Ontological Library Generator for Hypermedia-Based E-Learning System

Eugen Zaharescu¹, Georgeta-Atena Zaharescu²

(1) "Ovidius” University of Constanta, Mathematics and Informatics Faculty
124 Mamaia Blvd., Constanta 900527, Romania
E-mail: ezaharescu@univ-ovidius.ro http://math.univ-ovidius.ro/
(2) "DECEBAL" High School, Constanta, Romania
E-mail: aza@math.com

Abstract
This paper presents an automatically approach of metadata e-library generation providing online access to very large and organized video tutorials collections, covering the main aspects of e-learning processes. This complex metadata digital library defines the structures of a large hypermedia LU (Learning Units) database embedded in a hypermedia-based e-learning system. Different methods of semantic description and hypermedia educational contents integration are explored. Furthermore, some possibilities to build a large-scale hypermedia objects ontology based on lexical resources generated in the context of OWL (Web Ontology Language) systems (e.g. Protégé-OWL) are also described. Ontology-based e-library generator is designed in the context of Semantic Web technologies using a SOA (Service Oriented Architecture) approach. It provides e-learning management system with large hypermedia resources repositories and enables efficient knowledge reuse and exchange between e-universities.

Keywords: Semantic E-Learning, metadata, hypermedia, Semantic Web, Web Ontology Language, Service Oriented Architecture.

Introduction
The exploitation of new Semantic Web technologies in the context of E-learning requires a deeper understanding of the relevant issues as long as they will be able to incorporate even perception and pervasive or ubiquitous computing.

The ultimate objective of the Semantic Web research activities targets the improvement of the human experience and the enrichment of the living, with better ability to use heterogeneous content and knowledge applications.

Semantic Web Based E-learning
From the beginning, we have tried to summarize some key research themes in the convergence of Semantic Web and E-learning as shown in figure 1. Also, a set of research priorities are revealed here and more specifically, there are three cyclical areas that summarize the current research in Semantic E-Learning.

In this visual description it is used a matching of key issues that have significant roles in Semantic Web and E-learning research, respectively. They are presented as five pairs where the first part relates to the Semantic Web key issue and the second one to the E-learning key issue:

1. "Expression of Meaning"—"Content authoring"
2. "Policy Aware Infrastructure"—"Interoperability/Standards"
3. “Ontological Evolution”–“Adaptive Hypermedia”
4. “Web of Trust”–“Communities/Social Dimensions”
5. “Information Flow and Collaborative Life”–“Learning Context”.

Figure 1 The Semantic E-Learning research themes

"Expression of Meaning"–“Content authoring”. The obvious direct relation of Semantic Web and E-learning combines the traditional content authoring process with the critical objective of expression of meaning. Issues like Semantic Mark-Up, Semantic Retrieval, Personalized and (Semi)-Structured Annotation and Content Conversion are leading a big research stream, in which the main concern is the development of Semantic E-Learning content.

“Policy Aware Infrastructure”–“Interoperability/Standards”. The E-learning industry has many achievements in the area of interoperability and standards and it recognizes the need to secure a policy-aware infrastructure. The Semantic Web will only achieve its potential as an information space for the free flow of scientific and cultural information if its infrastructure supports a full range of fine-grained policy controls over its content. The research on types of Control Over Content, the Compliance To Semantic and Metadata Models as well as the issues of versioning and provenance require extensive research.

“Ontological Evolution”–“Adaptive Hypermedia”. The traditional Adaptive Hypermedia considerations in E-learning are combined with Ontological Engineering and a lot of flexible systems and accompanied methodologies have emerged. Issues like Ontology-Building, Ontology-Integration, Conceptual Modelling and Semantic Conceptualisation reveal a new research agenda, in which the specifications of conceptualisations (ontologies) promote the performance of learning systems.

“Web of Trust”–“Communities/Social Dimensions”. In the E-learning Industry this issue is of critical importance. The development of Virtual Learning Communities will require a Semantic Web language of describing trust in the form of Unique Identities of Resources and Intelligent Assistants.

“Information flow and collaborative Life”–“Learning Context”. As mentioned above, the instrumentation of knowledge flows has been set as one of the priorities of the SW W3C activity. In this area Semantic Services, (Semi) Automated Reasoning and Argumentation are critical themes on the semantic e-learning agenda.
Table 1 E-Learning requirements fulfilled by Semantic Web Agents conceptual characteristics

