Research Journal of Applied Sciences, Engineering and Technology 6(22): 4305-4317, 2013
DOI:10.19026/rjaset.6.3548
ISSN: 2040-7459; e-ISSN: 2040-7467

Submitted: April 06, 2013 Accepted: April 29, 2013 Published: December 05, 2013

Research Article

An Upper Ontology for E-Learning Material Semantic Annotations

T. Khdour and Ibrahim Tadros
Department of Information Technology, Prince Abdulah Bin Gazi Faculty of Information Technology,
Al-Balqa Applied University, Al-Salt 19117, Jordan

Abstract: Recent research reveals a great interest to introduce the Semantic Web as a promising technology for realizing eLearning requirements. The new, dynamic and distributed business world has motivated the research on developing eLearning. ELearning is efficient, task relevant and just-in-time learning. It gives the learner the ability to efficiently access the related educational resources just-in-time from any place. The vision of the Semantic Web is to make the Web data not only processable but also understandable so it can be used by machines not just for display purposes but for automation, integration and reuse of data across various applications. This study investigates the role of Semantic Web in realizing the e-learning requirements. It proposes an ontology-based e-learning framework that considers the main three component roles of the e-learning architecture: an author, a learner and a repository. The study also shed the light on improving the conventional metadata standards that are used to describe learning materials by proposing a semantic-based ontology to describe three different dimensions of the learning material: content, context and structure. Adopting the proposed ontology would result in facilitating both the process of finding suitable learning materials to build up a certain course and the process of navigating through the learning course.

Keywords: E-learning framework, learning object, metadata, semantic web, upper ontology

INTRODUCTION

e-Learning is responsible of synchronizing the workers' knowledge with their work environment. Hence, it is responsible of providing the right knowledge to the right people at the right time by dedicating the use of technology to enable people to learn anytime and anywhere. Moreover, eLearning, nowadays, becomes a fundamental part of the corporate strategy. However, adopting a new style of learning constitutes a major challenge to every industry. The new style of learning will be driven by the requirements of the new economy: efficiency, just-in-time delivery and task relevance (Stojanovic *et al*., 2001).

The organizational learning is the intentional use of learning processes at the individual, group and system level to continuously transform the organization in a direction that is increasingly satisfying to its stakeholders (Dixon, 1999). To encourage businesses invest into it, learning has to satisfy a set of requirements; the learning processes need to be efficient and just-in-time, the learning material must be organized and can be retrieved efficiently, it also must be customized i.e., it is initiated according to user profiles and finally, learning needs to be relevant to the semantic context of the business (Adelsberger *et al*., 2002).

E-Learning aims at replacing the time, place, static content, predetermined learning into just-in-time, at work place, customized, on demand learning (Stojanovic *et al*., 2001). Information Technology (besides management and culture) is one of the main pillars that e-Learning is built on. The IT is needed to implement the infrastructure that meets the e-Learning requirements: efficient, just-in-time and relevant. The existence of machine processable but not understandable information is one witness that current web based solutions do not satisfy the requirements of e-Learning. The Semantic Web comes to play its role in developing languages and approaches for expressing information in machine understandable forms.

The Semantic Web is a promising technology for realizing e-Learning requirements. According to its vision, the human and machine agents will communication on a semantic basis (Berners-Lee, 2000). "The Semantic Web is a vision: the idea of having data on the Web defined and linked in such a way that it can be used by machines not just for display purposes, but for automation, integration and reuse of data across various applications" (Kashyap *et al*., 2008). Ontologies play a main role in Semantic Web by enabling shared understanding of the domain problem. Ontology includes a description of a set of concepts (classes) and roles (relationships) in a specific problem

Corresponding Author: T. Khdour, Department of Information Technology, Prince Abdulah Bin Gazi Faculty of Information Technology, Al-Balqa Applied University, Al-Salt 19117, Jordan

domain. In eLearning scenarios, ontologies allow computer agents and programs to interpret the meaning of the eLearning materials unambiguously.

We believe that research efforts should focus on the problem of standardizing the way of semantically describing eLearning materials. Recently, we have seen many proposals towards standardizing the semantic descriptions of different Web resources. For example, OWL-S (Martin *et al*., 2004) is an OWL based ontology that aims at utilizing the Semantic Web to describe Web services with the objective of supporting more effective discovery, composition and interoperation. However, the main objective of this study is to propose a semantic-based approach that standardizes the way of semantically describing e-Learning materials. This approach takes the responsibility of facilitating both the process of locating suitable learning materials to build up a certain course and the process of navigating through the learning course.

LITRATURE REVIEW

There are some approaches in the literature that could be compared to our proposal. However, in this section, we try to highlight the most related ones. An ontology-based intelligent authoring tool is proposed in (Chen *et al*., 1998). The tool uses four main Ontologies (domain, teaching strategies, learner model and interfaces ontology) to construct the learning and the teaching strategy models. Still, the proposed tool fails in exploiting modern Web technology especially the Semantic Web technology to present the authoring tool.

The Karina project proposed in Crampes and Ranwez (2000) enables dynamically building courses that are adapted to the needs of users. The long range objective of the project was to propose several conceptual navigation strategies among which the system will choose the best adapted to the learner's need. Karina is based on the conceptual description of learning materials to fulfill the users' objectives in the navigation and search process through utilizing some prerequisite strategies. Sybil, the application example presented also in Crampes and Ranwez (2000), uses conceptual graphs to formalize the domain and the pedagogy Ontologies. Although the pedagogic ontology contains a hierarchy of pedagogic concepts, pedagogic rules and pedagogic strategies, it fails to explicitly describe the context of the learning course. Moreover, neither Karina nor Sybil systems have utilized the semantic web technologies in their implementations.

