A SPECIFIC ONTOLOGY AND RELATED WEB SERVICES FOR ASSESSING ACCESSIBILITY ISSUES IN CULTURAL HERITAGE ENVIRONMENTS

Paloma Martín [a], Beatriz Valverde [a], María Muñoz [a], Mercedes Martínez [a] and Javier Finat [a]

[a] MoBiVAP, University of Valladolid, 47011 Valladolid, Spain – {palomamartinlopez, b.valverde.olmedo, mariamunoznieto}@gmail.com

KEY WORDS: Cultural Heritage, Knowledge management, Accessibility ontology, Web Services, Context Awareness

ABSTRACT:

The "Design for All" (DfA, 2010) strategy proposes a general framework for accessibility to Cultural Heritage able to integrate 3D Documentation, Information and Management (3DIM) Systems in monitored environments. Interoperability between information systems requires a common reference framework developed as a semantic layer based on an ontology. This ontology integrates concepts and relations involving a formalization of knowledge on physical domain, user profiles and tasks to be developed. The developed ontology provides support for identifying problems, selecting the most appropriate techniques and solving troubles along interventions by technicians. In this work, we introduce a Web Service in the Semantic Web framework supported by the ontology formalized in OWL (Ontology Web Language) and implemented with Jena framework. This solution is currently being applied in the Maritime Museum of Barcelona and a small urban district of the historic city of Segovia (Spain).

1. INTRODUCTION

The functional approach to Cultural Heritage is focused towards to understand and provide a support for the interaction of citizens with the environment. The interaction involves both the physical domain and digital issues. It must be performed not only by removing physical obstacles, but also logical gaps which limit the understanding or relevance of Cultural Heritage. In this way, we intend to improve the integration through multiple connections which are compatible with the multiplicity of meanings involving complex cultural objects.

It is necessary to develop a common framework for Knowledge Management Systems able to integrate physical aspects, their relationships and different meanings for citizens, with a special regard to persons having some troubles for communicating or understanding complex realities. This is an ambitious program that implies a lot of things such as removing physical obstacles, providing services or filling the digital gap relative to more friendly visualization tools in order to solve accessibility issues in Cultural Heritage domains. This general problem is well known in linguistics, where syntactic analysis precedes to more elaborated formal, structural or functional approaches involving a specific thesauri (list of words to represent terms with its definitions or descriptions) and taxonomies (classification system composed by a hierarchy of categories), which are integrated in a common ontology.

Iconological studies are well known from the middle of sixties of the 20th century, and provide a support for lexicon. Nevertheless the contributions of structural approaches along the seventies and due to the very large diversity of shapes, contexts and meanings, there is a low understanding of relations between "components" in Cultural Heritage (CH), still. This heterogeneity is due to multiple factors, which are not easy to formalize. So, the multiplicity of meanings is translated in different interpretations which coexist, giving different relations between components. It is necessary to formalize and represent such relations for an efficient management in terms of different Systems -Documentation, Information and Management- which are relevant for generating new knowledge.

Following our reasoning, the integration of 3D Documentation, Information and Management (3DIM) Systems in monitored environments requires a common reference which is developed in a GIS framework augmented by a semantic layer for enabling interoperability, including facilities for friendly data management by non-expert users. These functionalities require a well-defined ontology able of interconnected services linked to the above three 3DIM Systems. The developed Ontology has been specifically designed for the PATRAC[1] project which integrates concepts and relationships involving a formalization of knowledge on physical domain, user profiles and tasks to be developed. Two important kinds of users are disabled persons wishing to enjoy contents and technicians which intervene in the physical domain. Georeferenced models arising from an image- and range-based surveying provide an objective representation of the monitored environment which supports the developed Information and Management Systems.

The semantic layer overlays all modules corresponding not only to the 3D DIM Systems, but the processes (Processing, Analysis) and applications (Visualization, Assessment) linked to multimedia database and services. The design and implementation of Web Services is the key for developing a universal solution, independent of technical devices and communications systems, requiring only an Internet access which is performed through mobile devices. On the other side, context awareness is crucial for providing services for solving accessibility issues to disabled or dependent persons. The developed ontology provides a support for identifying problems, selecting the most appropriate techniques and solving troubles along interventions by technicians.