<table>
<thead>
<tr>
<th>E-Learning Requirements</th>
<th>Semantic Web Agents conceptual characteristics</th>
</tr>
</thead>
<tbody>
<tr>
<td>Distributed Knowledge</td>
<td>The Semantic Web will be as decentralized as possible. Distributed nature of the Semantic Web enables continuous improvement of learning materials and effective co-operative content management.</td>
</tr>
<tr>
<td>Coordinated Interactivity</td>
<td>On the Semantic Web, software agents’ activities are coordinated as they may use commonly agreed service language and produce proactive delivery of updated learning materials. The vision is that each user has his own personalized Semantic Web agent that communicates with other agents to generate the answer.</td>
</tr>
<tr>
<td>Non-linear Delivery</td>
<td>Learning materials are distributed on the Web as linked objects to agreed ontologies. This enables construction of a user-specific course by semantic querying for topics of interest.</td>
</tr>
<tr>
<td>Dynamic Delivery</td>
<td>Based on personalized Semantic Web agents, the delivery of information will be proactive, creating a dynamic learning environment. The Semantic Web enables the use of knowledge provided in various forms by semantically annotated content.</td>
</tr>
<tr>
<td>Personalized Access</td>
<td>According to his own profile, user can describe goal of learning based on previous knowledge and perform semantic querying for the suitable learning material. The ontology is the link between the user needs and the characteristics of the learning material. Access to knowledge can be expanded by semantically defined navigation.</td>
</tr>
<tr>
<td>Integration</td>
<td>The Semantic Web offers the potential to become an integration platform for all learning activities in any organization.</td>
</tr>
</tbody>
</table>

Semantic Web Stack and Ontology Spectrum

The Semantic Web stack (proposed and gradually refined by Berners-Lee, 2003, figure 2) guides us through the process of increasing level of semantics.

Resources are at the basis of semantics, identified via their Uniform Resource Identifier (URI). The next semantic layer is the XML, a set of syntax rules for “creating semantically rich mark-up languages in a particular domain” (Daconta et al., 2003) together with its NS-Namespaces (“a simple mechanism for creating globally unique names for the elements and attributes of the mark-up language”, to avoid vocabulary conflicts). On top of XML is the Resource Description Framework (RDF), simply put, an XML language to describe whole resources (as opposed to only parts of them, as with XML). RDF Schema is a language that enables the creation of RDF vocabularies; RDF Schema is based on an object-oriented approach.

Figure 2 Semantic Web stack (Berners-Lee, 2003) – on the left side and Ontology Spectrum (Daconta et al., 2003) – on the right side
Semantics increases from the lower levels towards the top of the stack. **Ontologies** are constructed from structured vocabularies and their meanings, together with explicit, expressive and well-defined semantics. In particular, **Ontologies** make knowledge reusable by featuring classes (general things), instances (particular things), relationships between those things, properties for those things (with their values), functions involving those things and constraints on and rules involving those things.

**Ontologies** have their own spectrum of increasing semantics, as described in figure 2 (Daconta et al., 2003). **Taxonomies** contain structured data, where the semantics of the relationship between a parent and a child node is not well specified (“can be subclass of or part of”). **Thesauri** are controlled vocabularies, with clearly defined equivalence, homographic (the same spelling), hierarchical and associative relationships (e.g. **WordNet**). A **Conceptual Model** permits class-subclass hierarchies (as in **UML**). **Logical Local Domain** theories are directly interpretable semantically by the software, and represent the highest aspiration for ontologies.

**Distributed Learning Objects Metadata**

The distributed learning technologies and the learning objects standardization was developed by three major organizations:

- **Alliance of Remote Instructional Authoring and Distribution Networks for Europe (ARIADNE)** focused on metadata and learning object indexing systems;
- **IMS Global Learning Consortium (IMS-GLC)**, developed vocabularies and metadata for learning objects (**IMS Learning Resource Metadata specification**);
- **Advanced Distributed Learning (ADL)** realised the **Sharable Content Object Reference Model (SCORM)**, a web-oriented data model for content aggregation focusing on the structure and run-time environment for learning objects.

**SCORM (Sharable Content Object Reference Model)** is accepted as the standard for the **Educational Content Management** and represents a collection of specifications for web-based E-Learning. **SCORM** uses XML heavily, especially in defining **Course Structure Format**, a system for representing course structures so that educational material can interoperate between platforms and systems. It defines communications between client side content and a host system called the **run-time environment**, which is commonly supported by a **Learning Management System**. It also defines how content may be packaged into a transferable ZIP file called "**Package Interchange Format**".

**SCORM 2004** introduced a complex idea called **sequencing**, which is a set of rules that specifies the order in which a learner may experience content objects. These rules constrain the learner to a fixed set of paths through the training material, permit the learner to “bookmark” their progress when taking breaks, and assure the acceptability of test scores achieved by the learner.

**Learning Object Metadata (LOM)** from IEEE LTSC represents a standard for **Educational Content Metadata Management** and is a data model, usually encoded in XML, used to describe a learning object and similar digital resources used to support learning. The purpose of **Learning Object Metadata** is to support the reusability of learning objects, to aid discoverability and to facilitate their interoperability, usually in the context of online **Learning Management Systems (LMS)**.