The Collaborative Courseware Generating System proposed in Qu *et al*. (2001) focuses on taking advantages of recent Internet protocols and industry standards to facilitate the courseware generating process. Although the proposed system used various Web technologies (XML, XSLT, WebDAV) for describing course structures, it lacked explicit ontology definitions that describe the context and the structure of the learning materials.

An approach for implementing the eLearning scenario using Semantic Web technologies is presented in Stojanovic *et al*. (2001). The backbone of the proposed eLearning portal is the eLearning ontology. To achieve the goal of enabling easier and more comfortable search and navigation through the learning material, the authors presented an eLearning scenario that exploits Ontologies that describes three dimensions of the learning object: the content, the context and structure. Yet, the proposed e-Learning ontology is not considered to be comprehensive and does not exploit modern Semantic Web technologies.

An e-Learning framework that is based on the Semantic Web is presented in Naeve *et al*. (2001). The semantic web is utilized in order to develop tools, standards and environments that support four main areas: content management, knowledge navigation and experience-orientated environment. The presented project is part of a consortium comprising Swedish and German universities developing a P2P network for the exchange of educational resources. Although the framework sheds the light on the importance of the semantic web, we believe that the framework lacks concrete Ontologies that can be utilized to enable the desired tasks: search, retrieval, publication, replication and mapping of metadata.

The study presented in Henze *et al*. (2004) proposed a framework for building adaptive and personalized educational hypermedia systems. The authors utilized the semantic Web technologies to automatically generating hypertext structures from distributed metadata. Four main Ontologies were developed: the domain ontology, user ontology, observation ontology and the presentation ontology. We believe that the set of rules employed to reason over distributed information resources is not comprehensive and needs to be expanded to cover more scenarios. Moreover, utilizing OWL instead of RDF would add a great value to the proposed approach.

It is argued that current e-learning resources description standards fail to address the instructional purpose of a resource, for instance, whether a web page provides a definition or a counter-example of a concept. Hence, anontology of instructional objects (OIO) that captures the function of a learning resource is proposed in Ullrich (2004). However, this proposal is not specifically oriented to design teaching methods adapted to different learning style models.

The <e-aula> approach presented in Sancho *et al*. (2005) combines context ontology and a pedagogical ontology to create dynamic personalized courses using IMS LD specification. The approach mainly proposes the use of ontologies as the knowledge representation mechanism to allow the delivery of learning material that is relevant to the current situation of the learner. However, the project focuses only on the context dimension of the eLearning resource. Moreover, the proposed pedagogical ontology needs to be extended to better fit different learning styles.

Table 1: Characteristics and differences between traditional learning and eLearning (Druker, 2000)

Aspect	Traditional learing	e-Learning
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 to business tasks 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 originally authored form without regard to environmental changes	Dynamic: Content changes constantly through user input, experiences, new practices, business rules and heuristics

E-Learning: *"E-Learning is a broad combination of processes, content and infrastructure to use computers and networks to scale and/or improve one or more significant parts of a learning value chain, including management and delivery. Originally it aimed at lowering management cost while increasing accessibility and for measurability of employees"* (Adrich, 2004). Semantic is not the only difference between traditional learning and e-Learning, Table 1 depicts that they differ in eight different ways. The table also presents the characteristics and pitfalls of both approaches.

E-Learning concerns moving the tools and knowledge needed to perform work to workers whoever and wherever they are. It had its origins in Computer-Based Training (CBT), an attempt to automate education, replace a paid instructor and develop self-paced learning. However, e-Learning should not be confused with traditional forms of CBT, which is nothing more than recorded education (Druker, 2000). Two main benefits of eLearning that make it different from other forms of educations (such as academic education, CBT and distance learning): the elimination of the barriers of time and distance and personalization of the user's experience.

Traditional learning processes can be characterized by central authority where learning material is selected by the educator only. On the other hand, e-Learning processes are seen as a distributed student oriented processes where learning material is selected by the learner and according to his needs to fill his/her skills gap. As can be seen in Table 1, traditional learning processes adopt the push delivery where the instructor determines the learning material and pushes the knowledge to learners, while eLearning processes adopt the pull delivery where students determine the agenda and seek for the knowledge that suites their needs. Moreover, Traditional learning processes is criticized for the lack of personalization because the learning material must satisfy the needs of many, where the learning material in eLearning is personalized and selected to satisfy the needs of every user. Yet, another pitfall of traditional learning is that its content remains static and isolated from the environmental changes, while in e-Learning, the content changes constantly according to business rules and needs. Finally, the traditional learning processes have a defined progression of knowledge, while in eLearning the processes allow direct access to knowledge in whatever sequence. As we can see from this comparison between traditional learning and eLearning processes, eLearning satisfies the requirements of e-Learning mentioned in the previous section: efficient, just-in-time and task relevant.

Semantic web: The Semantic Web is an extension of the current Web, in which information is given well-defined meanings. The goal of the Semantic Web is to express information in machine-understandable forms. Moreover, the availability of machine-understandable descriptions (Semantic Web Activity Statement, 2013) is a must for discovering a learning material that is related to the learner's need. The W3C consortium presents a practical definition of the Semantic Web:

> "*The Semantic Web provides a common framework that allows data to be shared and reused across application, enterprise and community boundaries. It is a collaborative effort led by W3C with participation from a large number of researchers and industrial partners. It is based on the Resource Description Framework (RDF), which integrates a variety of applications using XML for syntax and URIs for naming*" (World Wide Web Consortium, 2001).

