SEMANTIC WEB MEETS VIRTUAL MUSEUMS: THE DOMUS NATURAE PROJECT

Cristina Ghiselli, Alberto Trombetta,

Loris Bozzato, Elisabetta Binaghi

Università degli Studi dell'Insubria,Varese, Italy

Abstract

The aim of this work is to present a complex, web-based virtual museum application, integrating several tools for flexible management of heterogeneous and highly structured knowledge. All the used tools are compliant to W3C's standards. In particular, the complex network of associations and relations among concepts and objects (as typically found in a virtual museum environment) has been faithfully represented adopting W3C's Semantic Web standards. The proposed ontology allows for constraining, expressing and analyzing the intended meaning of the shared vocabulary of concepts and relations in the project domain of knowledge. As a valuable byproduct, this allows the virtual museum's visitor to interact by means of highly expressive queries.

Keywords: Virtual Museum, Semantic Web, Ontologies, XML, OWL

I. Introduction

As the next step in the evolution of the World Wide Web, the Semantic Web supports the users (not only human readers, but software agents also) to find accurate information and combine easily related pieces of information in new objects. To achieve these objectives the content of the Web has to be better structured and marked up to allow a semantic processing of information. The relevant information has to represent in a declarative and semantically way through the generation of ontology-based metadata. In fact, an ontology defines relationships among different concepts and objects used to describe and represent an area of knowledge.

While implementing the Semantic Web on the Internet is still a vision, the building blocks for the Semantic Web are being deployed in small domains and prototypes.

In this work we explore the possibilities of using Semantic Web tools for representing and querying the complex relationships occurring among data in a cultural heritage domain. These tools are used on the top of a complex, web-based virtual museum application which is part of cultural heritage project called Domus Naturae.

The project was promoted to study, classify and store information about natural species painted on the walls of the about 60 rooms of Palazzo Arese Borromeo in Cesano Maderno (a small town near Milan). As a traditional "virtual museum" the hypermedia system is composed of logically related collection of digital objects and provides various points of information access, giving the visitors a first dynamic and multidisciplinary approach to the collection. In the next section the system architecture based on new technological standards of the WWW Consortium, like XML (Extensible Markup Language), is described.

Section 3 explains the applied web semantic techniques in more detail and then the features needed to support the additional scenario. In particular we introduce the proposed ontology describing relationship among artistic, botanic and zoological multimedia data by means of OWL (Ontology Web Language), part of the growing stack of W3C recommendations related to the Semantic Web. This ontology allows the visitor to interact by means of highly expressive queries. This section presents also the ontology development process.

Finally Section 4 highlights current research and implementation directions.

II. Domus Naturae Project

Domus Naturae project is the result of a collaborative work between Computer Science and Communication Department and Biology Department of Insubria University of Varese (Italy), coordinated by Prof. Andrea Spiriti. Then the project is a team effort for integrating different disciplines, ranging from zoological, botanic to historical and artistic knowledge.
In fact, it was promoted to study, classify and store information about natural species painted on the walls of the rooms of the 17th century Palazzo Arese Borromeo.
In our approach, the project Domus-Naturae is considered as a laboratory in which new technological tools supporting cultural needs can be proposed and experimented. Consequently, the infrastructure must keep pace with the trend of technological progress in order to understand the usefulness and applicative, or methodological, innovation of the instruments proposed and to see whether they can be adapted and reused in different environments, or whether further developments can be proposed.
The role of technologists and designers is also to create a link in continuous tension between speculation and pragmatism: while considering cultural needs, they are also introducing the most abstract of ideas into the operating laboratory to conceive solutions based on available and experimental infrastructures, methodologies, and languages.
The core of the Domus Naturae project is a hypermedia system allowing the access to artistic, botanic and zoological multimedia data and their scientific analysis through different keys of reading and communication, like guided walkthroughs or highly expressive queries over the multimedia data, as well (figure 1).

[ghiselliSELECT05_fig1.gif]

Fig. 1: The Home Page of the DOMUS NATURAE Web Application

The hypermedia is articulated on various levels of navigation in order to yield an enhanced museum experience: from a panoramic virtual visit through the floors and rooms to the dynamic visualization of particulars of a wall, to the analysis of a particular with the list of its occurrences in the rooms (figure 2).

