documents. The first stage is the salient concepts identification. The key concept is the edifice to which it is essential to model changes. All constituent elements of a building need to be described. They can delimit space or define religious function (baptismal, funerary, etc.). All changes of space or religious function determine a new edifice state. Edifices and their evolutions are described in a set of documents. Salient concepts let us to build a conceptual model, three groups of elements have been identified (figure 1):

1. spatial concepts without temporal relationship (light grey): concept EGS refers to edifice, group of edifices or space inside edifice such as nave or apse. Composition relationships can be identified between EGS.
2. spatio-temporal concepts (grey) called SEGS, represent variations of spatial concepts in time. Spatio-temporal concepts are linked to a date or a period. Dating elements can be determined by documents, or described by methods such as C14, thermoluminescence or stratigraphy.

- identification of vocabulary terms for the project domain (black). While concepts EGS and SEGS are used to structure the descriptions, terms are used to describe specific elements, properties of edifices, religious functions or manufacturing techniques.

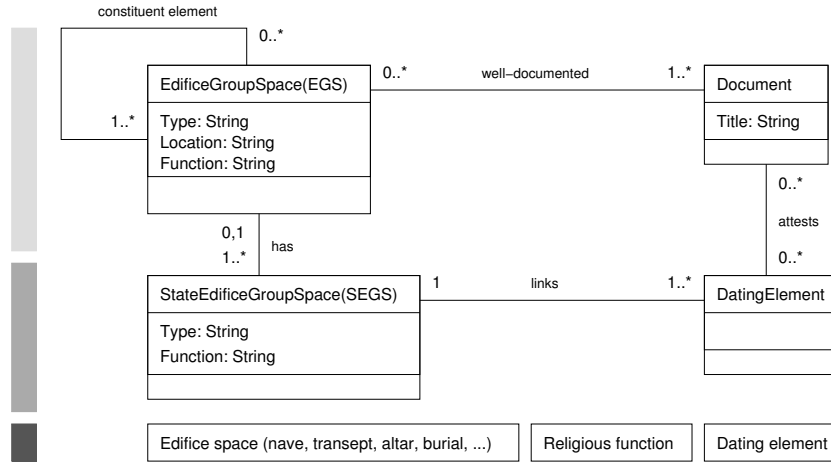


Fig. 1. Conceptual model of the CARE corpus

The description of the construction of the ontology is detailed in Section 4.

3 WikiBridge's architecture

In a survey authored by Uren et al. [15], authors study semantic annotation, identify a number of requirements, and review some semantic annotation systems. WikiBridge's design principles, in agreement with the CARE community, are following of the seven requirements given by Uren et al.: easy to use interface, user collaborative design, support of different user skills, support of heterogeneous format, compatibility with Semantic Web standards, annotation capabilities and storage, support for reasoning.

In the next subsection we develop the most important requirements with regards to the architectural design of WikiBridge.

3.1 Requirements

In a knowledge engineering process, it is common that non-technical domain experts work together with experienced knowledge engineers. To support different levels of users skill certain advanced functionalities should be hidden from novice users but made available to experienced users. Thus, we use an Access Control