There are several definitions for ontology that are source of confusion or ambiguity. We accept the definition proposed by

[1] In Spanish,"PATrimonio ACcesible: I+D+i para una cultura sin barreras" (Accessible Heritage)

Studer et Al. (Studer, 1998), based on previous definitions of Gruber and Borst (Gruber, 1993; Borst, 1997), which states that "an ontology is a formal, explicit specification of a shared conceptualization”. The same authors define conceptualization as the extraction process for the abstract model of a part of reality by the identification of the concepts. These concepts must be explicitly defined. The result should be machine-readable, so the ontology must be formalized. Finally, the conceptualization and the formal representation should be accepted by all users of the ontology, i.e., it should be shared and re-used by everyone.

The rest of the document is organized as follows: section 2 synthesizes the previous related work in the field of Semantic Web for cultural heritage; section 3 describes the development process of the ontology and taken design decisions; section 4 frames the ontology inside the global environment of the PATRAC project; section 5 is focused on the development and the behavior of the Web Service to exploit the previous described ontology; and finally, section 6 concludes the paper and shows future guidelines in our research.

2. RELATED WORK

The problem of finding and relating cultural heritage information in heterogeneous content with different data format creates an obstacle for end-users and a challenge to research communities. The literature introduces several approaches to ease these problems. (Lynch, 2002) highlights the importance of digitalizing cultural heritage documentation creating Digital Libraries and Digital Collections to make available cultural heritage content. It raises the need for an infrastructure based on a common vocabulary and vocabulary mapping, but out of the Semantic Web.

(Doerr, 2003) establishes the first ontology for cultural heritage data in collaboration with the International Council of Museums. This high level ontology called CIDOC Conceptual Reference Model is an annotation ontology standard ISO since 2006. It provides an underlying schema composed by over 200 concepts and relations into which other schemas can be transformed, but it does not contain domain ontologies for filling in property values or to detect accessibility issues. Other approaches like (Benjamins, 2004) extract ontology annotations automatically, integrating different repository contents, but obviating reasoning about them or reflecting accessibility issues.

Semantic portals (Hyvönen, 2009) collect contents of various publishers into a single site, based on Semantic Web standards in order to improve structure, extensibility, customization and usability of traditional portal designs. Although they provide reasoning task for recommendations or association discovery, they do not assess accessibility issues since ontology does not model them.

Geospatial Semantic Web (Kauppinen et Al, 2010) is the new emerging approach that merges two trends of the current state of the art, where *Geospatial* shows the important role of places in the Web and where the Semantic Web enables to be able to explain relationships. This proposal suggests that cultural heritage contents were annotated with its geographical position and processed using some kind of spatial ontology. Although this approach is nearer of our work, it does not consider accessibility issues.

3. PATRAC ONTOLOGY

In this work, we aim to develop an ontology called PATRAC Ontology with the goal of providing a semantic framework for 3D georeferenced information systems. The ontology represents knowledge on physical domain, user profiles and tasks to be developed in cultural heritage environments. The instances that populate the ontology are stored in a relational database with an entity-relation schema that maps the ontology. Thus, the PATRAC Ontology defines unambiguously the concepts which are referred by data repositories.

The development of the ontology is based on the methodology *Methontology* (Gómez-Pérez, 2004). This methodology enables the construction of ontologies at knowledge level and includes: the identification of the ontology development process, a life cycle based on evolving prototypes, and techniques to carry out each activity in the management, development-oriented, and support activities. The figure 1 shows the entire ontology life cycle (Gómez-Pérez, 1998) with the tools, methodologies and technologies around it. The development activities are: *specification*, *conceptualization*, *formalization*, *implementation* and *maintenance*.