[ghiselliSELECT05_fig2.gif]

Fig. 2: Example of different levels of navigation

System architecture

According to new technological standards of the WWW Consortium, the system is based on XML (Extensible Markup Language) for representing the structure of information and on XSL (Extensible StyleSheet Language) for its presentation. Graphic representations of dynamic maps are realized with SVG (Scalar Vector Graphics).
XML is a markup language that provides a simple, standard syntax for encoding the meaning of data, standardizing a structure suitable to express semantic information for documents. It defines new languages, including new meta-languages, and permits the definition of new syntactical elements, specifying separately the content, structure, and style of representation. Although XML syntax is flexible, all XML document must conform to basic grammar rules and are interpreted by standard XML interpreters. The structure of a particular XML document can be validated with a DTD (document-type definition) or an XMLSchema. This is particularly useful in data exchange scenarios; the schema languages provide and enforce a common vocabulary for the data to be exchanged. Another important aspect is that the representation of the various elements is left to languages devised for that purpose, one of which is XSL (extensible stylesheet language). The potential of XML associated with XSL lies in the fact that, by the application different stylesheets, the same XML document can be used by different applications and can be presented in different media formats. A XSLT processor transforms a XML document with the instruction given in a style sheet, in a result document formatting into different format like XHTML or PDF. For example, SVG, a language for describing two-dimensional graphics based on XML, can also be generated via an XSLT transformation, providing a way to generate graphical representations of a wide range of XML data.
In the application the user interaction with the classification of particulars painted on the wall is improved by the use of SVG as language for graphical interfaces (figure 3).
With SVG vector graphics, images and text are grouped, styled and transformed with filter effects and it is possible to perform advanced animation and dynamic graphics.

.

[ghiselliSELECT05_fig3.gif]

Fig. 3: Example of SVG dynamic map.

The architecture of the application is divided into two basic components. One component concerns the automatic and interactive generation of XML-based data that represent the knowledge structures of the application, while the other component concerns the XSLT-based transformation of the data in the result document format, like XHTML or SVG.

III. Semantic Web techniques

Semantic Web tools are used on the top of the hypermedia system for representing and querying the complex relationships occurring among the above mentioned data.

In the Semantic Web scenario, relationships among different concepts and objects are represented by means of ontologies. In fact, the wide array of information residing on the Web and the perceived need to make it more machine-processable have acted as a strong impetus for the development of ontology languages, like RDF or OWL.

An ontology describes a formal, shared conceptualization of a particular domain of interest, for example cultural heritage objects held in art museums. In particular, an ontology allows for constraining, expressing and analyzing the intended meaning of the shared vocabulary of concepts and relations in a domain of knowledge.

We have developed an ontology describing relationship among artistic, botanic and zoological multimedia data by means of OWL (Ontology Web Language), while queries are expressed through the (far less standard) ontology query language RDQL.

As a positive remark, deploying such technologies has allowed us to express far more expressive queries than those usually possible within an RDBMS-based web application. On the other side, we have experimented the extreme inefficiency of the available Semantic Web tools, even in the execution of the simplest queries.

Furthermore, a very relevant issue regards what methodology adopts when developing an ontology. There are many proposals and none of them has been widely adopted. As such, we chose to employ an "ad-hoc" methodology comparable to the well-known ER methodology in the relational model setting.

As already said, the most relevant feature of our semantic Web-based application is its capability of expressing queries that are difficult or impossible to answer using SQL as query language. The main reason for this augmented expressive power comes from the possibility offered by the deployed Semantic Web tools of directly representing ontologies as derived from the conceptual map describing the application's overall scenario (figure 4) and from the reasoning mechanisms provided by such tools.

[ghiselliSELECT05_fig4.gif]

Fig. 4: Ontology schema describing the domain of the application.

An example of interesting query is the one who search for all the relations between two given persons. In this query we request a relation, not an object related to a particular person: this is possible primarily for the unique structure of the representation of data as triples. Similar to this is the query which looks for all the relations of a person (figure 5).

[ghiselliSELECT05_fig5.gif]

Fig. 5: Examples of queries interface.

Other queries are based on the ontology structure, like the ones which search sub-relations or subsumption between classes. Other are based on the reasoning techniques, such as the one that looks for all the relatives of a person: it uses the implicit symmetric and transitive relation defined for kinship.

An example of an “usual” SQL-like query asks for the details of a particular object and its species. Similar to the latter but more oriented to the research of hidden relations is the query for all the reference images of a book or the query for particulars of a room. A newly suggested query is the one which requests all the working places of an artist and vice versa all the artists who have worked in a specific place: it’s important to note that while this kind of queries in a relational database often require more expensive join operations to be performed, in an ontology the query system must just follow the relations between the objects.

Another interesting point relates to the ontology development process.

Such process has undergone through many steps: we started from the conceptual map and from data samples about the species represented on the paintings.

From this we developed the specifications of the prototype: we followed a development based on the ER model development for relational databases. So we developed a schema of the ontology and then converted it on the classes and properties taxonomy, adding also restrictions on classes and properties to better describe their features.

Once done this formalization of the ontology schema and having further refined it, we implemented the schema in the OWL ontology by means of the ontology editor Protegè.

Then we added part of the data (as we said, we’re working on completing the data entry) and we implemented the described queries on the semantic search application (using the RDQL ontology query language) on the top of it.

