eLearning based on the Semantic Web

Ljiljana Stojanovic
FZI Research Center for Information Technologies
at the University of Karlsruhe, 76131 Karlsruhe, Germany
Ljiljana.Stojanovic@fzi.de

Steffen Staab
Institute AIFB - University of Karlsruhe and Ontoprise GmbH
76128 Karlsruhe, Germany
Staab@aifb.uni-karlsruhe.de

Rudi Studer
Institute AIFB - University of Karlsruhe, FZI Research Center
for Information Technologies and Ontoprise GmbH, 76128 Karlsruhe, Germany
Studer@aifb.uni-karlsruhe.de

Abstract: eLearning is fast, relevant and just-in-time learning grown from the learning-requirements of the new, dynamically changing, distributed business world. The term „Semantic Web" encompasses efforts to build a new WWW architecture that supports content with formal semantics, which enables better possibilities for searching and navigating through the cyberspace. As such, the Semantic Web represents a promising technology for realizing eLearning requirements.

This paper presents an approach for implementing the eLearning scenario using Semantic Web technologies. It is primarily based on ontology-based descriptions of content, context and structure of the learning materials and benefits the providing of and accessing to the learning materials.

Introduction

It is clear that new styles of learning are some of the next challenges for every industry. Learning is a critical support mechanism for organizations to compete not only from the point of view of education, but also from the point of view of the new economy (Drucker 2000). Incredible velocity and volatility of today's markets require just-in-time methods for supporting the need-to-know of employees, partners and distribution paths. It is also clear that this new style of learning will be driven by the requirements of the new economy: fast, just-in-time and relevant.

Time, or the lack of it, is the reason given by most businesses for failing to invest in learning. Therefore, learning processes need to be fast and just-in-time. Speed requires not only a suitable content of the learning material (highly specified, not too general), but also a powerful mechanism for organizing such material. Also, learning must be a customized on-line service, initiated by user profiles and business demands. In addition, it must be integrated into day-to-day work patterns and needs to represent a clear competitive edge for the business. Learning needs to be relevant to the (semantic) context of the business (Adelsberger et al. 2001).

eLearning aims at replacing old-fashioned time/place/content predetermined learning with a just-in-time/at-work-place/customized/on-demand process of learning. It builds on several pillars, viz. management, culture and IT (Maurer&Sapper 2001). eLearning needs management support (vision and plan for learning, integrating learning into daily work). It requires changes in organizational behavior establishing a culture of "learn in the morning, do in the afternoon". An IT platform, which enables efficient implementation of learning infrastructure, is also needed. Our focus here lies on IT (web) technology that enable fast, just-in-time and relevant learning. Current web based solutions don't meet the above mentioned requirements. Some pitfalls are e.g. information overload, lack of accurate information, content that is not machine-understandable.

The new generation of the web, the so-called Semantic Web, appears as a promising technology for implementing eLearning. The Semantic Web constitutes an environment in which human and machine agents will communicate on a semantic basis (Berners-Lee 2000). One of its primary characteristics, viz. shared understanding based on the ontology backbone. Ontology enables the organization of learning materials around small pieces of semantically annotated (enriched) learning objects (Neidl 2001). Items can be easily organized into customized learning courses (fast and just-in-time) and delivered on demand to the user, according to her/his profile and business needs (relevant).

The paper will outline how the Semantic Web can be used as a technology for realizing a sophisticated eLearning scenarios. In the following, we will first sketch requirements for eLearning. Thereafter, we analyze the requirements the Semantic Web puts on representational structures (common semantic, machine-processable and -understandable data) and discuss layers of the Semantic Web architecture. In the subsequent section the advantages of using ontologies for describing eLearning materials are presented. We continue with a description of an ontology-based solution for eLearning. After a discussion of related work, concluding remarks summarize the importance of the presented topic and outline some future work.

eLearning and eLearning requirements

"eLearning is just-in-time education integrated with high velocity value chains. It is the delivery of individualized, comprehensive, dynamic learning content in real time, aiding the development of communities of knowledge, linking learners and practitioners with experts" (Drucker 2000).

Standard or traditional learning process could be characterised with centralised authority (content is selected by the educator), strong push delivery (instructors push knowledge to students), lack of a personalisation (content must satisfy the needs of many) and the linear/static learning process (unchanged content). A detailed view on standard learning is given in Tab.1. The consequences of such organisation on the learning are expensive, slow and too unfocused (problem-independent) learning process. But dynamically changed business environment puts completely different challenges on learning process – fast, just-in-time (cheap) and relevant (problem-dependent) learning, as mentioned in first section. This can be solved with the distributed, student-oriented, personalised, non-linear/dynamic learning process – eLearning. Tab. 1 shows the characteristics (or pitfalls) of the standard learning and improvements achieved using the eLearning environment. These are also the most important characteristics of eLearning.