The Semantic Web should enable greater access to the Web contents including the learning materials. In the context of eLearning, the Semantic Web can be utilized to discover learning materials based on their contents rather than their plain text keywords. Learning materials that are annotated with semantic information become meaningful to computer programs. Indeed, augmenting learning materials with semantic information enables computer programs to autonomously decide whether or not a particular learning material satisfies certain requirements. This

can be achieved by annotating learning materials with semantic information. Using shared Ontologies, these learning materials are given a predefined meaning and thus, they become machine-understandable.
An ontology is defined as:

"*a formal explicit description of concepts in a domain of discourse (classes (sometimes called concepts)), properties of each concept describing various features and attributes of the concept (slots (sometimes called roles or properties)) and restrictions on slots (facets (sometimes called role restrictions))*" (Noy and McGuinness, 2001).

Ontologies are the basis of the Semantic Web and they allow computer agents and programs to interpret the meanings of different Web resources unambiguously. One main goal of developing Ontologies is to share common understanding of the structure of information among people or software agents.

One interesting feature of Ontologies is that they can import other Ontologies transitively and they can use the concepts and relationships defined in those imported Ontologies. This allows for the creation of a few domain-independent Ontologies, where very general concepts and relationships are defined and many domain-dependent Ontologies, where more specific concepts and relationships are defined.

Semantic web architecture: The underlying framework of the Semantic Web is composed of the eXtensible Markup Language (XML) (Bray *et al*., 2008) and Resource Description Framework (RDF) (Manola and Miller, 2004). XML allows users to add arbitrary structure to their documents but says nothing about what the structures mean (Erdmann and Studer, 2001). XML Schema (XML-S) is an extension to XML that defines structure with a richer language. Both XML and XML-S lack a semantic model, they help in presenting data in a machine processable but not understandable format. Hence they are not the solution to augment learning materials (and Web resources in general) with semantics.

RDF is an XML-based language for describing resources on the Web. It is considered as the first step towards realizing the semantic Web. RDF identifies objects using Uniform Resource Identifiers (URIs) and describes resources using simple properties and property values. An RDF document represents its knowledge in forms of Subject-Verb-Object (SVO) triples. Triples are used to define the relationships between described concepts. RDF is very similar to a basic directed graph; nodes in the graph represent both subjects and objects and arcs represent predicates. Basic class hierarchies and relations between classes and objects can be expressed in RDFS (Brickley *et al*., 2004). Still, RDFS suffers from the lack of formal semantics for its modeling primitives.

The lack of more sophisticated constructs such as data types, enumerations and restrictions, which are needed to describe more complex Web resources, has limited the use of RDF and motivated the development of more expressive description frameworks. Many ontology languages have been proposed in the last few years. Examples include Ontology Exchange Language (XOL), Simple HTML Ontology Extension (SHOE), Ontology Markup Language (OML), Resource Description Framework (RDF) and RDF Schema (RDFS), OIL, DAML+OIL and OWL. However, DAML-S and DAML+OIL are built on top of RDF.

Web Ontology Language (OWL): The DARPA Agent Markup Language (DAML) was developed by the Defense Advanced Research Project Agency (Defense Advanced Research Projects Agency, 2013) based on RDF and RDF Schema. The Ontology Inference Layer (OIL) was developed to cover the shortcomings of RDF and RDF-S. DAML+OIL (Horrocks *et al*., 2001) is a description logics markup language that combines DAML and OIL to define Web ontologies. The goal of DAML+OIL is to semantically markup Web resources. To realize the vision of the Semantic Web, the Web Ontology Language (OWL) (McGuinness and Harmelen, 2004) was developed on top of DAML+OIL by the US/UK ad hoc Joint Working Group on Agent Markup Languages. The main goal of developing OWL is to give explicit meanings to the information in the Web, making it possible for machines to automate processes and integrate the information available on the Web more easily.

The OWL ontology generally consists of three main components: Individuals; Properties and Classes. Individuals represent objects in the problem domain that we are interested in. Classes contain a set of individuals and properties are binary relations that link individuals to individuals or to data types. Moreover, OWL distinguishes between three main types of properties: Object Property that links an individual to an individual. (e.g., John hasCoAuthor Edward), Datatype Property that links an individual to an XML Schema Datatype value (e.g., “Introduction to eLearning” hasISBN 987321) and Annotation Property that can be used to add meta data (data about data) about an ontology. (e.g., Book dc:title "Introduction to e-Learning").

Another feature of OWL is that it distinguishes between inverse, functional, transitive and symmetric property. If a property can be an inverse of another then the property is inverse. For example, if Inheritance has Next Polymorphism then we can infer that Polymorphism has Previous Inheritance because the inverse of the has Next property is the has Previous property. A functional property means that an individual can have only one value through this property. For example, if “Introduction to eLearning” has Publisher “John Wily” and “Introduction to e-Learning” has Publisher JW then we can infer that John Wily and JW are the same individual (i.e., the same publisher) if we define the has Publisher property as functional property. If a property P is transitive and it links individual x to individual y and it also links individual y to individual z, then we can infer that x is linked to the property z by the property P. For example, if Interface has Previous Polymorphism and Polymorphism has Previous Inheritance, we can infer that Interface has Previous Inheritance if the property has Previous is a transitive property. Finally, if a property P is a symmetric property and it links an individual x with an individual y, then we can infer that an individual y is linked to the individual x through the property P. For example, if eLearning has Related Topic Distance Learning, we can infer that Distance Learning has Related Topic eLearning because the property has Related Topic is symmetric.