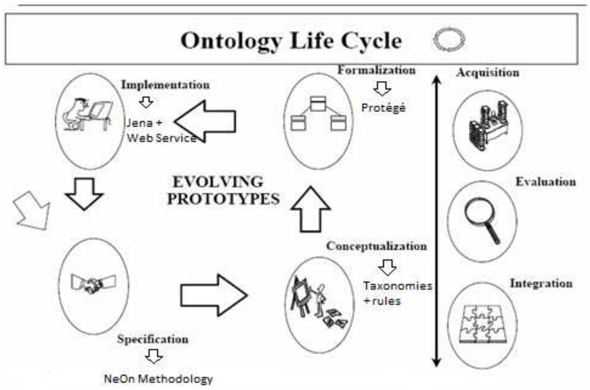

Figure 1. *Methontology* life cycle, adapted from (Gómez-Pérez, 1998)

3.1 Specification

Along this phase of ontology development we must analyze requirements and needs which our ontology must response. In accordance to the methodology described by *NeOn Methodology* (Suárez-Figueroa, 2008) this involves: explain the main purpose, application domain, level of formality, users, uses and competency questions.

The *main purpose* is to provide/explain a formal knowledge model to represent accessibility issues in cultural heritage environments in architecture and possible interventions to solve or improve these issues. The PATRAC project framework provides support for information management and web services related with information systems.

The *application domain* refers to accessibility conditions in outside or inside of cultural heritage environments (see section 3.2 for a more detailed analysis).

The *level of formality* in which the ontology is expressed is semi-formal since the formalization has been done with OWL W3C language specification as it shows in section 3.3

We identify the following application *users*:

- *Technicians* or professionals in charge of maintenance and rehabilitation interventions in monuments.
- *Visitors* or people interested only in cultural value of monument. Taking into account visitor disabilities is essential to determine the accessibility degree of monuments. Thus, we divide disabilities between *physics* and *psychics*.
- *Service providers* or professionals who schedule and design interventions and services in order to publish, advertise, advance, etc. the cultural heritage.
- *Content creators* or users who design and publish data and applications linked to the environment.

The possible *uses* are structured by the different user roles:

- *Technician*, which is responsible for the description of monument and its environment; the management of the information and knowledge that emerges from its activity, document searches, etc.; the analysis of architectural environment which is object of intervention; the definition of intervention proposals concerning the monument and its environment.
- *Visitor*, who has the ability to query about monument accessibility issues or about context awareness points of interest for tourist.
- *Service Provider*, which retrieves cultural heritage information in order to design and define possible activities and services after an analysis.
- *Content creator* that manages social interesting contents, especially multimedia data.

Finally, in order to identify the *competency questions*, we have to define the questions which the ontology must response. They are a first approach to the vocabulary, the relationships and properties we have to explain. For instance, what kind of accessibility issues could a person in a wheelchair finding around the Maritime Museum of Barcelona?

3.2 Conceptualization

From the application domain analysis for PATRAC ontology we take into account the need to subdivide it into three different subdomains:

- *The physical domain Ontology*, which contains monument description including both architectural features and accessibility issues. Also, it establishes a relationship between the most of its concepts and its geographical position. Figure 2 shows a part of the concepts of Physical Domain Ontology. This schema shows that monuments have spaces which themselves contains construction elements, accessibility issues and cultural objects. Pathologies are linked to construction elements and each concept could have attached multimedia resources.
- *The task Ontology*, which describes interventions in order to enhance accessibility conditions or architectural structure with properties like urgency, difficulty, etc.
- *The user Ontology*, which classifies and models properties for users involved in intervention processes, creation content and access to services

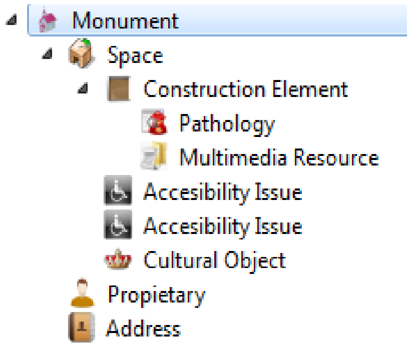

Figure 2. A part of the concepts of Physical Domain Ontology

3.3 Formalization

There are different languages to formalize ontology contents inside the Semantic Web Framework, such as RDF(S) and OWL. Next, we will summarize the main aspects of each one.