Nowadays there is no well established ontology development methodology. As a simple solution, we followed a top-down approach in order to create first the main classes, then their properties and then adding more details to the scheme adding restrictions.

III. Conclusions

The proposed ontology represents a first step towards a more complete formalization of the relationship occurring among cultural heritage objects as found in art museums.

Towards this end, the system provides a community space for field experts, which may help grow and/or amend the knowledge stored in the hypermedia system, e.g., by means of annotations. The basic idea is that the users/visitors can't only browse, but can also create shared information. Another future work will be to extend these annotation tools from simple instruments that enhance collaboration within a group of experts to complex instruments that permit authoring data by using ontology guidance.

A very relevant issue regards the problem of adopting an ontology development methodology. We chose to employ an "ad-hoc" methodology since nowadays there is no a well-established design methodology for this kind of data organization comparable, say, to the ER methodology in the relational model setting.

The experimental prototype of the multimedia system, developed and maintained by Computer Science and Communication Department of Insubria University of Varese, proves how the synergy of new technologies, communication and specific disciplines can activate processes of knowledge, exploiting the interdisciplinary nature and the potentials for interaction, which is a distictive mark of the Web.

References

Books:

Berners-Lee, T. (2002). Weaving the Web. San Francisco: Harper San Francisco.

Daconta, M. C., Obrst, L.J., Smith, K.T., (2003) The Semantic Web: a guide to the future of XML, Web Services, and Knowledge Management. Indianapolis: Wiley Publishing.

Chapters in a Book:

McGuinness , D., (2001) Ontologies Come of Age. In Dieter Fensel, Jim Hendler, Henry Lieberman, and Wolfgang Wahlster, editors. Spinning the Semantic Web: Bringing the World Wide Web to Its Full Potential. MIT Press, 2003

Journal Articles:

Gruber, T. R. (1993): A translation approach to portable ontology specifications, in Knowledge Acquisition n. 5, pp. 199-220

Schuler, D. (1994) Community Networks: Building a New Participatory Medium, Communications of the ACM, 37 (1), pp. 39-51.

Proceedings:

Patel-Schneider, F., Fensel, D.(2002) Layering the Semantic Web: Problems and Directions, in Proceedings of the First International Semantic Web Conference, Lecture Notes In Computer Science vol. 2342, Springer-Verlag, Heidelberg.

Schweibenz, W., (1998). The "Virtual Museum": New Perspectives For Museums to Present Objects and Information Using the Internet as a Knowledge Base and Communication System. In Zimmermann, Harald H./Schramm, Volker (eds., 1998): Knowledge Management und Kommunikationssysteme. Workflow Management, Multimedia, Knowledge Transfer. Proceedings des 6. Internationalen Symposiums für Informationswissenschaft (ISI '98) Prag, 3.-7. November 1998. Konstanz: UKV. S. 185-200.

Spiriti, A. (2003) Domus Naturae. Arte e scienza a Palazzo Arese Borromeo: la flora e la fauna della grande boscareccia, Comune di Cesano Maderno, Cesano Maderno.

Technical Reports:

Fikes, R., Hayes, P., Horrocks, I. (2003) OWL-QL. A Language for Deductive Query Answering on the Semantic Web, Stanford Knowledge Systems Laboratory Technical Report KSL-03-14, Stanford, CA.

Web Sites:

Adler, S. et al. (2001) Extensible Stylesheet Language (XSL) 1.0 (http://www.w3.org/TR/xsl/)

Bray, T., Paoli, J., Sperberg-Mcqueen, C. M. (2000) Extensible Markup Language (XML) 1.0 (http://www.w3.org/TR/REC-xml, 2000.)

Ferraiolo, J. et al. (Editors) (2003) Scalable Vector Graphics (SVG) 1.1 Specification (http://www.w3.org/TR/SVG11)

Horridge, M.(2004): A Practical Guide To Building OWL Ontologies With The Protégé-OWL Plugin. Edition 1.0, University of Manchester, Manchester, june 2004. (http://www.co-ode.org/resources/tutorials/ProtegeOWLTutorial.pdf.)

Stanford Medical Informatics (2004). The Protégé Ontology Editor and Knowledge Acquisition System. Version 29-12-2004. (http://protege.stanford.edu/).

McGuinness, D. van Harmelen (2004), F. OWL Web Ontology Language Overview, W3C Recommendation del 10 febbraio 2004. last updated: 10-02-2004. consulted May 2005 (http://www.w3.org/TR/2004/REC-owl-features-20040210/)

Seaborne, A (2004). RDQL. A Query Language for RDF, W3C Member Submission del 9 january 2004. last updated: 09-01-2004. consulted May 2005 (http://www.w3.org/Submission/2004/SUBM-RDQL-20040109/)

Swartz, A. (2002) The Semantic Web In Breadth. last updated: 01-05-2002. consulted May 2005 (http://logicerror.com/semanticWeb-long.)