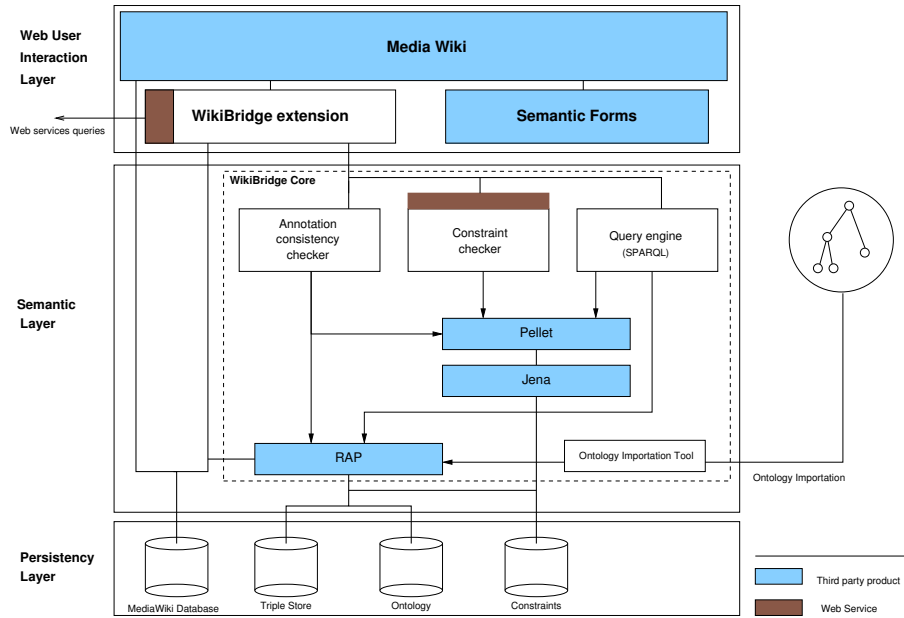


Fig. 2. WikiBridge’s Architecture

List (ACL) mechanism to describe privilege control depending on user identity and group affiliation. Advanced users can define forms to help users to enter descriptions of edifices, each part of a form generates automatically annotations. During the annotation process, a wizard suggest terms of the ontology according to highlighted section in the form.

To be able to exchange data with other applications (e.g. ontology editors, Web Services, other wikis), a compliance to Semantic Web standards is required. WikiBridge is purely based on existing Semantic Web standards such as the Web Ontology Language OWL for describing ontologies and W3C’s RDF for annotations.

We consider reasoning as one of the most important functionalities as it allows: 1) to emerge knowledge that is not explicit in the data; 2) to check the meaning of annotations with regards to the context of the annotation; and 3) to enhance navigation and search.

3.2 Architecture

One of the most famous semantic wiki is Semantic MediaWiki (SMW) which is based on MediaWiki [16, 9]. In 2009, when we have started the project, complex annotation and consistency checking were identified as mandatory functionalities. In 2009, SMW doesn’t provide complex annotation and doesn’t have consistency checking in its roadmap. We have extended MediaWiki with the following semantic components: form based acquisition interface with automatic

annotation, annotations wizard, annotations validation based on the context of a document, semantic rules and a semantic query engine. The semantic components are structured in three layers.

User interaction layer The user interaction layer is covered by MediaWiki and structured data control needed for unexperienced users is managed by Semantic Forms (an extension¹ for MediaWiki). A specific description language allows administrators and advanced users to define new forms (figure 3). Modules corresponding to the interaction layer are represented on the top of figure 2.