RDF stands for Resource Description Framework (W3C, 2004c). Since RDF data model does not have mechanisms for defining vocabulary and constraints in the domain, range and relationships between concepts, then it emerges RDF Schema (W3C, 2004d). RDF(S) combines semantic networks with frames but it does not provide all the primitives that are usually found in frame-based knowledge representation systems (Gómez-Pérez, 2004).

The expressivity of RDF and RDFS is deliberately very limited: RDF is (roughly) limited to binary ground predicates, and RDF Schema (roughly) is limited to a subclass hierarchy and a property hierarchy, with domain and range definitions of these properties. OWL (W3C, 2004b) overcomes these limitations since it allows defining disjoint classes, cardinality constraints and inverse or transitive relationships (Antoniou, 2004). Our ontology contains both disjoint classes and transitive properties. For instance, a sketch resource can not be an audio resource at the same time; the property *isComposedBy* is a transitive property of monument and space. Thus, we decided to formalize the ontology in OWL.

Protégé (Stanford, 2010) is the ontology editor and knowledge-base framework that we choose in this stage. The choice of Protégé was taken due to several reasons. In first place, Protégé is the most widely used editor for ontologies by the Semantic Web research community. Besides, Protégé is Open Source and freely redistributable software, which is available for everyone. And finally, Protégé is well documented and supported by a wide research community.

3.4 Implementation

There are several Semantic Web frameworks to implement reasoning task with Ontology, such as Sesame (opeRDF.org, 2010), Mulgara (Mulgara, 2010), AllegroGraph (Franz Inc, 2010), Jena (HP Labs, 2009), etc. The first three frameworks only allow RDF data management, but the ontology requires OWL, so we finally chose Jena.

Jena is a Java framework enabling Semantic Web application development. The reasons why we chose Jena are: it provides OWL in memory or persistence storage management; it includes a rule-based engine allowing inference reasoning; it has an

SPARQL query engine. Jena manages OWL ontology models stored in persistent storage like in relational databases. In order to export the ontology model from Protégé we have used the Protégé2jena (Barhatov, 2006) plug-in for Protégé.

SPARQL (W3C, 2008) is a query language and a protocol for accessing RDF. As a query language, SPARQL is "data-oriented" in that it only queries the information held in the models; there is no inference in the query language itself. Of course, the Jena model may be 'smart' in that it provides the impression that certain triples exist by creating them on-demand, including OWL reasoning. The information required by the clients in a query is returned in the form of a set of bindings or an RDF graph.

4. ONTOLOGY IN THE PATRAC FRAMEWORK

The PATRAC Ontology defines the semantic layer for 3D information and management system, called GIRAPIM, part of PATRAC project. The aim of this system is to simplify and ease the task of populate the ontology inside a 3D GIS environment.

There are two ways of storing individuals and ontology in a relational database. First of them is Jena persistence subsystem that stores both individuals and ontology model with the same database schema based on RDF triples. The second one treats the ontology model with the Jena database schema and, on the other hand, the individuals are stored following a specific entity-relation schema. While the first solution eases ontology management and reasoning tasks, the second solution provides efficient access through traditional SQL query engine. Also, the last solution enables to exploit the power of PL/SQL language and domain-specific extensions like GIS.

The efficient access and geographical reference support for the ontology individuals are two specific requirements for GIRAPIM as a 3D geographic information system. This application requires a heavy storage and update activity in relational databases that can not be sent through Jena to avoid a bottleneck. For these reasons, the second solution with two databases has been selected (see figure 3). The first database stores the ontology graph managed by Jena, and the second one stores ontology individuals representing information about the cultural heritage and linked multimedia resources.