One more interesting feature of OWL is that its description can include restrictions. Restrictions can be thought of as anonymous super classes of the class that is being described. The restriction consists of the restriction type, the property and the filler. For example, the restriction *∃* has Content *Network* contains the Existential Quantifier, the property hasContent and the filler Network. OWL supports three main types of restrictions: Quantifier (Existential and Universal), cardinality and hasValue. The Existential Quantifer (∃) is read as “at least one", “there exists" or “some”. For example, the restriction *∃* has Content *Network* describes the class of individuals that have at least one content that is an individual from the Network class. The Universal Quantifier (∀) is read as “only". For example, the restriction *∀ has Content Network* describes the class of individuals all of whose has Content relationships are to members of the Network class. The cardinality restriction is used to describe the number of relationships an individual must have for a given property. Three cardinality restrictions can be used: Exact Cardinality; Minimum Cardinality and Maximum Cardinality. For example, if we describe a class called Book, then the descriptions*: = hasChapters 3*, *> = hasChapters 4*, *< = hasChapters 4* state that the book has to have exactly two chapters, minimum 4 chapters and maximum 4 chapters, respectively. The hasValue Restriction is used to link a set of individuals that have at least one relationship through a given property to specific individual. The hasValue restriction is denoted by the symbol ∈. For example, *∈ hasAttached CD* describes a set of individuals that have at least one relationship along the hasAttached property to the specific individual CD.

Descriptive metadata standards: *“Metadata is structured information that describes, explains, locates, or otherwise makes it easier to retrieve, use, or manage an information resource. Metadata is often called data about data or information about information”* (The National Information Standards Organization, 2004).

Metadata is important to ensure that the described resource will be discovered and continued to be accessible. For the content of a learning material to be reached, it has to be indexed and searched easily. This is especially true as the volume of learning materials increases. Traditionally, Metadata can be used to describe any types of textual and non-textual objects including published books, electronic documents, educational and training materials.

Standard metadata used in eLearning suffers from the problem of shared understanding. Indeed, shared understanding of concepts used in describing eLearning materials is essential for computer programs and agents because it enables them to create the learning course according to the learner needs and preferences. However, annotating metadata with semantics is an ideal solution for such a problem. The use of ontology based metadata is an essential part of our approach as it enables learners to search and combine learning materials based on the meaning rather than on the content. Moreover, this increases the chance of finding related learning materials that would not be found when dealing with learning materials that are annotated with only standard metadata. The reasons for creating descriptive metadata as stated by the National Information Standards Organization (NISO) (National Information Standards Organization, 2004) are to facilitate discovery of relevant information, organize electronic resources, facilitate interoperability and legacy resource integration, provide digital identification and finally to support archiving and preservation. Many metadata schemas (standards) are being developed to describe information resources. Examples include the Dublin Core, the Text Encoding Initiative (TEI), Metadata Encoding and Transmission Standard (METS), Metadata Object Description Schema (MODS) and the Encoded Archival Description (EAD). However, three metadata schemas are developed for eLearning: IEEE Learning Object Metadata (LOM) (Institute of Electrical and Electronics Engineers, 2002), ARIADNE and IMS (IMS Global Learning Consortium, 2006). ARIADNE Foundation (2013) submitted an early version of its specification to the IEEE LTSC Learning Object Metadata (LOM). Together with a similar specification contributed by the IMS Project, that early ARIADNE version was the basis of the LOM standard. The metadata specification of the IMS project is based on the IEEE LOM scheme with only minor modifications. The goal of the LOM standard is to enable the use and reuse of technology-supported learning resources such as computer-based training and distance learning. The LOM includes a set attributes grouped into eight categories:

- **General:** Containing information about the object as a whole.
- **Lifecycle:** Containing metadata about the objects evolution.

- **Technical:** With descriptions of the technical characteristics and requirements.
- **Educational:** Containing the educational/ pedagogical attributes.
- **Rights:** Describing the intellectual property rights and use conditions.
- **Relation:** Identifying related objects.
- **Annotation:** Containing comments and the date and author of the comments.
- **Classification:** Which identifies other classification system identifiers for the object.

Conventional metadata standards have contributed to achieve the goal of facilitating the systems interoperability. This means that eLearning systems with different hardware, software and interfaces are now able to exchange data without losing the content and the functionality. However, these standards enable interoperability within specific domains but they failed to address the issue of compatibility between heterogeneous domains. One reason for this failure is that they lack a formal semantic model in their descriptions. This absence of semantic descriptions yield to the problem of shared understanding between the terms within a certain metadata ontology (vocabulary) as well as between terms in different metadata vocabularies. To solve the problems of both compatibility and shared understanding, annotating metadata vocabularies with semantics using Ontologies is a must.

Ontology-based metadata: Three dimensions of metadata should be taken to consideration when annotating learning materials with semantics: content, context and structure (Fig. 1). Learners may search for eLearning materials based on one or more of these three dimensions. The content dimension describes the concrete components of the described learning material. In other words it describes what the learning material is all about. The context dimension describes the presentation form of the described learning material. Learning materials can exist in different presentation forms, for example an introduction, an explanation, discussion, details, figure, example, discussion, survey and flow chart. In eLearning, the learning material is composed of chunks of knowledge that need to be linked together someway and according to the learner's needs and preferences to build up the training course. The structure dimension is used to describe the location and order of a particular learning material in relation to other learning materials.