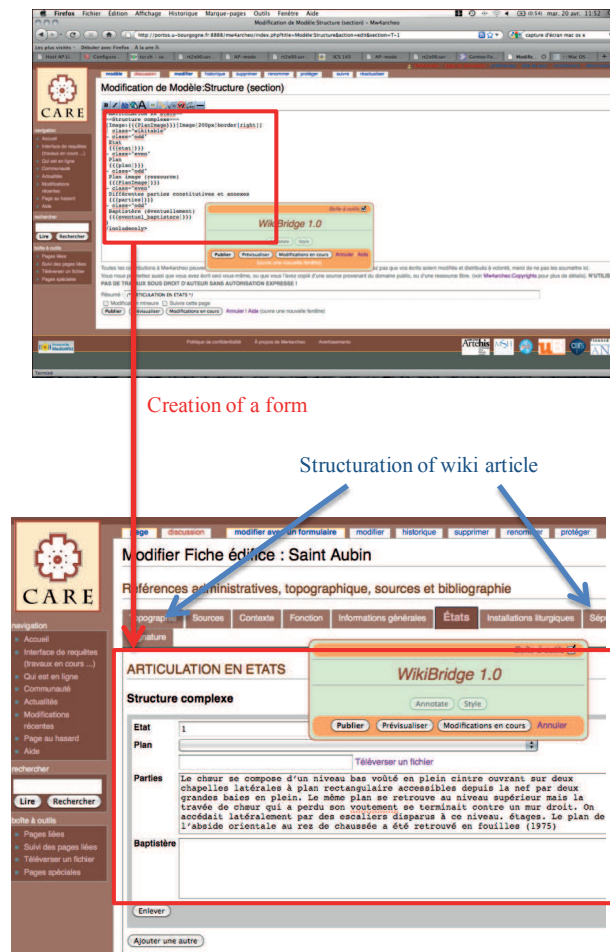


Fig. 3. User interaction layer

¹ http://www.mediawiki.org/wiki/Extension:Semantic_Forms

Semantic Layer To improve the quality of information during the input and annotation processes, we propose three semantic components in WikiBridge Core (white boxes in figure 2) developed on the top of third party tools (RAP - RDF API for PHP, Pellet and Jena). The syntax and the semantics of annotations, made by experts, are guaranteed by an application ontology. Values lists used in annotation wizard are filled by individuals retrieved from ontology concepts, through SPARQL queries that are used to select individuals. The annotation wizard helps users to construct simple or complex annotations by selecting ontology terms in lists and giving them properties and values (figure 4). Each document is identified by its URL in the wiki and annotations use this URL as a basement for identifying part of content. Simple annotation allows to annotate a subject by describing a property by a literal. Complex annotation allows to annotate a subject with two or more values and references to other elements (subjects). For example we can annotate an altar with its dimension, its building material, its location in the nave. The nave is detailed in another part of the document. Annotation construction is a context sensitive process, initial terms displayed to users are directly connected to fields of forms.

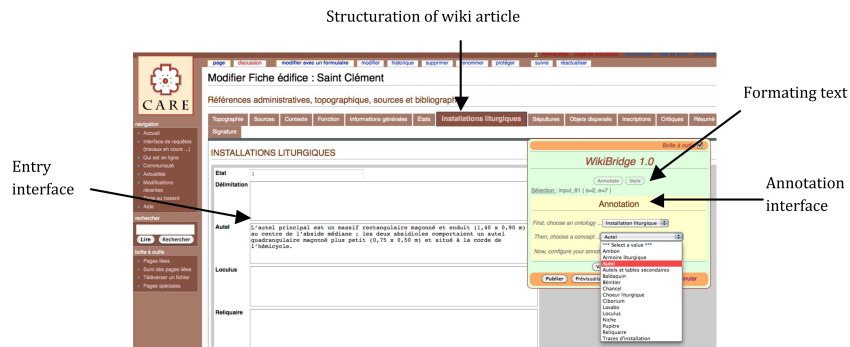


Fig. 4. Annotation wizard

Annotation consistency checking is then operated by a set of specific components interacting with RAP, Pellet and Jena. A subset of first order logic constraints are checked with a Java component that interacts with Pellet and Jena. So, WikiBridge extension connects to the Java constraint checker by the means of a web service. Moreover, rules can be added to query ontology and annotations in order to test new facts and thus to produce new knowledge that can be inserted in the set of semantic constraints.

Two kinds of constraints can be check by using the ontology knowledge: 1) domain values of properties using ABox capabilities; and 2) structural consistency of properties using TBox capabilities (for instance, a cathedral can have a nave but cannot have an atrium). Nevertheless, some domain dependent constraint cannot be embedded in the structure. For example "In France, there is no

church with rammed earth wall for the studied period however, this technique is used in Ireland and in other countries” can be translated by the following constraint that must remains consistent:

$$\text{hasRammedWall}(?x) \wedge \text{edificeType}(?x, ?t) \wedge \\ \text{edificeCountry}(?x, ?c) \wedge c='France' \wedge t \neq 'church'$$

Persistency layer The persistency layer includes four types of storage (bottom of figure 2): documents content, semantic annotations, ontology, and constraints.

