

Modeling Aspects of Semantic Web-based E-learning System

Lect. Dr. Sabin Buraga¹

Lect. Dr. Marius Cioca²

¹Faculty of Computer Science, "Alexandru Ioan Cuza" University of Iași, Romania
busaco@infoiasi.ro

²Faculty of Engineering, "L. Blaga" University of Brașov, Romania
marius.cioca@ulbsibiu.ro

Abstract

This paper investigates the possibility to utilize standardized semantic Web-based languages, such as RDF (*Resource Description Framework*) and OWL (*Web Ontology Language*), in order to model a complex e-learning system and to improve the inter-connectivity between heterogeneous components of that system. The paper proposes an XML-based model for attaching metadata, expressing the relations between e-learning system's components, and using ontologies to express knowledge within such an application.

INTRODUCTION

The World Wide Web space has attracted a great deal of consideration as a vehicle for delivering distance courseware, in a synchronous and asynchronous mode. In any e-learning system there are certain "IT-oriented islands" – legacy applications, database management systems, client programs, etc. – hard to be integrated in a consistent semantic-based manner. The integration of such a Web-based application with its tutors and students or with other entities from different geographical areas is more difficult, for the reason that the transferred information (such as documents, multimedia streams, user's data, etc.) is very complex (Cioca, 2004).

The actual paper will investigate the possibility to use different standardized semantic Web-based languages, such as RDF – *Resource Description Framework* (Beckett, 2004) and OWL – *Web Ontology Language* (Smith, 2004), in order to model a multifaceted e-learning system and to enhance the inter-connectivity between the heterogeneous components of such a system.

FROM A TRADITIONAL E-LEARNING SYSTEM TO A SEMANTIC WEB-BASED APPLICATION

A traditional open-source e-learning Web application can consist of several distributed Web components, such as CGI (*Common Gateway Interface*) scripts written in Perl, C or Python languages, Web application servers (notably, PHP and JSP), Web services or Web agents – details in (Buraga, 2001; Buraga, 2003b; Buraga, 2004).

From the point of view of data sources, such of system implies the existence of:

- multiple database relational servers (e.g., *MySQL* or *PostgreSQL* servers),
- real-time multimedia streams (audio, video, or animated documents),
- XML-based documents (on the server, the content can be easily stored into XML documents without the layout or even within *native XML databases*, such as *eXists* or *Apache Xindice*),
- plain text files (e.g., used for storage of configuration parameters or log information).

The Web interface needs to comply with a standard for interoperability, due of multiple browsers:

- classical desktop graphical-oriented and HTML-based browsers, such as *Mozilla* or *Firefox*,
- text-oriented browsers (e.g., *Lynx* or *links*),

- WML (*Wireless Markup Language*) browsers that can be used in a wireless environment.

More over, in order to present the graphical information, the provided materials can be written in several XML-based languages. For example, to create vector-based charts, diagrams or drawings we can use SVG (*Scalable Vector Graphics*), a standardized recommendation of the World-Wide Web Consortium (W3C, 2005). Also, to design complex animations or to synchronize multiple multimedia sources a possible solution is the use of SMIL (*Synchronized Multimedia Integration Language*) – standardized specification of the World-Wide Web Consortium (W3C, 2005).

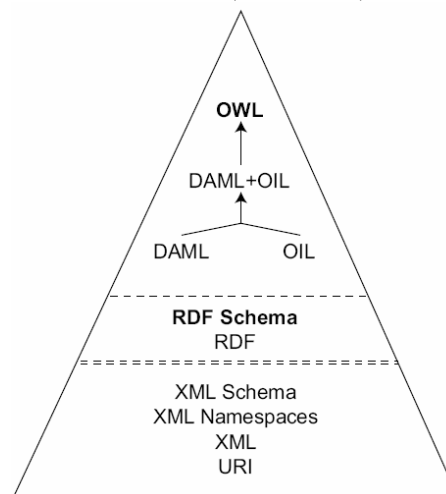


Figure 1. The pyramid of Semantic Web languages (Geroimenko, 2004)

USING SEMANTIC WEB TECHNOLOGIES TO DESCRIBE AN E-LEARNING SYSTEM

The actual WWW space is mainly compounded by pages (documents that contain markups) with information in the form of natural language text and multimedia – still images, sound, animations, video clips, etc. – intended for humans to read and to understand. Computers are principally used to render this hypermedia information, not to reason about it. Information retrieval has become omnipresent and information needs no longer to be intended for human readers only, but also for machine processing, enabling intelligent information services, personalized Web sites, and semantically empowered search engines – this is the seminal idea of *Semantic Web* (Davies *et al.*, 2003; Da Conta *et al.*, 2003; Buraga, 2004).