Dimension	**Training**	**eLearning**
Delivery	Push – Instructor determines agenda	Pull – Student determines agenda
Responsiveness	Anticipatory – Assumes to know the problem	Reactionary – Responds to problem at hand
Access	Linear – Has defined progression of knowledge	Non-linear – Allows direct access to knowledge in whatever sequence makes sense to the situation at hand
Symmetry	Asymmetric – Training occurs as a separate activity	Symmetric – Learning occurs as an integrated activity
Modality	Discrete – Training takes place in dedicated chunks with defined starts and stops	Continuous – Learning runs in the parallel loops and never stops
Authority	Centralized – Content is selected from a library of materials developed by the educator	Distributed – Content comes from the interaction of the participants and the educators
Personalization	Mass produced – Content must satisfy the needs of many	Personalized – Content is determined by the individual user's needs and aims to satisfy the needs of every user
Adaptivity	Static – Content and organization/taxonomy remains in their original authored form without regard to environmental changes	Dynamic – Content changes constantly through user input, experiences, new practices, business rules and heuristics

Table 1 Differences between training and eLearning (Drucker 2000)

The principle behind eLearning is that the tools and knowledge needed to perform work are moved to the workers – wherever and whoever they are. Simply put, eLearning revolves around people. This is in stark contrast to the way learning has typically involved people flocking around the learning, i.e. a typical scholastic environment.

eLearning has its origins in computer-based training (CBT), which was an attempt to automate education, replace a paid instructor, and develop self-paced learning. But, focus of eLearning is not only on education (or recorded education – as in CBT), but also on education without barriers of time and distance, and customized to user's and business' needs (Barker 2000). Key to success is the ability to reduce the cycle time for learning and to adapt "content, size and style" of learning to a user and to the business.

Semantic Web architecture - XML, RDF and Ontologies

The term „Semantic Web" encompasses efforts to build a new WWW architecture that supports content with formal semantics. That means, content suitable for automated systems to consume, as opposed to content intended for human consumption. This will enable automated agents to reason about Web content, and produce an intelligent response to unforeseen situations.

Layers of the Semantic Web

"Expressing meaning" is the main task of the Semantic Web. In order to achieve that several layers are needed. They are presented in the figure 1 (Berners-Lee 2000), among which the following layers are the basic ones:

- the XML layer, which represents data;

- the RDF layer, which represents the meaning of data;
- the Ontology layer, which represents the formal common agreement about meaning of data;
- the Logic layer, which enables intelligent reasoning with meaningful data.

It is worth to note that the real power of the Semantic Web will be realized when people create many systems that collect Web content from diverse sources, process the information and exchange the results with other human or machine agents. Thereby, the effectiveness of the Semantic Web will increase drastically as more machine-readable Web content and automated services (including other agents) become available. This level of inter-agent communication will require the exchange of "proofs". Two important technologies for developing the Semantic Web are already in place: eXtensible Markup Language (XML) and the Resource Description Framework (RDF).

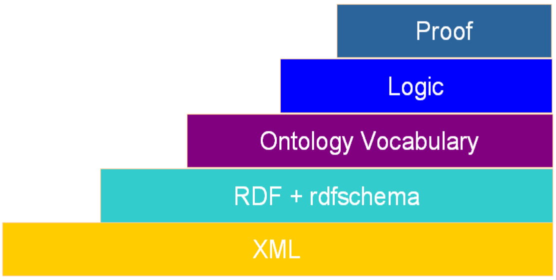

Figure 1 Layers of the Semantic Web architecture

XML (http://www.w3.org/XML/) lets everyone create their own tags that annotate Web pages or sections of text on a page. Programs can make use of these tags in sophisticated ways, but the programmer has to know what the page writer uses each tag for. In short, XML allows users to add arbitrary structure to their documents but says nothing about what the structures mean (Erdmann & Studer 2000). Meaning of XML-documents is intuitively clear, due to "semantic" mark-up and tags, which are domain-terms. However, computers do not have intuition. Tag-names per se do not provide semantics.

DTDs are a possibility to structure content of the documents. However, structure and semantics are not always aligned, they can be orthogonal. A DTD is not appropriate as a semantic language. The same holds for XML-Schema (http://www.w3.org/XML/Schema) – it only defines structure, though with a richer language. XML lacks a semantic model, it has only a "surface model", a tree. So, XML is not the solution for propagating semantics through the Semantic Web. It only has the role as a "transport mechanism", viz. as an easily machine-processable data format.

The Resource Description Framework (RDF) (http://www.xml.com/xml /pub/98/06/rdf.html) provides a means for adding semantics to a document. RDF is an infrastructure that enables encoding, exchange and reuse of structured metadata (described later). Principally, information is stored in the form of RDF statements, which are machine understandable. Search engines, intelligent agents, information broker, browsers and human users can understand and use that semantic information. RDF is implementation independent and may be serialized in XML (i.e., its syntax is defined in XML). A process in which semantic information is added to the web documents is called semantic annotation (Handschuh et al. 2001). RDF, in combination with RDFS (http://www.w3.org/TR/PR-rdf-schema/), offers modeling primitives that can be extended according to need. Basic class hierarchies and relations between classes and objects are expressible in RDFS. However, the model suffers from a lack of distinction between object and meta level, which makes it unintuitive. In general, RDF(S) seems to suffer from a lack of formal semantics for its modeling primitives, making interpretation of how to use them properly an error-prone process.