The content of documents is stored by MediaWiki specific database. Annotations are stored as triple in RAP triple store and they can be retrieved by WikiBridge user layer to display annotations with icons and colors (different colors are used to distinguish centuries) in a document or by the SPARQL query engine. The ontology imported as an OWL file is stored in a specific schema managed by RAP. Ontology terms can be then queried using SPARQL and results can generate wiki pages. Constraints are stored in plain text using Jena rules syntax. A type attribute specifies if the rule can be applied to check the ontology structure or to check annotation consistency.

3.3 Users with different skills

Information access has been designed with taking into account some features about users. We have thus identified a usage typology in accordance to 1) kind of usage (reader, investigator, annotator); 2) knowledge degree of the domain (domain specialists like historian researchers and non specialists). To handle these different types of users, we offer three types of queries:

1. faceted browsing allows users to explore by filtering available information with the ontology structure;
2. form based searching provides semantic search by filling in parameters of parametric queries identified during the analysis of requirements;
3. aggregate view for each article, all annotations related to the edifice are displayed in a factbox.

Nevertheless all types of queries rely on the SPARQL query engine which also allows to process in line queries into wiki pages in order to summarize informations.

To operate spatial and temporal analysis on annotations a set of web services has been developed (figure 5). Some specific services allow to retrieve edifices and their coordinates according to a set of conjunctive properties. Moreover, a generic web service have been developed to handle SPARQL queries.

4 Offering semantic tools for archaeology

Cultural heritage collections can be annotated with different thesauri. The web page <http://tinyurl.com/5u8bjer> maintains about twenty thesauri, classified by content and organization.

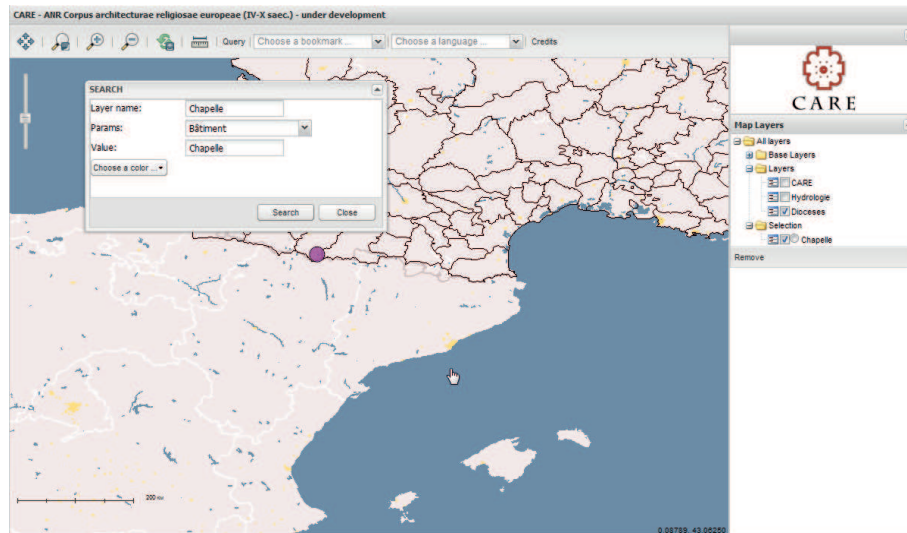


Fig. 5. OpenLayers interface interacting with a web service of WikiBridge

The CIDOC Conceptual Reference Model (CRM) provides an extensible ontology for concepts and relationships in cultural heritage domain. Since 2006, it is an ISO standard (21127:2006) for exchange of cultural heritage information [6]. The CIDOC CRM emerged from the ICOM/CIDOC Documentation Standards Working Group (<http://www.cidoc-crm.org>). CIDOC-CRM aims at treating all types of material collected and/or displayed by museums: sites, monuments as well as collections of fine and applied arts. It is intended to encompass the detailed description both of individual materials as well as groups of materials as a whole. It covers also contextual information: historical, geographical context in which materials are placed and which gives them much of their significance and value. The event notion in CIDOC-CRM is represented by the concept of *event*. An event describes environment of the material over the time and what could happen to him. The central notion is complemented by: 1) the *TimeSpan* concept describes the moment it happened; 2) the concept of *Place*; 3) who did it (*Actor*); and 4) what is being described. CRM offers notions of *Physical Objects* that can be natural or have been manufactured by man, and *Conceptual Objects*.