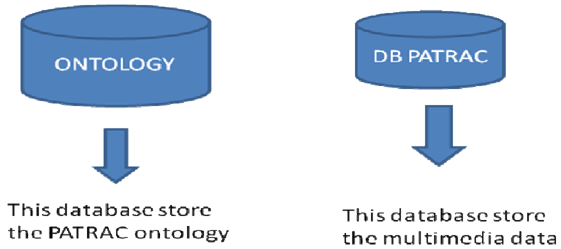

Figure 3. Database and ontology

5. ACCESSIBILITY ASSESSMENT WEB SERVICES

The Web Services are designed to offer context aware services, taking into account the user type, its position, and the task to be developed. Since there are different user needs, the service logic has been decoupled into two different services, one for the technicians and another for the visitors. The first is tailored to technician tasks and provides support to manage data related to new accessibility issues, pathologies or interventions linked to a monument. The second one provides services allowing visitors to get information suited to its interest, its disability and geographical position.

W3C defines a WS (Web Service) as "a software system designed to support interoperable machine-to-machine interaction over a network. It has an interface described in a machine-processable format (specifically WSDL). Clients interact with the WS using SOAP messages in the way described by the WSDL, typically covered using HTTP with an XML serialization in conjunction with other Web-related standards" (W3C, 2004a). The SOAP specification defines the envelope structure, encoding rules, and conventions for representing remote procedure calls and responses. The WS introduce several advantages over conventional solutions, among which the following:

- Application interoperability with independence of the development platform, company policy, product vendor or even legacy systems. This independence can be achieved thanks to the use of open standards like HTTP and XML over Internet.
- Combination with other WS in order to provide more complex integrated services. There are two ways of combination: orchestration and choreography. These services could be provided by different vendors and located in several different places.
- Maximizing system flexibility, scalability and reusability of different components in a SOA environment due to the encapsulation of capabilities. For example, they could wrap complex legacy systems in enterprise organizations.

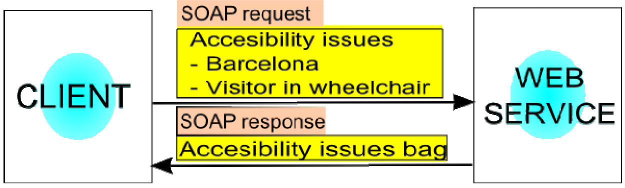

Figure 4. Client-server interaction

5.1 Web Service building

The Web Services are built in Java, using JAX-WS (Java, 2006) specification that stands for Java API for XML Web Services that is part of the Java EE platform. JAX-WS is a technology for building web services and clients where data are transferred by mean of XML documents. A remote procedure invocation and response are represented by an XML-based protocol such as SOAP and transmitted over HTTP. For example, figure 4 shows a common client-server interaction with SOAP protocol in order to get religious monuments from city of Burgos. The Web Services are deployed using the GlassFish Web Application Server.

5.2 Web Service Architecture

The both services are designed with the same software architecture (see figure 5) that is composed by three main components:

1. The Controller receives requests and sends responses to clients. It coordinates the interactions between the other two components.
2. The SPARQL manager retrieves ontology model using Jena framework and it executes a semantic query in SPARQL getting semantic relevant information for the client in its current context.
3. Finally, the SQL manager retrieves instance data from PATRAC repository with an automatically generated SQL query. The connection with the relational database is made with a typical JDBC driver.

In this way, while the SPARQL manager provides reasoning through inference on top of the ontology, the SQL manager allows accessing to individuals populating the ontology. Thus, the first returns the classes or concepts from the ontology with which the second creates the suited SQL query. This collaboration scheme between the two main components allows the recovery and provision of context aware contents and accessibility issues to clients.