A solution to this problem is provided by the third basic component of the Semantic Web, viz. ontologies. In philosophy, an ontology is a theory about the nature of existence, about what types of things exist; ontology as a discipline studies such theories. Artificial Intelligence and Web researchers have co-opted the term for their own jargon, and for them an ontology describes a formal, shared conceptualization of a particular domain of interest.

Ontologies are specifications of the conceptualization and corresponding vocabulary used to describe a domain (Gruber 1993). They are well-suited for describing heterogeneous, distributed and semistructured information sources that can be found on the Web. By defining shared and common domain theories, ontologies help both people and machines to communicate concisely, supporting the exchange of semantics and not only syntax. It is therefore important that any semantic for the Web is based on an explicitly specified ontology. By this way consumer and producer agents (which are assumed for the Semantic Web) can reach a shared understanding by exchanging ontologies that provide the vocabulary needed for discussion.

Ontologies typically consist of definitions of concepts relevant for the domain, their relations, and axioms about these concepts and relationships. Several representation languages and systems are defined. A recent proposal extending RDF and RDF Schema is OIL (Ontology Interchange Language) (Fensel et al. 2001). OIL unifies the epistemologically rich modeling primitives of frames, the formal semantics and efficient reasoning support of description logics and mapping to the standard Web metadata language proposals. The DAML+OIL language (http://www.daml.org/2001/03/reference.html) has also been developed as an extension to XML and RDF. It is a representation language for describing web resources and supporting inference over those resources. It provides a rich set of constructs for creating ontologies and to markup ontologies so it is machine readable and understandable.

Semantic Web & eLearning

Key property of the Semantic Web architecture (common-shared-meaning, machine-processable metadata), enabled by a set of suitable agents, seems to be powerful enough to satisfy the eLearning requirements: fast, just-in-time and relevant learning. Learning material is semantically annotated and for a new learning demand it may be easily combined in a new learning course. According to his/her preferences, user can find useful learning material very easily. The process is based on semantic querying and navigation through learning materials, enabled by the ontological background.

In fact, the Semantic Web could be treated as a very suitable platform for implementing an eLearning system, because it provides all means for (eLearning) ontology development, ontology-based annotation of learning materials, their composition in learning courses and (pro)active delivery of the learning materials through eLearning portals. More details about the eLearning scenario will be given in the last section. In the following (Tab. 2) a summary view of the possibility to use the Semantic Web for realizing the eLearning requirements is presented.

Requirements	eLearning	Semantic Web
Delivery	Pull – Student determines agenda	Knowledge items (learning materials) are distributed on the web, but they are linked to commonly agreed ontologie(s). This enables construction of a user-specific course, by semantic querying for topics of interest.
Responsiveness	Reactionary – Responds to problem at hand	Software agents on the Semantic Web may use commonly agreed service language, which enables co-ordination between agents and proactive delivery of learning materials in the context of actual problems. The vision is that each user has his own personalised agent that communicates with other agents.
Access	Non-linear – Allows direct access to knowledge in whatever sequence makes sense to the situation at hand	User can describe situation at hand (goal of learning, previous knowledge,...) and perform semantic querying for the suitable learning material. The user profile is also accounted for. Access to knowledge can be expanded by semantically defined navigation.
Symmetry	Symmetric – Learning occurs as an integrated activity	The Semantic Web (semantic intranet) offers the potential to become an integration platform for all business processes in an organisation, including learning activities.
Modality	Continuous – Learning runs in parallel and never stops	Active delivery of information (based on personalised agents) creates a dynamic learning environment.
Authority	Distributed – Content comes from the interaction of the participants and the educators	The Semantic Web will be as decentralised as possible. This enables an effective co-operative content management.
Personalization	Personalized – Content is determined by the individual user's needs and aims to satisfy the needs of every user	A user (using personalised agent) searches for learning material customised for her/his needs. The ontology is the link between user needs and characteristics of the learning material.
Adaptivity	Dynamic – Content changes constantly through user input, experiences, new practices, business rules and heuristics	The Semantic Web enables the use of knowledge provided in various forms, by semantical annotation of content. Distributed nature of the Semantic Web enables continuous improvement of learning materials.

Table 2 Benefits of using Semantic Web as a technology for eLearning

Metadata & eLearning

This section gives overview of the current metadata standards for eLearning, problems in shared-understanding, which arise when using these conventional metadata and the enhancement, achieved using ontology-based solution (ontology-based metadata) applied in our e-Learning scenario (following section).

Conventional Metadata for eLearning

Compared to traditional learning in which the instructor plays the intermediate role between the learner and the learning material, the learning scenario in eLearning is completely different: instructors no longer control the delivery of material and learners have a possibility to combine learning material in courses on their own. So the content of learning material must stand on its own. However, regardless of the time or expense put into creating advanced training material the content is useless unless it can be searched and indexed easily. This is especially true as the volume and types of learning content increase.