Since CRM is a reference in the field, we use it as a starting point to establish CARE ontology. The view of the ontology CARE as a specialization of CRM allows us to relate to a standard and comply with it. The CARE ontology has two parts: 1) religious concepts, their spatial relationships and characteristics (class EGS in figure 1) and 2) timeline to track evolutions.

4.1 Religious concepts in the CARE project

Religious concepts in CARE are edifices, represented by the concept **Building**, with its decomposition into different constituent elements (nave, transept, apse, etc.) represented by the concept of **Structure**. To detail parts of an edifice, we introduced the concept of **ElementArchitectural** to describe masonry, floor, opening, etc. Liturgical installations (altar, ambo, ciborium, etc.) are represented by the concept of **InstallationLiturgique** and burials represented by the concept **Tombe**. These concepts have been placed under the concept *E24 Physical Man-Made Thing* CIDOC-CRM. Indeed, CIDOC-CRM defines this concept as "all persistent physical items that are purposely created by human activity". Figure 6 represents all these concepts (concepts with EXX are CIDOC-CRM concepts).

4.2 Modeling spatial relationships in archaeology

The geometry implementation in the textual descriptions that are analyzed is a complex geometry. Indeed, these descriptions do not refer to an absolute and orthonormal space: it is rather, a space perception or a cognitive space whose structure is largely based on the functional aspects and objects described, and the perspective of the archaeologist. From the analysis of textual descriptions of religious concepts (description of the position and shape) given by archaeologists we have found four types of spatial properties:

- **orientation properties:** forward, back, bottom, next to, on one side, under, below, at a lower level, at the same level, above, right, left, center, prior, and the cardinal directions;
- **boundary properties:** outside, inside;
- **distance properties:** near, far, next to, around;
- **topological properties:** flank, join, open on the side, link to, stand against, surround, isolated. To represent the topological properties, we used the work of Hegenhofer and Herring [8]. The authors have defined a minimum set of eight relations (disconnected, externally connected, partially overlap, equal, tangential proper part, non-tangential part, tangential proper part inverse, non-tangential part inverse) describing the relations between two regions.

4.3 Modeling temporal knowledge to track evolutions in CARE project

When writing of his excavation report, the archaeologist graphically summarizes the results obtained with a timeline which is often organized by anterior/posterior relationships: materials are considered in relation to each other. In the CARE project, time model is based on following criteria: some absolute benchmarks and a relative chronology based on intervals. We have established a convention for century division and boundaries. Centuries start at year 1 and end at year 100. We also have established subdivisions terms such as early (1 → 32), mid (33 →