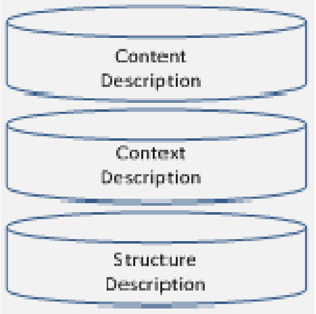

Fig. 1: Dimensions of e-learning metadata

The content metadata: The description of the content of the learning material is essential to solve the problem of shared understanding mentioned previously. As depicted in Figure 2, the author describes the content of the learning material semantically. Learners (or software agents who are responsible of building up the course) search this description and decide whether it satisfies their query. Hence, content metadata description is not only used when providing learning materials but also when searching for them. When adopting our upper ontology, proposed in the next section, to annotate metadata with semantics, the search for learning materials will not be based on simple keyword (syntax) search, instead, it will be based on the semantic description of their metadata.

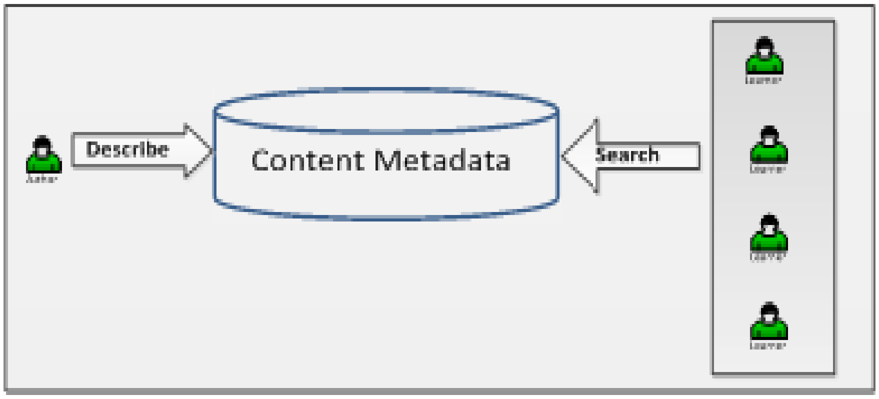

Fig. 2: The use of content description in providing and searching e-learning materials

The keyword based search fails to retrieve materials described using synonyms (“old” and “ancient”), abbreviations (“DL” and “Description Logic”), different languages (“vehicle” in English and “vehikel” in Dutch), morphological variations (“semantic-based metadata” and “semantic based metadata”), singular and plurals (“mouse” and “mice”) of the search string. Moreover, two metadata descriptions can have the same keywords to describe different topics. For example, two descriptions for two different learning materials can use the keyword “soap”; while the first description tries to use this keyword to denote the cleaning product soap, the other one tries to use the same keyword to denote the SOAP protocol (Simple Object Access Protocol). Keyword-based search suffers from two well known problems in information retrieval systems: low precision and reduced recall. The former means that many irrelevant metadata may include the query keywords in its description. The latter means that the query keywords are semantically equivalent but syntactically different from the words in the described metadata. To overcome such problems, we annotate metadata descriptions with semantics to avoid the problems of the keyword-based search. The domain ontology used to describe the learning materials metadata should encapsulate all the above mentioned relations: *synonyms*, *abbreviations*, *language*, *morphological variations*, *singular* and *plurals*

The presentation context metadata: Learning materials can be presented in different presentation contexts; examples include *an introduction, an explanation, discussion*, *details*, *figure*, *example*, *discussion*, *survey* and *flow chart*. Learners may specify their preferred context when searching for eLearning materials. Sometimes, the type of the users and their level of knowledge determine the type of presentation context they search for. For example, experienced learners may search for a detailed explanation of a certain topic while beginner learners may search for an introduction to the topic. Context metadata ontology is needed to describe the presentation context of the described eLearning material. This ontology plays a main role in achieving a shared understanding between vocabularies used in describing the metadata of different eLearning materials. For example, the domain ontology used to describe the context metadata of a particular eLearning material should define “Fig” and “Figure” as synonyms.

The structure metadata: In contrast to the static traditional learning material, eLearning material is dynamic (Table 1). It is built up based on the learner’s level of knowledge and his/her preferences. We should not expect that learners will go through the eLearning material sequentially. They may move from one topic to another according to their interest and understanding of the topic. Hence, the relationships between the chunks of knowledge that constitutes the e-Learning materials should be defined. The LOM defined a set of structure relations that are Table 2. More structure relations can be defined such as: *is Narrower Than*, *is Broader Than*, *is Alternative To*, *illustrates*, *isIllustrated By*, *isles Specific Than and isMore SpecificThan* (Engelhardt *et al*., 2006). The properties and characteristics of these relations are distributed across metadata repositories.

Table 2: The LOM relations derived from the original dublin core relations

The Relation	The Inverse
Has part	Is part of
Has version	IS version of
Has format	IS format of
References	Is referenced by
Is basis for	Is based on
Requires	Is required by

A description logic reasoner such as RACER (Haarslev and Moller, 2001) and Pellet (Sirin *et al*., 2007) and JENA framework (Apache Software Foundation, 2009) can infer new hidden structure relationships between the eLearning chunks of knowledge. For example, in our ontology, we define that the properties *isNarrowerThan* and *isBroaderThan* as inverse properties. Assume that we describe the generality of two learning materials D1 and D2, if we assert that a learning material D1 isNarrowerThan D2, the reasoner can infer that *D2* is *BroaderThan D1*. Such new knowledge does not only enable searchers to search for a narrower learner material than D1 (which is stated in the description) but also for a wider learner material than D2 (which is inferred by the reasoner). However, to be able to infer new knowledge, defining a set of inference rules is required. Let us assume that D1, D2 and D3 represent eLearning documents (i.e., chucks of knowledge), the following are examples of inference rules that are written in normal English rather than Boolean algebra or F-logic:

- If D1 is narrower than D2 and D2 is format of D3 then D1 is narrower than D3.
- If D1 is based on D2 and D3 has part D2 then D1 is based on D3.
- If D1 requires D2 and D2 is based on D3 then D1 is based on D3.
- If D1 is more specific that D2 and D2 is format of D3 then D1 is more specific than D3.
- If D1 is version of D2 (or D1 has version D2) and D1 is format of D3 and D2 is format of D3 then D1 is alternative to D2.