One solution lies in using metadata. Metadata is the Internet-age term for information that librarians traditionally have used to classify books and other print documents. At its most basic level, metadata provides a common set of tags that can be applied to any resource, regardless of who created it, what tools they used, or where it's stored. Tags are, in essence, data describing data. Metadata tagging enables organizations to describe, index, and search their resources and this is essential for reusing them.

In the eLearning community three metadata standards are emerging to describe eLearning resources: IEEE LOM (http://ltsc.ieee.org/doc/wg12/LOM3.6.html), ARIADNE (http://ariadne.unil.ch/Metadata/) and IMS (http://www.imsproject.org/metadata/imsmdv1p2/imsmd_infov1p2.html). Those meta-models define how learning materials can be described in an interoperable way. All the metadata elements necessary to describe a resource can be classified into several categories, each offering a distinct resource viewpoint.

For example, the LOM standard contains the following metadata levels:
- general - groups all context-independent features plus the semantic descriptors for the resource;
- lifecycle - groups the features linked to the lifecycle of the resource;
- meta-metadata - groups the data elements describing the metadata that indexes the document;
- technical - groups data elements describing the technical features of the document;
- educational - groups educational and pedagogic data elements for the resource;
- rights - groups data elements pertaining to the conditions of use for the resource;
- relation - groups data elements that describe the linkage between the subject and other resources;
- annotation - groups data elements that allow comments on the educational use of the resources;
- classification - groups data elements that describe the position of the resource in an existing classification system.

Different communities have developed their own standardized metadata vocabularies to meet their specific needs. However, most of those metadata standards lack a formal semantics. Although these standards enable interoperability within domains, they introduce the problem of incompatibility between disparate and heterogeneous metadata descriptions or schemas across domains.

This lack of a shared understanding between terms in one vocabulary as well as terms in various metadata vocabularies might be avoided by using ontologies as a conceptual backbone in an eLearning scenario.

Ontology-based metadata

The role of an ontology is to formally describe shared meaning of used vocabulary (set of symbols). In fact, an ontology constrains the set of possible mapping between symbols and their meanings. But the shared-understanding problem in eLearning occurs on several orthogonal levels, which describe several aspects of document usage, as described in Fig. 2.

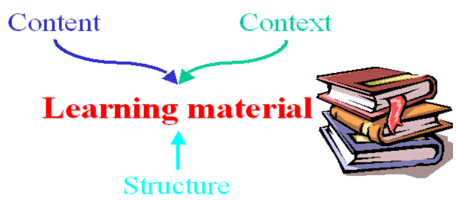

Figure 2 From the student point of view the most important things for searching learning materials are: what the learning material is about (content) and in which form this topic is presented (context). However, while learning material does not appear in isolation, another dimension (structure) is needed to encompass a set of learning materials in a learning course.

Metadata for describing content of learning materials

The shared-understanding problem in eLearning occurs when one tries to define the content of a learning document in the process of providing learning materials as well as in the process of accessing to (searching for) particular learning material.

In an eLearning environment there is a high risk that two authors express the same topic in different ways. This means semantically identical concepts (i.e. topics of eLearning-content) may be expressed by different keywords. For example, one may use the following semantically equivalent terms for "Agent": *agent, actor, contributor, creator, player, doer, worker, performer*. The problem could be solved using domain (content) ontologies in which mappings from domain vocabulary(s) in the commonly-agree terms are defined extensionally (e.g. *agent, actor, contributor, creator, player, doer, worker, performer* are symbols used in real-world and they map to the concept *Agent* in the domain ontology). Also, in the process of providing information ontological axioms play an important role. For example, an axiom that states that two relations are mutually inverse relations is used for checking consistency of provided information, as described in the next section.

From the point of view of the user there is the problem of what terms or keywords to use when searching for learning materials. Simple keyword queries are valuable in situations where users have a clear idea of what they are seeking and the information is well-defined. It doesn't hold for eLearning, where the viewpoints and the knowledge levels of author and users of learning materials may be completely different and some mechanism for establishing shared understanding is needed. Second, simple keyword searches cannot pick up synonyms ("*Agent*" and "*Actor*"), abbreviations ("*World Wide Web*" and "*WWW*"), different languages („*house*" (English) and „*Haus*" (German)) and often not even morphological variations ("*Point-to-Point Network*" and "*Point to Point Network*"), not to mention the context of the query. This problem could be resolved by defining corresponding relations (e.g., synonym, abbreviation) in the domain ontology. Ontological relations are also used in the process of navigating through learning materials (for example, from topic "*Network*" it is reasonable to "jump" to topic "*Protocol*").

Metadata for describing context of learning materials

Learning material could be presented in the various learning contexts: as introduction, as analysis, as discussion; or in the various presentation contexts: as example, as figure. The context description enables context-relevant searching for learning material according to the preferences of the user. For example, if the user needs a more detailed explanation of the topic, it is reasonable to find learning material which describes an example of the given topic. In order to achieve shared-understanding about meaning of the context vocabulary (e.g. intro or introduction) a context-ontology is used.