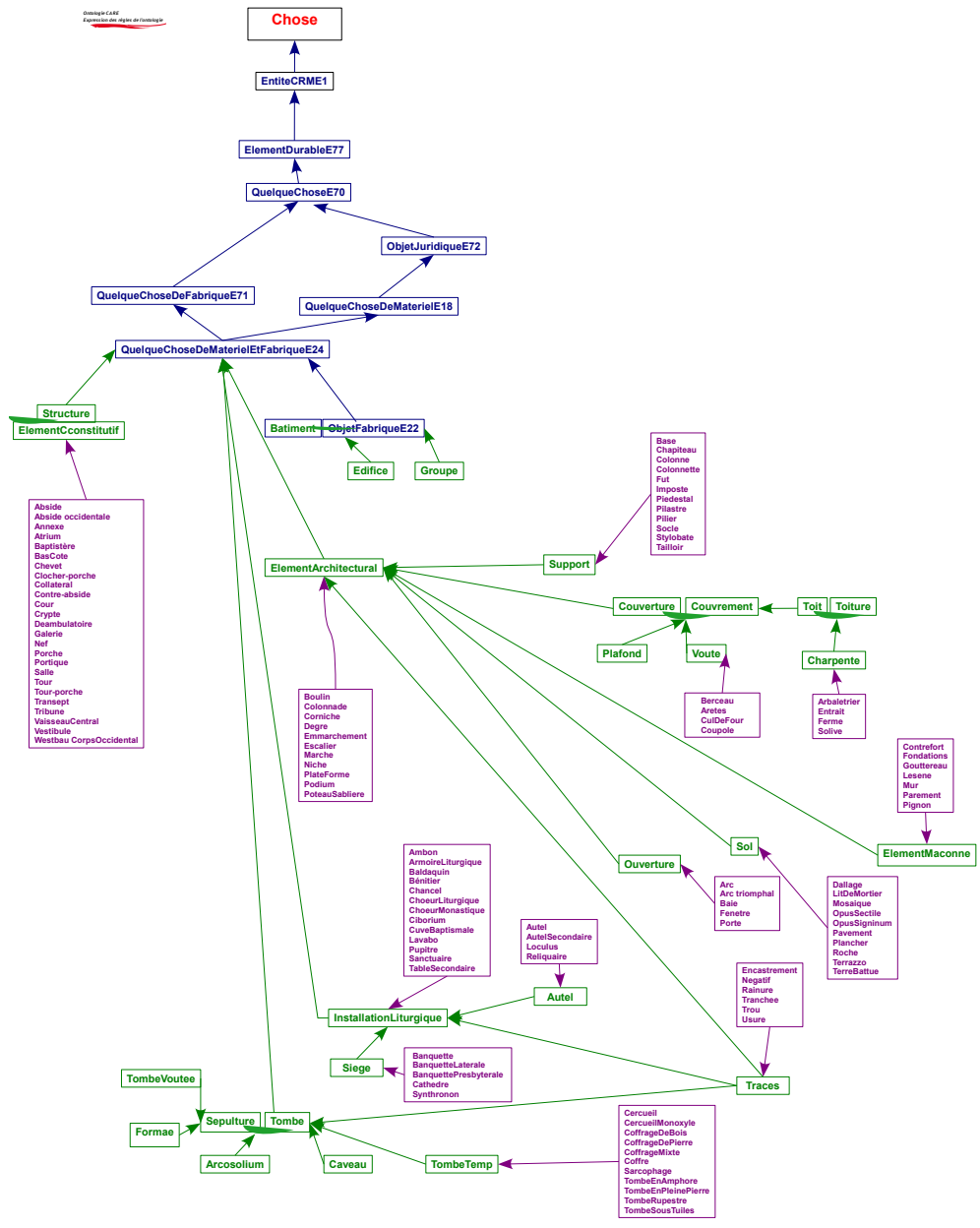


Fig. 6. Religious concepts in CARE project

66) and late (67 → 100). These century divisions are placed under *E52 TimeSpan*. CIDOC-CRM offers specific concepts related to time [3, 7]. The concept *E2 Temporal Entity* describes all phenomena which happen over a limited extent in time. It includes notions such as: *E3 State Conservation* comprises the states of objects characterized by a certain condition over a timespan; *E4 Period* includes sets of coherent phenomena or cultural manifestations bounded in time and space. Some Allen' relationships [1] are properties used.