AN ONTOLOGY-BASED E-LEARNING FRAMEWORK

The basic e-Learning architecture involves three component roles: an author; a learner and a repository.

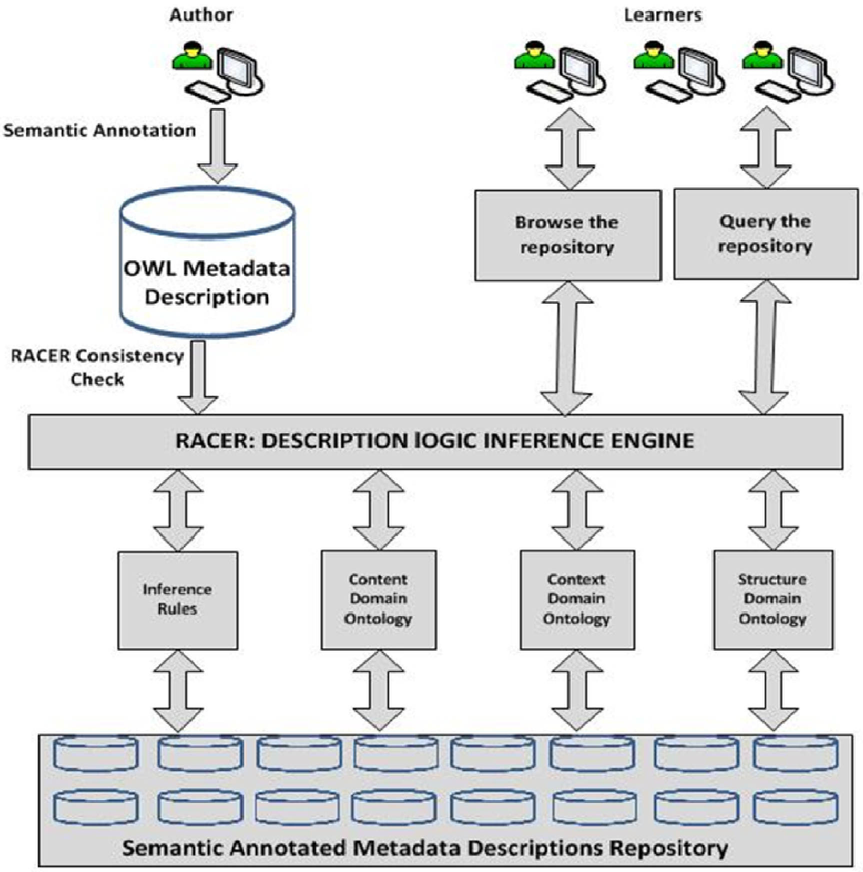

Fig. 3: A semantic-based e-learning portal architecture

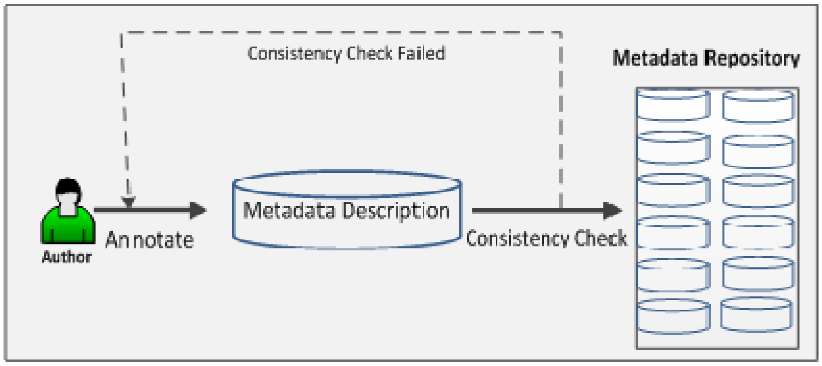

Fig. 4: A schema of providing descriptions to the repository

The author provides a semantic-based description of the eLearning material to the repository; this description includes the description of the content, context and structure metadata. The learner poses its query to the repository to search for a provided description that satisfies its needs and preferences. The repository is responsible to apply certain matching algorithms to return eLearning materials that satisfy the learner's query. Checking all available eLearning material descriptions in the repository when receiving a learner's query can be time and effort consuming. Hence, the repository can be filtered when receiving a query according to certain semantic similarities between the query and the available descriptions. Moreover, returned eLearning materials can be ranked according to their degree of match to the learner's query.

Based on this architecture we have proposed an ontology-based eLearning architecture that enables three main functions: providing descriptions, querying the repository and browsing (navigating) the repository. Figure 3 depicts the proposed eLearning architecture. Learners can query the repository to find eLearning materials that satisfy their needs. The query should encapsulate a semantic description of the content, the context and the structure of the required material. The chance to return accurate results increases when both authors and learners use the same domain ontology to describe their eLearning materials and queries respectively. Both learners and authors have the ability to browse the repository to tune their provided material and queries.