Metadata for describing structure of learning materials

Because eLearning is often a self-paced environment, training needs to be broken down into small bits of information ("lego" learning) that can be tailored to meet individual skills gaps and delivered as needed. These chunks of knowledge should be connected in order to create the whole course. Learning material is usually more complex in its structure than continuous prose, so it requires greater care in its design and appearance. Much of it will not be read continuously. The structure isn't a static one, because it depends on user type, users' knowledge level, users' preferences and prerequisite materials. But, again shared understanding about used terms is also needed for describing the structure of a learning course.

Several kinds of structuring relations between elementary learning materials may be identified. Some of them are: *Prev*, *Next*, *IsPartOf*, *HasPart*, *References*, *IsReferencedBy*, *IsBasedOn*, *IsBasisFor*, *Requires*, *IsRequiredBy*. There are semantic connections between some of these relations defined by axioms: for example, *IsPartOf* and *HasPart* are mutually inverse relations. This corresponding axiom may help in searching for information. Without the definition of the inverse relation searching for information would depend on the strategy of providing metadata information. If one defines that some learning material named "*X*" "*IsBasedOn*" some other learning material named "*Y*", there is no possibility (without programming or explicit specification) to find all learning materials the learning material "Y" "*IsBasisFor*".

The reader may note that these three dimensions of metadata also appear in the conventional metadata model (content = classification metadata, context = educational/pedagogical metadata, structure = relational metadata). However, our metadata are ontology-based metadata and describe the whole domain (including axioms), not only data. Consequences are, as mentioned previously, better (in the mean of the semantic) describing of learning materials and better searching for useful materials according to user preferences. Our ontology-based approach could be very easily extended to the situation that all of the conventional metadata levels (e.g., general, technical) are used (in ontology-based manner) in annotation of learning materials.

Semantic Web-based eLearning scenario and preliminary experiences

Based on the discussion in the previous section, this section presents overall architecture of our ontology-based eLearning scenario. The architecture of the system is represented in Fig. 3. The knowledge warehouse acts as a metadata repository and the Ontobroker system (Decker et al. 1999) is an principal inferencing mechanism.

The backbone of the system is the course ontology presented in the Tab 3. The ontology definition contains an is-a hierarchy of relevant domain concepts, possible relation between concepts, further properties of concepts (attributes with value ranges), and the derivation rules to infer new knowledge. The leftmost column shows the concepts of the domain organized in the is-a hierarchy. For example, "*PhDStudent*" is a subconcept of a concept "*Student*". Attributes and relations of concepts are inherited by subconcepts. Multiple inheritances are allowed as a concept may fit into different branches of the taxonomy. Attributes and relations of the concepts appear in the middle column in the Tab. 3. Relations refer to the other concepts, like "*hasAuthor*" denoting a relation between concept "*Document*" and concept "*Author*". The rightmost column shows course ontology's rules. For example, the fourth rule in Tab. 3 ensures that whenever a document is known to have a child document then a "child document" also has "parent document" that is the particular document. This kind of rules completes the knowledge and frees a knowledge provider to provide the same information at different places reducing development as well as maintenance efforts. The ontology representation language is F-Logic (Kifer et al. 1995). In the simplified interpretation, statements *ConceptX::ParentX* and *ConceptX[relationXY=>>ConceptY]* could be read as *ConceptX* is a subconcept of the concept *ParentX* and *ConceptX* is in the relation r*elationXY* with *ConceptY*, respectively.

The course ontology consists of content, context and structure ontology, mentioned in the previous section. The content ontology is visible in the description of domain terms like "*Protocol*", "*Service*", "*Topology*". The relation "*hasTopic*" and the first two rules are also a part of the content ontology. The first rule determines the transitive property of the "*hasTopic*" relation (Maedche et al. 2001). For example, based on the first rule and on the facts that "*eLearning hasTopic TeleTeaching*" and that "*TeleTeaching hasTopic WebBasedLearning*", the fact "*eLearning hasTopic WebBasedLearning*" is concluded (http://www.aifb.uni-karlsruhe.de/Personen/index.html). The second rule ensures that whenever a document with the content "*eLearning*" is searched for, then the documents about "*TeleTeaching*" and "*WebBasedLearning*" can be also found.

The context ontology is based on the pedagogical model and concepts like "*Introduction*", "*Explanation*", "*Example*" are used to describe several types or several context of the learning materials. The most important part of the structure ontology are the relations between learning materials ("*preDocument*", "*nextDocument*", "*IsBasedOn*", "*IsBasisFort*") and corresponding rules. The learning materials are organized in a tree structure. The relations "*preDocument*" and "*nextDocument*" describe a sequence of the documents at the same level in the tree of the learning materials. The relations "*parentDocument*" and "*firtsChildDocument*" correspond to the references between two successive levels. The rules in the structure ontology enable moving through the learning materials organized in a course. For example, the rule "*FORALL D1, D2 D1:Document[prevDocument->>D2] <-> D2:Document[nextDocument->>D1].*" enables to go through the learning materials in two direction (forward or backward), even though only one "path" is defined.