The CARE project aims to follow the evolutions (creation, modification, deletion) of an edifice and its constituent elements. The concept of activity is important for CARE project because a state may be characterized by an activity in the CIDOC-CRM ontology. The concept *E7 Activity* is defined as follows: "The action or sequence of actions intentionally carried out by Actors that result in changes of state in the cultural, social, material systems which interest". This notion includes both complex and long-lasting actions such as building an edifice, as well as simple and short-lived actions. Following the concepts of the CIDOC-CRM ontology, we use seven concepts to model states of edifices:

- *E6 Destruction* includes events that destroy one or more instances of the concept *E18 Physical thing*;
- *E11 Modification* includes all instances of *E7 Activity* that modify *E24 Physical Man-made thing*;
- *E81 Transformation* corresponds to the destruction of one or more objects and the simultaneous production of others using parts or materials from the first. Transformation preserves recognizable substance from the first element, but has a different nature;
- *E63 Beginning of Existence* includes events that bring into existence to any *E77 Persistent Item*;
- *E64 End of Existence* includes events that end the existence of any *E77 Persistent Item*;
- *E79 Part Addition* includes activities that result in the fact that an instance of *E24 Physical Man-made Thing* is enlarged or increased by the addition of a party;
- *E80 Part Removal* includes activities that result in the fact that an instance of *E18 Physical Thing* is reduced by the removal of a party.

The CARE ontology has been designed using Protégé and it actually encompasses 124 classes and 715 individuals.

5 Related works

Several semantic wikis have been developed or used specifically for cultural heritage applications.

Witte et al. [17] present an approach to cultural heritage data management which integrates different technologies: a wiki user interface, text mining support using a Natural Language Processing (NLP) framework and ontologies based on OWL and RDF. Authors have implemented the ideas for the

German Handbuch der Architektur, a comprehensive multi-volume encyclopedia of architecture. A volume (506 pages) of the encyclopedia is converted into wiki pages. Authors have to capture two sub-domains by ontologies: the domain of document management (i.e. sentence, noun, page number, etc.) and architectural domain (i.e. wall, building material, etc.). NLP allows to connect architectural concepts with document-specific one, e.g. sentences that mention construction elements of a certain material. A public version is available at <http://durm.semanticsoftware.info/wiki>.

The HermesWiki [13] is a semantic wiki in the historical domain in German language. The main objective is to provide an overview on Ancient Greek History for teaching purposes of undergraduate students. The wiki consists of three parts: a collection about twenty essays giving a comprehensive domain walk-through, translations of the describing ancient sources and a glossary. The entries in the glossary are tagged. It has been implemented as a plugin for KnowWE [2], reusing as much of the core components as possible. A public version is available at <http://hermeswiki.informatik.uni-wuerzburg.de>.

NavEditOW is a framework for ontology driven web site. It has been exploited to support a semantic description of two projects: 1) a web portal and a set of advanced services supporting the sharing of knowledge about Prehistory and Protohistory in the Italian context [4]. In particular, one of the services is represented by a digital library, in which entries (i.e. bibliographic description of publications) will be ontologically described. The system is currently on-line at URL <http://www.archeoserver.it> and; 2) SilkRoDE (Silk Roads in the Digital Area) project that aims to collect, structure and diffuse all knowledge about the Cultural Heritage of Central Asia from fields such as archeology, geography or history [5]. The ontological approach provides the required expressiveness and flexibility to support rich forms of navigation among stored contents. The framework integrates a wiki engine for rendering documents stored in the ontological tier.

In the same view, MANTIC is a web application that realizes a portal for archaeological information about the city of Milan [12]. MANTIC integrates different data sources and the global schema is based on CIDOC-CRM.

Our approach of semantic wiki is directed towards scientific application domains which contribute to produce knowledge. These kind of applications rely on core ontologies that act as a consensus. Knowledge is enhanced by querying and analyzing data, new concepts can emerge and new constraints can be found out. In contrast with above projects, WikiBridge provides a rich annotation model [10] based on semantic values [14] and supports semantic consistency checking upon the ontology and constraints.

6 Conclusion

In this article, we have presented WikiBridge which provides users with advanced functionalities such as rich annotation model, consistency checking. The first experiment of the use of WikiBridge shows many interesting possibilities for

scientific community, mainly the possibility given to scientists share and collaboratively build annotated knowledge. A knowledge that can be re-organized at any time, to fit the needs of the scientists. We have demonstrated that flexibility and data quality required by scientific applications can be achieved by using wiki with Semantic Web technologies. The semantics of annotation is guaranteed by an ontology including constraints which allow to describe accurately domain knowledge. Our dual approach allows to cope with evolution of knowledge by dynamically modifying the ontology and annotations without modifying database schema.