To ensure that the repository contains only valid eLearning material descriptions, it accepts only descriptions that pass a consistency check. This is one of the tasks that can be performed by description logic reasoners. Consistency check is used to ensure that all concepts and relations used in the provided descriptions are valid. Indeed, this limits the possibility of returning eLearning materials that do not satisfy the learner's need. Figure 4 depicts the process of providing descriptions to the repository.

AN E-LEARNING MATERIAL DESCRIPTION UPPER ONTOLOGY

Knowledge-based systems distinguish between two kinds of knowledge: the intensional (T Box) and extensional (A Box) knowledge. While the intensional knowledge represents general knowledge, the extensional knowledge represents more specific knowledge about the knowledge domain. T Box encapsulates the intensional knowledge in the form of terminologies that includes declaration of concepts and roles. It is the structure of the knowledge domain and the asserted axioms. A Box is a concrete example of the knowledge domain of the asserted axioms; it encapsulates examples of individuals and their properties. In the context of eLearning, Ontologies are seen as examples of T Boxes, while the provided eLearning materials submitted to the repository are seen as examples of A Boxes.

In this section, we propose an upper ontology (T Box) for describing eLearning materials. This ontology is considered as a step towards standardizing the process of describing e-Learning materials to facilitate the process of discovering and composing them efficiently. Authors of different e-Learning materials should describe their provided descriptions using the same description standard. Moreover, the process of discovering and composing e-Learning materials to build up eLearning courses would become much easier when distinct e-Learning materials are described using the same ontology.

Based on the architecture of the eLearning material portal presented in the previous section, Fig. 5 depicts an upper ontology for describing eLearning materials. As can be seen in the figure, the author provides a description of the eLearning materials. The description includes the basic three dimensions of semantic metadata: content, context and structure. The proposed ontology is implemented using Protégé (Tudorache *et al*., 2001). The ontology defines four main aspects: a hierarchy of concepts, relationships between concepts, properties of concepts and inference rules to infer new knowledge.

The ontology concepts and relationships: Figure 6, depicts part of the ontology concepts and their sub sumption and supersumption relationships. The concepts content, context and structure correspond to the description of the content, context and structure metadata respectively. The concepts atomic, collection, network, hierarchical and linear are sub concepts of the *structure* concept. They are used to describe the complexity of the eLearning materials. When the

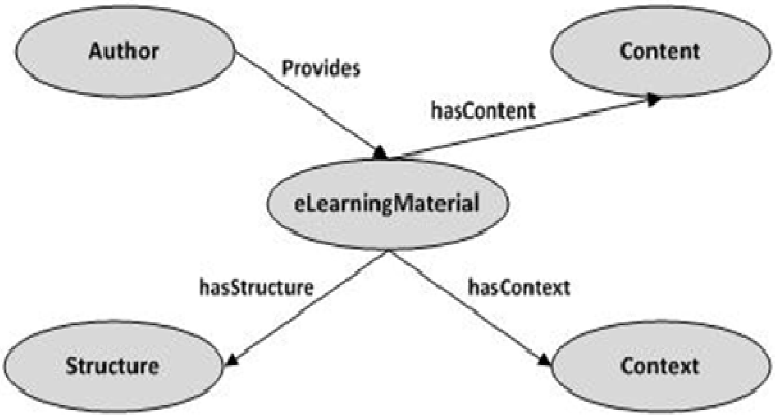

Fig. 5: The ontology main concepts

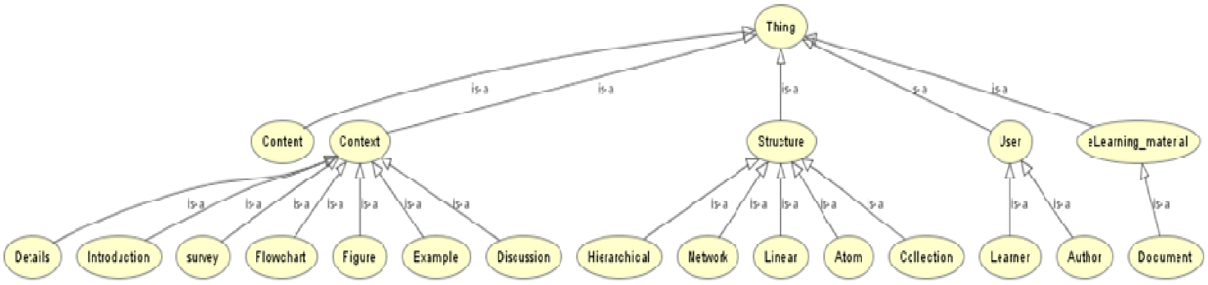

Fig. 6: Part of the e-learning ontology generated by protégé OWL Viz

Table 3: The OWL object properties (relations) used in the eLearning material ontology

Object Property	Inverse	Domain	Range
hasContent	------------------	*elearning material*	*Content*
hasStructure	------------------	*elearning material*	*Structure*
hasContext	-----------------	*elearning material*	*Context*
has Author	*isAuthoredBy*	*Elearning material*	*Author*
HasTopic	-----------------	*Content*	*Content*
HasPart	*isPartof*	*Document*	*Document*
HasVersion	*isVersionof*	*Document*	*Document*
References	*isReferencedBy*	*Document*	*Document*
isBasedon	*isBasisfor*	*Document*	*Document*
Requires	*isRequiredBy*	*Document*	*Document*
isNarrowerThan	*isBroaderThan*	*Document*	*Document*
isAlternativeTo	-------------	*Document*	*Document*
isFormatof	--------------	*Document*	*Document*
illustrates	*isillustratedBy*	*Document*	*Document*
isLessSpecificThan	*isMoreSpecificThan*	*Document*	*Document*
PrevDocument	*NextDocument*	*Document*	*Document*
FirstChildDocument	*ParentDocument*	*Document*	*Document*
RelatedDocument	-----------	*Document*	*Document*

described learning object is indivisible and does not contain any other learning objects then it is described as *atomic*, which is the simplest eLearning material type. While the *collection* concept is used to describe materials with no specified relationship between them, the *network* concept is used to describe materials with relationships that are unspecified. The concept *hierarchical* is used to describe materials whose relationships can be presented by a tree structure. Finally, eLearning materials that are fully ordered are described using the concept *linear*.