The concepts "*Course*", "*Module*" and "*Atom*" are also part of the structure ontology. They are used to indicate the complexity of the learning materials. The simplest type of the learning materials is an "*Atom*". It is a learning material that doesn't contain any other learning material. The "*Modul*" consists of several atoms organized in a sequence and "*Course*" is a sequence of the modules or another courses. In this way a course is a tree structure of a learning materials. Complexity of the learning materials can be defined automatically using, for example, the last rule in Tab. 3.

All others elements of the course ontology, represented in the Tab. 3, correspond to the common metadata. For example, attributes "*name*", "*title*", "*path*", which describe the "*Document*" concept, are equivalent to the general metadata level in the previously mentioned LOM standard.

Concept	**Relation**	**Rule**

Object [ ].
Document :: Object.
...
Content :: Object.
Protocol :: Content.
Service :: Content.
Topology :: Content.
Bustopology :: Topology.
Circletopology ::Topology.
...
Context::Object.
Introduction:: Context.
Explanation:: Context.
Example:: Context.
Figure::Example.
...
Structure::Object.
Course:: Structure.
Module:: Structure.
Atom:: Structure.
...
Person::Object.
Author :: Person.
Student :: Person.
PhDStudent :: Student.
...
	Document [
	name=>>String;
	title=>>String;
	path=>>String;
	hasAuthor=>>Author;
	
	content=>>Content;
	context=>> Context;
	structure=>> Structure;
	...
	prevDocument =>> Document;
	nextDocument =>> Document;
	firstchildDocument =>> Document;
	parentDocument =>> Document;
	relatedDocuments =>> Document;
	...
	IsBasedOn=>>Document;
	IsBasisFor=>>Document;
	...].
	
	Content[
	hasTopic=>>Content].
		FORALL A, B, C
		A[hasTopic->>C] <- A:Content and A[hasTopic ->>B] and B:Content and B[hasTopic ->> C] and C: Content.
		
		FORALL D, C1, C2
		D:Document[content->>C1] <- C1:Content and C2 :Content and D:Document[content->>C2] and C1[hasTopic->>C2].
		
		FORALL D1, D2
		D1:Document[prevDocument->>D2] <-
		EXISTS E1, E2, C C:Content andD2:Document[context->>E2] and E2:Example and D1[context->>E1] and E1:Explanation and D1[content->>C] and D2[content->>C].
		
		FORALL D1, D2
		D1:Document[parentDocument->>D2] <-
		D2:Document[firstchildDocument->>D1].
		
		FORALL D1, D2
		D1:Document[prevDocument->>D2] <->
		D2:Document[nextDocument->>D1].
		
		FORALL D, S
		D:Document[structure->>S:Course] <-
		Exists D1, S1 D1:Document and (S1:Course or S1:Module) and D1[structure->>S1] and D1[parent->>D].
		...

Table 3 Ontology in the eLearning scenario

Core modules, as depicted in Fig. 3, correspond to primary activities in an eLearning environment:

- providing information from authors
- accessing the learning materials by readers and authors by querying and by browsing.

Providing

The first phase is the production of learning materials that may be used or reused in the construction of training courses. In order to provide learning material, which could be suitable for metadata-searching, each learning material has to be described or "enriched" with the following metadata information:

- what is the learning material about (content annotation),
- which context has the learning material (context annotation) and
- how is it connected to other learning materials (structure annotation).

This "enriching" consists of explicitly adding to each learning material a set of metadata information referring to course ontology. Providing information is for now constrained on manually entering metadata information (facts) through automatically generated templates, based on the definition of concepts in the course ontology. The metadata may be placed within the document itself (e.g. HTML <META> tags or an extended <A> tag) or in some external metadata repository (e.g. an RDF repository) (Handschuh et al. 2001). In our approach these information are stored externally in the knowledge warehouse. First, it is easier to scan a separate meta-description stored in a database and it takes less space to store it. Second, the point of view may vary according to different authors who reuse the same learning material. It means that it is possible to have different descriptions of the learning material according to the different contexts.

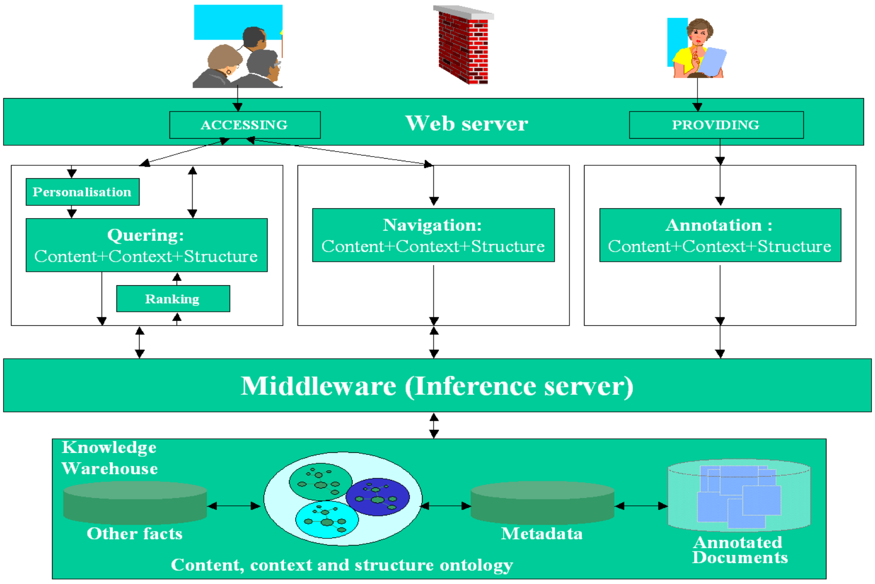

Figure 3 Architecture of an eLearning Portal

Accessing the learning materials

In the process of information accessing, the ontology is used for:

1) Semantic search for learning materials
- based on three dimensional searching space (content, context, structure);
- implemented as an easy-to-use interface on the query capabilities of the F-Logic query interface of Ontobroker

2) Conceptual navigation through the collection of learning materials
- based on ontological relations between concepts in the (a) content and (b) context ontologies

(a) postulates an assumption that the semantic relevant links for the learning material correspond to the ontological relations. From the interface level, each learning material, which is indexed using an concept (or concept instance) from ontology, has (hyper)links to the learning materials indexed by the concepts (or instances of a concept) this concept is related with. For example, in telematic domain for the learning material which describes OSI layers there are hyperlinks to corresponding protocol-, service- and interface- learning materials, while, according to telematic ontology, concept "OSI layer" has relations with the concepts "protocol",

(b) This type of navigation is based on the rules in the context ontology. The rules describe how to organize the learning materials about the same content in a proper structure. For example, from the pedagogical point of view the learning material that explains some content must precede the learning material that is an example of the same content. It means that each learning material with a context "*Explanation*" has (hyper)links to the learning materials about the same content that have the context "*Example*".

- based on explicit (navigational) structure defined from the author in the structure ontology

The navigational structure consists of the ordering of learning material in the learning course (first, next, parent), but it is also created by authors who define related learning materials that do not obligatory correspond to content ontology.

Other components

The knowledge warehouse serves as repository for data represented in the form of RDF statements. The knowledge warehouse itself hosts the ontology, the metadata, as well as the data proper.

The system uses the inference engine of Ontobroker system (Decker et al. 1999). Particularly, the inference engine answers queries and it performs derivations of new knowledge by an intelligent combination of facts in knowledge warehouse with an ontology definition. The possibility to derive additional factual knowledge that is only provided implicitly frees knowledge providers from the burden of specifying each fact explicitly. Semantic ranking and personalization are methods for refining results of querying the knowledge warehouse are described in (Maedche et al. 2001).

Related work

There are only a few approaches that could be compared to our eLearning scenario. The most similar approach is the system Karina (Crampes et al. 2000), which enables dynamical building of the learning courses according to user preferences. It is based on the conceptual description of learning material using conceptual graphs and uses some (prerequisite) strategies to fulfil the users' objectives in the search/navigation process. A sibling of Karina, the Sybil system (Crampes et al. 2000) uses an ontology of pedagogy for defining the context of the learning course. However, both approaches do not describe explicit structure of the course (structure ontology in our case).

The Collaborative Courseware Generating System (Qu et al. 2001) uses modern web technologies (XML, XSLT, WebDAV) for describing course structure, but without explicit ontology support. It also does not define context and structure of the learning materials explicitly.

Ontology-based Intelligent Authoring Tool (Chen et al. 1998) is the represent of the using an intelligent training system in the eLearning scenario. Good characteristic is using four ontologies (domain, teaching strategies, learner model and interfaces ontology) for construction of learning model and teaching strategy model, but it failed in using modern Web technologies.

As a summary remark, to note that no one of the mentioned systems use the advantages of the Semantic Web, which is the main point in our approach.

Conclusion

"Making content machine-understandable" is a popular paraphrase of the fundamental prerequisite for the Semantic Web. In spite of its potential philosophical ramifications this phrase must be taken very pragmatically: content (of whatever type of media) is 'machine-understandable' if it is bound (attached, pointing, etc.) to some formal description of itself.

This vision requires development of new technologies for web-friendly data description. The Resource Description Framework (RDF) metadata standard is a core technology used along with other web technologies like XML. Ontologies are (meta)data schemas, providing a controlled vocabulary of concepts, each with an explicitly defined and machine processable semantics. By defining shared and common domain theories, ontologies help both people and machines to communicate concisely, supporting the exchange of semantics and not only syntax.

In the same time, promising areas for applying the Semantic Web are unlimited. In fact, each area, in which a lot of information should be provided and accessed in a distributed manner, searches for some semantic-based solution.

In this paper we presented an eLearning scenario that exploits ontologies in three ways:

- for describing the semantics (content) of the learning materials. This is the domain dependent ontology,
- for defining learning context of the learning material and
- for structuring learning materials in the learning courses.

This three-dimensional space enables easier and more comfortable search and navigation through learning material.