Semantic Web technologies are based on the XML (*Extensible Markup Language*) and is structured on three main layers (see figure 1):

- *the metadata layer* offers an extensible framework in order to express simple semantic assertions (e.g., vocabularies or taxonomies); this conceptual model can be used to attach metadata (data about data) to each Web resource;
- *the schema layer* can help to specify simple ontologies in order to define a hierarchical description of the concepts and properties for a given resource;
- *the logical layer* introduces ontological languages that are capable to model complex ontologies; at this layer, in the future will exist different *reasoning services* to be used by the applications oriented to Semantic Web.

XML Technologies

Any semantic Web-based application is based on XML – Extensible Markup Language (Bray, 2004; Buraga, 2001) language, a recommendation of the World Wide Web Consortium (W3C, 2005) for a meta-language to define mark-ups (annotations) for content publishing particularly on the World Wide Web space.

The main objective of the XML meta-language is to provide some benefits not available in HTML (HyperText Markup Language), such as arbitrary extensions of a document's elements (tags) and their attributes, support for documents with complex structure, and validation of document structure with respect to an optional document-structure grammar, called a *DTD (Document Type Definition)*. Also, instead of DTD, an object-oriented method for validation of XML documents can be used: an *XML Schema* (Geroimenko, 2004; W3C, 2005).

As a standard recommended by the Web Consortium, XML is considered as the data format for information interchanging between various Internet and Web applications. The XML popularity is primarily due to its flexibility in the representation of many data types (see below). The uses of mark-ups give to the XML language the possibility of self-description, and its extensible nature makes possible the definition of new document types, with a particular destination (e.g. user profiles, business rules, multimedia, data-flow etc.).

Using XML, the semantics and the structure of the data exchanged by diverse Web business applications is preserved. One of the key advantages is that the data can be organized as in an object-oriented database. As XML is format-independent, there is possible to generate multiple – XHTML, SMIL, WML or XUL – outputs smoothly by transforming XML documents via *XSL (Extensible Stylesheet Language)* constructs (Buraga, 2001; W3C, 2005). Similarly to the *CSS (Cascading Style Sheets)*, the XSL documents separate the content from representation.

Since 1998, XML has grown into a great family of standards integrating key technologies from three previously independent domains: documents, databases, and the Internet.

Some examples of XML-based languages are (Geroimenko, 2004):

- *MathML (Mathematics Markup Language)*,
- *SMIL (Synchronized Multimedia Integration Language)*,
- *RDF (Resource Description Framework)*,
- *XUL (Extensible User-interface Language)*,
- *WML (Wireless Markup Language)*.

In order to move towards the Semantic Web, there were developed a series of XML-based languages specialized in the modeling of knowledge – for example, *RDF (Resource Description Framework)* and *OWL (Web Ontology Language)*.

From the authors' point of view, these languages could be the best solution to store and process various information within an e-learning system.

Use of Metadata

Each component of the e-learning system can be described with the help of metadata. The *metadata level* is the first level of a semantic Web-based application (Buraga, 2004). This metadata can be attached to each software component of the e-learning system in order to store several important characteristics (e.g., information regarding uptime, ownership, execution platform, etc.).

Also, for each user we can retain the information about his/her status. For example, we can store the user role – administrator, database manager, security monitor, regular user (tutor, student or visitor etc. Also, the system can retain personal data (e.g., age, user e-mail address, location, etc.), and user-interface preferences (layout, chromatic and interaction preferences, etc.).

To associate and store metadata, we use RDF (Beckett, 2004; Buraga, 2004) – an XML-based model for processing metadata. RDF standard provides interoperability between applications that exchange machine-understandable information on the World-Wide Web. RDF is intended to be used to capture and express the conceptual structure of information offered by the Web.