**LOM** defines a hierarchy of data elements for learning objects metadata named **Base Schema**. At the top level of the hierarchy there are nine categories and for each data element, **LOM** specifies a name, explanation, size, example value, data type, and other key details: **1-General, 2-Lifecycle, 3-Meta-Metadata, 4-Technical, 5-Educational, 6-Rights, 7-Relation, 8-Annotation, 9-Classification**
Authoring Adaptive Educational Hypermedia

Adaptive Educational Hypermedia (AEH) is dedicated to personalization of distributed learning materials in the open hypermedia corpus, e.g., the WWW. The most important goal is make an easier authoring process (“authoring once, delivering many”) with two major possible approaches: first, a common language used by all authors of AEH, and secondly, the use of converters between AEHs.

We can describe several approaches of Adaptive Educational Hypermedia systems:

- **AHA! “Adaptive Hypermedia Architecture”**, was originally developed to support on-line courses with some user guidance through conditional (extra) explanations and conditional hided links. AHA! has many extensions and tools that have turned the system into a versatile adaptive hypermedia platform. AHA! can be used to add different adaptive “features” to applications such as on-line courses, museum sites, encyclopaedias, etc.

- **InterBook** is a system for authoring and delivering adaptive electronic textbooks on the WWW providing a technology for developing electronic textbooks from a plain text to a specially annotated HTML, *Adaptive Hypertext and Hypermedia*. An HTTP server for adaptive delivery of these electronic textbooks over WWW is also provided. Adaptive navigation support techniques applied in InterBook proved to be efficient for educational applications of hypertext and hypermedia, transforming them in an intelligent learning support media. InterBook is used to deliver adaptive Web-based courses on “*ACT-R theory of cognitive modelling*” Carnegie Mellon University, USA.

- **MOT (My Online Teacher)** is a general authoring system for adaptive hypermedia. MOT can author for different adaptation engines. To achieve this, MOT is exporting to a generic format, called CAF, which realises the static representation of the data. Together with the adaptation strategy, written in the adaptation language, LAG, this system can provide specifications of adaptation for various types of user-model and presentation-model related adaptations.

- **Claroline** is an Open Source E-Learning and E-Working platform allowing teachers to build effective online courses and to manage learning and collaborative activities on the web. Claroline has a large worldwide users’ and developers’ community.

- **WHURLE (Web-based Hierarchical Universal Reactive Learning Environment)** is an adaptive learning environment, which is pedagogically effective, suitable to learner needs and all subjects. WHURLE is implemented using XSLT, is designed for Coccon 1.x. (Java publishing framework) and is developed as a research tool in the Web Technology Group of the University of Nottingham.

Semantic E-Learning Conceptual Platform Architecture

In this section, we present a conceptual Semantic E-Learning architecture which provides high-level services for appropriate online information retrieving.

This architecture integrates semantic services like: semantic browsing, semantic search or smart question answering and is structured on three levels: (i)-Access Interface Level, (ii)-Service Manager Level and (iii)-Knowledge Base Level.

The very top level of this architecture is the Access Interface Level representing the integrated interface with the User and Provider Category. Through this personalized interface the learners, the readers as well as the authors / managers of the academic institutions can access, upload or modify the data with particular authority.

The second level, Service Manager Level, will generate a very complex and personalized set of services for each interacting actor (particular searches, notification service, course annotation, etc.).
Between this second level and the third/last level we can observe two key elements:

(i) Search Engine and (ii) Inference Engine

Search Engine provides an API with methods for querying the knowledge base. RDQL (RDF Data Query Language) can be used as an ontology query language. Also, it integrates Ontological Hypermedia Library Generator.

Inference Engine answers to very complex queries and is responsible for inferring new facts by an intelligent combination of facts already have in the knowledge base.

The fundamental and core level in this Semantic E-Learning platform is Knowledge Base Level that will manage the conceptual elements of the whole architecture. In fact, it is a repository where ontologies, metadata, inference rules, educational resources and course descriptions, user profiles are stored.
Ontological Hypermedia Library Generator

Ontology-based e-library generator is based on a SOA (Service Oriented Architecture) design and is integrated in the Search Engine module of the Semantic E-Learning architecture. It uses Google Video and Image Search agents to retrieve from the WWW the hypermedia elements corresponding to a given ontology or taxonomy generated with WordNet or other similar systems. Finally, these hypermedia objects are stored in a large LU (Learning Units) repository. Subsequently, it provides a complex metadata digital library that defines the structures of this large hypermedia reusable LU database embedded in the Semantic E-Learning system. This metadata digital library uses Learning Object Metadata (LOM) from IEEE LTSC as a standard for the management of hypermedia e-learning objects as shown in the figure 4.

Figure 4  LOM standard representation of a hypermedia e-learning object

![LOM standard representation of a hypermedia e-learning object](image1)

Acknowledgment

This paper was accomplished as part of the research project no. 551/2009 granted by Romanian CNCSIS.
References

Books:  

Journal Articles:  

Conference Proceedings:  

Internet Sources:  