References

1. J. F. Allen. Maintaining Knowledge about Temporal Intervals. *Commun. ACM*, 26(11):832–843, 1983.
2. J. Baumeister, J. Reutelshoefer, and F. Puppe. KnowWE: community-based knowledge capture with knowledge wikis. In *K-CAP*, pages 189–190, 2007.
3. C. Binding. Implementing Archaeological Time Periods Using CIDOC CRM and SKOS. In *Proceedings of the 7th Extended Semantic Web Conference (ESWC)*, pages 273–287, 2010.
4. A. Bonomi, G. Mantegari, and G. Vizzari. A Framework for Ontological Description of Archaeological Scientific Publications. In *SWAP*, 2006.
5. A. Bonomi, A. Mosca, M. Palmonari, and G. Vizzari. Integrating a Wiki in an Ontology Driven Web Site: Approach, Architecture and Application in the Archaeological Domain. In *Proceedings of the Third Workshop on Semantic Wikis: The Wiki Way of Semantics (SemWiki)*, 2008.
6. CRM. The CIDOC Conceptual Reference Model (ISO/CD 21127). Technical report, International Council of Museum.
7. M. Doerr, A. Kritsotaki, and S. Stead. Thesauri of Historical Periods - A Proposal for Standardization. In *Proceedings of CIDOC Conference*, 2005.
8. M. Hegenhofer and J. Herring. Categorizing Binary Topological Relations Between Regions, Lines and Points in Geographic Databases. Technical report, National Center for Geographic Information and Analysis, CA, 1991.
9. M. Krötzsch, D. Vrandečić, and M. Völkel. Semantic MediaWiki. In *International Semantic Web Conference*, pages 935–942, 2006.
10. E. Leclercq and M. Savonnet. Access and annotation of archaeological corpus via a semantic wiki. In *Fifth Workshop on Semantic Wikis - Linking Data and People (Semwiki)*, 2010.
11. M. Linster. Viewing Knowledge Engineering as a Symbiosis of *Modeling to Make Sense* and *Modeling to Implement Systems*. In *GWAI*, pages 87–99, 1992.
12. G. Mantegari, M. Palmonari, and G. Vizzari. Rapid Prototyping a Semantic Web Application for Cultural Heritage: The Case of MANTIC. In *Proceedings of the 7th Extended Semantic Web Conference (ESWC)*, pages 406–410, 2010.
13. J. Reutelshoefer, F. Lemmerich, J. Baumeister, J. Wintjes, and L. Haas. Taking OWL to Athens: Semantic Web technology takes Ancient Greek history to students. In *Proceedings of the 7th Extended Semantic Web Conference (ESWC)*, pages 333–347, 2010.
14. E. Sciore, M. Siegel, and A. Rosenthal. Using semantic values to facilitate interoperability among heterogeneous information systems. *ACM Trans. Database Syst.*, 19(2):254–290, 1994.

15. V. Uren, P. Cimiano, J. Iria, S. Handschuh, M. Vargas-Vera, E. Motta, and F. Ciravegna. Semantic Annotation for Knowledge Management: Requirements and a Survey of the State of the Art. *Web Semantics: Science, Services and Agents on the World Wide Web*, 4(1):14–28, 2006.
16. M. Völkel, M. Krötzsch, D. Vrandečić, H. Haller, and R. Studer. Semantic wikipedia. In *Proceedings of the 15th international conference on World Wide Web*, WWW '06, pages 585–594, 2006.
17. R. Witte, R. Krestel, T. Kappler, and P. C. Lockemann. Converting a Historical Architecture Encyclopedia into a Semantic Knowledge Base. *IEEE Intelligent Systems*, 25(1):58–67, 2010.