As suggested in (Buraga, 2003a), RDF metadata can also describe client's session information:

- host and user profiles (tutors, students);
- user profiles of participants of previous and current sessions;
- session profiles (e.g., timing, available resources, navigation history, etc.);
- private user applications and resources (viewers, plugins, content editors, etc.).

For example, the list of the on-line courses accessed by a particular student is expressed by several RDF constructs (Buraga, 2003a):

```
<rdf:RDF>
  <rdf:Bag ID="courses">
    <!-- Web Technologies course -->
    <rdf:li
      rdf:resource=
        "http://www.infoiasi.ro/courses/web" />
    <!-- Operating Systems course -->
    <rdf:li
      rdf:resource=
        "http://www.cs.pub.ro/teach/os" />
    <!-- Perl Programming course -->
    <rdf:li
      rdf:resource=
        "http://www.infoiasi.ro/perl_prog" />
    <!-- other courses... -->
  </rdf:Bag>
  <rdf:Description rdf:about="#courses">
    <t:User t:id="33" t:type="student">
    <rdf:Description
      rdf:about=
        "http://students.infoiasi.ro/~stud">
    <t:Name> ... </t:Name>
    <t:Year> 4 </t:Year>
    <!-- other useful information -->
  </rdf:Description>
</rdf:RDF>
```

The namespace prefix *t* refers to a specific namespace prefix chosen by the author of the RDF expression and defined in an XML namespace declaration such as *xmlns:t="http://some.host.com/xml-schema"*.

The *rdf* namespace is defined by the World-Wide Web Consortium to be specified in every RDF statement. The XML namespaces are used to avoid parsing conflicts for identical elements or attributes names included in the same XML document (Geroimenko, 2004; Buraga, 2001).

Also, RDF can be used to express a high-level platform-independent model for accessing and discovering hypermedia resources of an Internet-based distributed system.

A Model for Expressing the Relations between E-learning System's Components

The relationships established between the components of the system can be also expressed by RDF assertions. For example, we can trace a relation of dependence between different e-learning modules (e.g., user authentication module, workflow monitor, student evaluation module, etc.) and the involved resources. The relations are useful to maintain the connection between the e-learning application and its users. Using FOAF – *Friend Of A Friend* (FOAF, 2005) statements, we can create the graph of inter-connections between the persons that participate into a task or use a certain component (such as the courseware management software tool or the course annotator module).

Using Ontologies

A superior level of modeling is to create or (re)use *ontologies* to represent the knowledge within the application. The semantic structure achieved by ontologies differs from the superficial composition and formatting of information (viewed as data) afforded by relational and native XML databases. Ontologies are able to provide an objective specification of domain information (in our case, the e-learning domain) by representing a consensual agreement of the concepts, characteristics, and relations characterizing the way knowledge in that domain is expressed (Daum & Merten, 2003; Si Shi & Murthy, 2003).

The RDF – *Resource Description Framework* (Beckett, 2004) and OWL – *Web Ontology Language* (Smith *et al.*, 2004) facilities are significant in modeling of a knowledge-based e-learning system.

Using RDF and OWL statements, we can represent – in a standardized way – several information such as (Buraga, 2003a, Buraga & Cioca, 2003):

- instructional content divided into different modules for multiple use and re-use,
- abstract pedagogical entities,
- tutor and student profiles,
- domain, pedagogical, and student knowledge bases (facts and rules) used by the system's inference engines.

Also, the e-learning system can use several upper-level ontologies – such as the *ABC ontology* (ABC, 2001) – that can be helpful to model concepts, properties and relations between these concepts.

The ABC ontology has been designed as a primitive (general) ontology so that individual communities are able to build on top of it more specific ontological constructs. The primitive category at the core of the ABC ontology is an entity. At the next level three main categories are offered: *temporality*, *actuality* and *abstraction*.

Inter-connectivity of the E-learning System's components

We can also use certain ontological statements expressed by RDF and/or OWL languages to improve the inter-connectivity between the components of the given application.