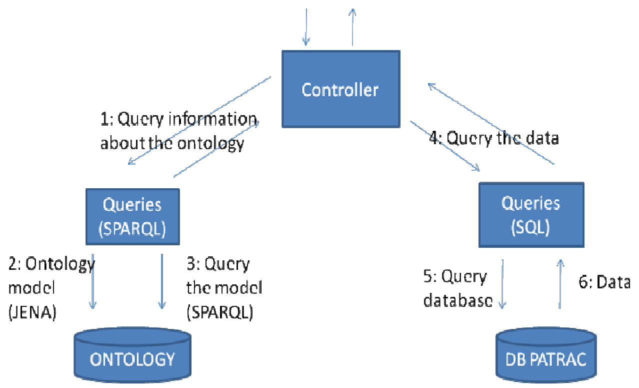

Figure 5. Web Service architecture

We illustrate the Web Service workflow with a practical example of a query about the Maritime Museum of Barcelona. The client asks the WS for the accessibility issues for visitors in a wheelchair (see figure 6). First, the controller queries for all kind of accessibility issues affecting visitors in wheelchair. Second, the PATRAC ontology model stored in relational database is recovered by Jena to execute a semantic query in SPARQL language. Thus, this query returns *stair, ramp with a slope greater than 6%* and so on, inferred from the ontology. Next, the controller runs a SQL query with the results in the database repository, through the SQL manager, in order to retrieve well known accessibility issues. Finally, the SQL query returns every item related with the monument that shows the accessibility issue.

6. CONCLUSIONS

A far-reaching goal for an efficient design and implementation in Knowledge Systems in Cultural Heritage must include an integration of Documentation, Information and Management Systems for relational databases. Georreferenced data relative to multimedia contents and corresponding metadata provide a physical support for superimposing additional contents. In this work we sketch some elements of the Ontology for Knowledge Management Systems applied to Cultural Heritage which has been developed for solving Physical and Digital Accessibility Issues in Cultural Heritage domains in the “Design for All” framework.

Our approach has been designed and implemented in the Web 3.0 framework for solving interoperability and reuse issues. Our application provides a support for technicians and disabled persons, but due to space limitations, in this work we have developed only some ideas relative to Knowledge Management Systems for providing a support for disabled persons. The developed Ontology can be applied not only to assessing Accessibility issues, but also to interventions. In a complementary paper, we develop an approach for assessing interventions to technicians following the classical distinction between non-destructive and semi-destructive techniques. In addition, the Ontology has been validated in the Maritime Museum of Barcelona and a small urban district of the historic city of Segovia (Spain).

The use of Web Services allows us the development of two clients with different technologies like J2EE and .NET working under the same semantic framework. Moreover, the creation of two databases provides logic reasoning through Jena and efficient data recovery through SQL at the same time. However this involves the execution of at least two queries, one in SPARQL and another in SQL. In the next future, ontology could include the geographical location of its concepts, allowing task such as geospatial reasoning.

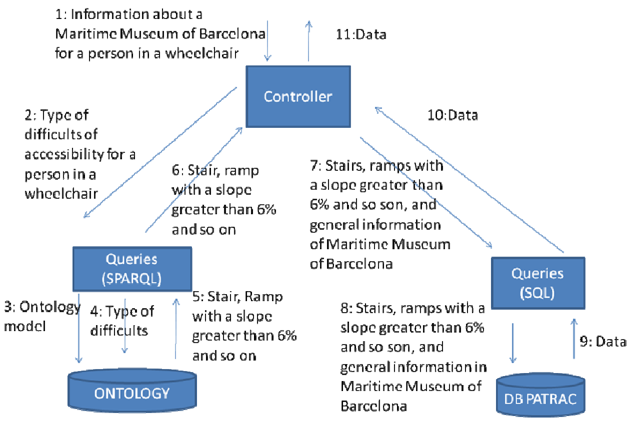

Figure 6. Web Service collaboration diagram

ACKNOWLEDGMENTS

This work is mainly supported by the "Proyecto Singular Estratégico PATRAC (Patrimonio Accesible: I+D+i para una cultura sin barreras)", PS-380000-2009-2 of the Spanish Ministry of Science and Innovation which has been co-funded with FEDER grants. The authors acknowledge also to the ADISPA (CICYT) Project with reference BIA2009-14254-C02-01 for their partial financial support. Also, we want to thanks to F.J.Delgado for the suggested comments during the peer review process.

REFERENCES

Antoniou, G., Van Harmelen, F., 2004. *A Semantic Web Primer*. MIT, London.