The purpose of this paper was to clarify possibilities of using ontologies as a semantic backbone for eLearning. Primarily, the objectives are to facilitate the contribution of and efficient access to information. But, in a broader or in Semantic Web's view, an ontology-based learning process could be a relevant (problem-dependent), a personalised (user-customised) and an active (context-sensitive) process. These are prerequisites for efficient learning in the dynamically changed business. This new view enables us to go a step further and consider or interpret the learning process as a process of managing knowledge in the right place, at the right time, in the right manner in order to satisfy business objectives - knowledge management. It means the merging of eLearning and knowedge management using the Semantic Web should be the promising integration.

Acknowledgements

The research presented in this paper would not have been possible without our colleagues and students at the Institute AIFB, University of Karlsruhe, and Ontoprise GmbH. We thank all of them. Research for this paper was partially financed by EU in the IST-2000-28293 project "OntoLogging" and by US Air Force in the DARPA-DAML project "OntoAgent".

References

Adelsberger H., Bick M., Körner F., Pawlowski J.M. (2001). Virtual Education in Business Information Systems (VAWI) - Facilitating collaborative development processes using the Essen Learning Model, *In Proceedings of the 20th ICDE World Conference on Open Learning and Distance Education*, Düsseldorf, Germany, April 2001.

Berners-Lee T. (2000). What the semantic web can represent, *http://www.w3.org/DesignIssues/RDFnot.html.*

Barker Ph. (2000). Designing Teaching Webs: Advantages, Problems and Pitfalls; *Educational Multimedia, Hypermedia & Telecommunication*, Association for the Advancement of Computing in Education, Charlottesville, VA, pp. 54-59.

Chen W., Hayashi Y., Jin L., Mitsuru I., Mizoguchi R., (1998), An Ontology-based Intelligent Authoring Tool, *Proceedings of the Sixth International Conference on Computers in Education*, Beijing, pp. 41-49.

Crampes, M. and Ranwez, S., Ontology-Supported and Ontology-Driven Conceptual Navigation on the World Wide Web, *11th ACM Hypertext Conference*, May 30 -- June 4. 2000, ACM Press, San Antonio, Texas. pp. 191-199.

Decker S., Fensel D., van Harmelen F., Horrocks I., Melnik S., Klein M., and Broekstra J. (2000). Knowledge representation on the web, *IEEE Internet Computing,* September/October 2000.

Decker, S., Erdmann, M., Fensel, D., & Studer, R. (1999). Ontobroker: Ontology based access to distributed and semi-structured information, in Meersman, R. et al. (Eds.), *Database Semantics: Semantic Issues in Multimedia Systems*, Boston, USA, pp. 351–369. Kluwer Academic Publisher.

Drucker P. (2000). Need to Know: Integrating e-Learning with High Velocity Value Chains, A Delphi Group White Paper, *http://www.delphigroup.com/pubs/whitepapers/20001213-e-learning-wp.pdf* .

Erdmann M. and Studer R. (2000). How to structure and access XML documents with ontologies. *Data and Knowledge Engineering - Special Issue on Intelligent Information Integration DKE(36)* (3): 317-335

Fensel, D., van Harmelen, F., Horrocks, I., McGuinness, D. L., & Patel-Schneider, P. F. (2001). OIL: An ontology infrastructure for the semantic web. *IEEE Intelligent Systems*, 16(2):38–44.

Gruber, T. (1993). A translation approach to portable ontology specifications. *Knowledge Acquisition*, 5:199–220.

Handschuh S., Staab S., Mädche A. (2001). CREAM — Creating relational metadata with a component-based, ontology-driven annotation framework. To appear in *ACM K-CAP 2001*. October, Vancouver.

Heflin, J., and Hendler, J. (2000), Semantic Interoperability on the Web. *In Proceedings of Extreme Markup Languages*. Graphic Communications Association, 2000. pp. 111-120.

Nejdl W. (2001). Learning Repositories – technologies and Context, To appear *in Proceedings of ED-MEDIA 2001 World Conference on Educational Multimedia, Hypermedia & Telecomuunications*, June 25-30, Tampere, Finland

Kifer, M., Lausen, G., & Wu, J. (1995). Logical foundations of object-oriented and frame-based languages. *Journal of the ACM*, 42:741–843.

Maedche, A., Staab, S., Stojanovic, N., & Studer, R. (2001). SEAL - A Framework for Developing SEmantic portALs. In *Proceedings of the 18th British National Conference on Databases*, July, Oxford, UK, LNCS 2097, Springer, pp. 1-22.

Maurer H. and Sapper M (2001). E-Learning Has to be Seen as Part of General Knowledge Management, *In Proceedings of ED-MEDIA 2001* World Conference on Educational Multimedia, Hypermedia & Telecomuunications, Tampere, AACE, Charlottesville, VA (2001), pp. 1249-1253.

Qu C., Gamper J. and Neidl W. (2001), A Collaborative Courseware Generating System based on WebDAV, XML, and JSP, to appear in *Proceedings of IEEE International Conference on Advanced Learning Technologies - ICALT 2001*, Madison, USA

Staab S., Schnurr H.-P., Studer R., and Sure. Y. (2001). Knowledge processes and ontologies. *IEEE Intelligent* Systems, Vol. 16, No. 1, January/February 2001