One of the major concerns of the e-learning application's analysts, designers and programmers is to use a platform- and language-independent protocol for exchanging information between the software components (scripts, agents, Web services, etc.) of the e-learning system or between these components and other entities, i.e. final users (tutors, students, etc.), administrators, processes, modules – see also (Buraga, 2003) and (Buraga & Cioca, 2003).

The interactions between the users or between the users and the software components comply with a standard for interoperability (Buraga & Alboae, 2004). One solution is to consider an XML-based protocol, such as SOAP – *Simple Object Access Protocol* (W3C, 2005). SOAP is a simple lightweight protocol that can be used for structured and strong-type information exchange in a decentralized, distributed environment. The SOAP messages can encapsulate “intelligent” data expressed by XML constructs (metadata, schemas, ontologies, etc.).

Using the process ontology described by the OWL-S specification (OWLS, 2004), the relations between the components of the e-learning application can be easily modeled and can be integrated on the semantic basis.

CONCLUSION

The actual paper presented a manner of using semantic Web-based standardized languages for modeling the information within an e-learning system. The main focus of the paper was to use RDF and OWL languages for attaching metadata, describing the relationships between e-learning system's components, and using ontologies to denote knowledge within such a Web-based application.

The XML family of languages was proposed to be adopted for information exchanging and interoperability. XML is considered the best solution for information interchange between diverse components of the e-learning system and for semantic representation of data.

An agent-based e-learning system that can use the proposed model will be considered as a further direction of research, following (Buraga, 2003a) and (Buraga & Cioca, 2004).

References

1. Beckett, D. (ed.), *RDF/XML Syntax Specification (Revised)*, W3C Recommendation, Boston, 2004: <http://www.w3.org/TR/rdf-syntax-grammar>
2. Bray, T. *et al.* (eds.), *Extensible Markup Language (XML) 1.0*, W3C Recommendation, Boston (2004): <http://www.w3.org/TR/REC-xml>
3. Buraga, S., Developing Agent-Oriented E-Learning Systems, *Proceedings of The 14th International Conference on Control Systems and Computer Science – vol.II*, I.Dumitrache and C.Buiu (eds.), Politehnica Press, Bucharest (2003).
4. Buraga, S., *Tehnologii Web* (in Romanian), Matrix Rom, București (2001).
5. Buraga, S., *Semantic Web* (in Romanian), Matrix Rom, București (2004).
6. Buraga, S. (ed.), *Aplicații Web la cheie* (in Romanian), Polirom, Iași (2003).

7. Buraga, S., Alboaie, L., Alboaie, S., *An XML-based Serialization of Information Exchanged by Software Agents*. International Informatica Journal, 28, 1 (2004).
8. Buraga, S., Cioca, M., Open Methodologies for Developing Web-based E-Learning Systems. *Proceedings of the International Conference on Manufacturing Science and Education Challenges of the European Integration*, "L.Bлага" University Press, Sibiu (2003).
9. Buraga, S., Cioca, M., Using Open Methodologies to Build Web-based E-learning Systems, *Proceedings of the International Symposium on European Integration*, ASEM Press, Chisinau (2004).
10. Cioca, M., *Conducerea asistată a unităților economice* (in Romanian), Editura Universității "L. Blaga", Sibiu (2004).
11. Daconta, M. *et al.*, *The Semantic Web*, Wiley & Sons (2003).
12. Daum, B. Merten, U., *System Architecture with XML*, Elsevier Science (2003).
13. Davies, J. *et al.* (eds.), *Towards the Semantic Web*, Wiley & Sons (2003).
14. Geroimenko, V., *Dictionary of XML Technologies and the Semantic Web*, Springer (2004).
15. Si Shi, N., Murthy, V.K., *Architectural Issues of Web-Enabled Electronic Business*, Idea Group Publishing (2003).
16. Smith, M. *et al.* (eds.), *OWL Web Ontology Language Guide*, W3C Recommendation, Boston (2004): <http://www.w3.org/TR/owl-guide/>
17. * * *, *ABC Modeling Examples*: http://metadata.net/harmony/cimi_modelling.htm
18. * * *, *FOAF (Friend Of A Friend) Specification*: <http://rdfweb.org/>
19. * * *, *OWL-S Specification*: <http://www.daml.org/services>
20. * * *, *World-Wide Web Consortium* (2005): <http://www.w3.org/>