Barhatov Andrey aka edhel, 2006. *Protégé2Jena*. http://semweb.krasu.ru/protege2jena (accessed 26 May 2010)

Benjamins, VR, Contreras, J. et Al., 2004. Cultural heritage and the semantic web: *The Semantic Web: Research and Applications*, p.433-444, Springer.

Borst, W. N., 1997. *Construction of engineering ontologies for knowledge sharing and reuse: Ph.D. Dissertation*. University of Twente, p. 12.

DfA, 2010. Design for All Foundation. http://www.designforall.org (accessed 3 June 2010)

Doerr, M., 2003. The CIDOC conceptual reference module: an ontological approach to semantic interoperability of metadata: *AI Magazine*, 24(3)

Franz Inc., 2010. “AllegroGraph RDFStore Web 3.0’s Database”. http://www.franz.com/agraph/allegrograph (accessed 22 Feb. 2010)

Gómez-Perez, A., 1998*. A knowledge Sharing and Reuse* in The Handbook of Applied Expert Systems. Ed. J. Liebowitz, CRC Press, chapter 10.

Gómez-Pérez, A., Fernández-López, M., and Corcho, O., 2004. *Ontological Engineering with examples from the areas of knowledge Management, e-Commerce and the Semantic Web*. Springer, Londres, pp. 125-142.

Gruber, T.R., 1993. A Translation Approach to Portable Ontology Specifications. *Knowledge Acquisiton*, 5(2), pp. 199-220.

Hyvönen, E., 2009. Semantic portals for cultural heritage: *Handbook of Ontologies* (2), Springer.

Java, 2006. Java for XML-based Web Services “JAX-WS 2.0 (JSR-224)”.https://jax-ws.dev.java.net (accessed 25 May 2010)

HP Labs Semantic Web Programme, 2009. “Jena: Semantic Web Framework for Java”. http://jena.sourceforge.net (accessed 22 Feb. 2010)

Lynch, C., 2002. Digital collections, digital libraries and the digitization of cultural heritage information: *First Monday* 7(6)

Mulgara Semantic Store, 2010. “Mulgara”. http://www.mulgara.org (accesed 22 Feb. 2010)

OpenRDF.org, 2010. "Sesame: RDF Schema Querying and Storage". http://www.openrdf.org (accessed 25 May 2010)

Stanford Center for Biomedical Informatics Research, 2010. “Protégé”. http://protege.stanford.edu (accessed 29 May 2010).

Studer, R., Benjamins, VR. and Fensel, D., 1998. Knowledge engineering: principles and methods. *Data and knowledge engineering*, 25 (1-2), p. 185.

Suárez-Figueroa, M.C, Dellschaft, K., Montiel-Ponsoda, E., et. al., 2008. NeOn Deliverable d5.4.1. NeOn Methodology for Building Contextualized Ontology Networks, "Ontology Specification", chapter 5, http://www.neon-project.org/web-content/images/Publications/neon_2008_d5.4.1.pdf (accessed 20 May 2010)

Kauppinen, T. Paakkarinen, P., 2010. Geospatio-temporal Semantic Web for Cultural Heritage: *Digital Culture and E-Tourism: Technologies, Applications and Management Approaches*, IGI Global.

W3C Consortium, 2004a. W3C Working Group Note "Web Services Architecture". http://www.w3.org/TR/ws-arch (accessed 22 Feb. 2010)

W3C Consortium, 2004b. W3C Recommendation "OWL Web Ontology Language Overview". http://www.w3.org/TR/owl-features (accessed 25 May 2010)

W3C Consortium, 2004c. RDF Working Group “Resource Description Framework (RDF)”. http://www.w3.org/RDF (accessed 31 May 2010)

W3C Consortium, 2004d. W3C Recommendation “RDF Vocabulary Description Language 1.0: RDF Schema”. http://www.w3.org/TR/rdf-schema (accessed 27 May 2010)

W3C Consortium, 2008. W3C Recommendation "SPARQL Query Language for RDF". http://www.w3.org/TR/rdf-sparql-query (accessed 25 May 2010)