E-Learning materials can exist in different presentation forms. These forms are defined as sub concepts of the concept *context*. Hence the concept *context* has the concepts introduction, explanation, discussion, details, figure, example, discussion, survey and flow chart as sub concepts. In the ontology, we distinguish between two types of users: the author and the learner. While authors provide eLearning material descriptions to the repository, learners search for e-Learning materials to build up a course according to their preferences. learning material is composed of chunks of knowledge that need to be linked together someway and according to the learner's needs and preferences to build up the training course. Hence, we define the concept Document (which denotes a certain chunk of knowledge) to be a subconcept of the e-Learning material concept.

The ontology concepts properties: In our ontology, we distinguish between two types of properties: data type properties that link individuals to data values and object properties that link individuals to individuals. The data type properties defined in our ontology contains elements that are taken from two main resources: The general category of the LOM (Draft Standard for Learning Object Metadata, 2002) and the Dublin Core Metadata (Dublin Core Metadata Initiative, 2013). The e-learning material concept is defined to have the following data type properties: Identifier, Catalog, Entry, Title, language, Description, Keyword, Coverage, Structure, Aggregation level, Contributor, Creator, Date, Format, Publisher, Relation, Rights, Source, Subject and Type. One benefit of adopting the general category elements of the LOM standard in our ontology is that learners still can use the traditional search techniques that are based on LOM standard to search for eLearning materials in addition to the semantic-based search technique proposed in our approach. All properties and relations are inherited by subconcepts. For example, all properties mentioned above are inherited by the concept Document because it

Table 4: Some of the description logic inference rules defined in the ontology

Structure related rules	
The Rule	Description
For All D1, D2 D1: Document [parentDocument->>D2] <- D2: Document [firstchildDocument->>D1].	The relations *parentDocument* and *firstchildDocument* describe the sequence of two learner materials at two successive structure levels. Whenever a document D2 has the document D1 as its first child document, then the document D1 has the document D2 as its parent Document. This rule facilitates a semantic navigation through elearning materials (up and down one level navigation).
For All D1, D2 D1: Document [prevDocument->>D2] <- D2: Document [nextDocument->>D1].	The relations *prevDocument* and *nextDocument* describe the sequence of two learning materials at the same level in the tree structure of the elearning materials. Whenever a document D2 is known to have a document D1 as its next documents, then the document D1 has the document D2 as its previous document. This rule facilitates a semantic navigation through elearning materials (backward and forward).
For All D1, D2, D3 D1: Document [isBasedon->>D3]<- D2: Document and D1: Document [isBasedon->>D2] and D3: Document [hasPart->>D2]	Whenever a document D1 is known to be based on a document D2 and a document D3 is known to be part of the document D2 then the document D1 is based on the document D3.
For All D1, D2, D3 D1: Document [isBasedon->>D3]<- D3: Document and D1: Document [Requires->>D2] and D2: Document [isBasedon->>D3].	Whenever a document D1 requires a document D2 and the document D2 is based on a document D3 then the document D1 is based on the document D3.
For All D1, D2, D3 D1: Document [isNarrowerThan->>D3]<- D3: Document and D1: Document [isNarrowerThan->>D2] and D2: Document [isFormatof->>D3].	Whenever a document D1 is narrower than a document D2, and the document D2 is format of a document D3 then the document D1 is narrower than the document D3.
For All D1, D2, D3 D1: Document [isMoreSpecificThan->>D3]<- D3: Docuement D1: Document [isMoreSpecificThan->>D2] and D2: Document [isFormatof->>D3].	Whenever a document D1 is more specific than a document D2, and the document D2 is format of a document D3, then the document D1 is more specific than the document D3.
For All D1, D2, D3 D1: Document [isAlternativeTo->>D2]<- D3: Document and D1: Document [isVersionof->>D2] or D1: Document [hasVersion->>D2] and D1: Document [isFormatof->>D3] and D2: Document [isFormatof->>D3].	Whenever a document D1 is a version of a document D2 or the document D1 has version the document D2 and the document D1 is format of a document D3, and the document D2 is format of the document D3 then the document D1 is alternative to the document D2.
Content Related Rules	
For All C1, C2, C3 C1: Content [hasTopic->>C3]-> C3: Conent and C1: Content [hasTopic->>C2] and C2: Content [hasTopic->>C3].	This relation depicts the transitive property of the *hasTopic* relation. Whenever a content C1 has topic a content c2 and the content c2 has topic a content c3 then the content c1 has topic the content c3. For example, based on the facts that "*collections hasTopic DirectAccessCollections*" and "*DirectAccessCollections hasTopic Array"* then the fact that "*Collections hasTopic Array"* can be concluded.
For All D1, C1, C2 D1: Document [hasContent->>C1]-> C2: Conent and D1: Document [hasTopic->>C2] and C1: Conent [hasTopic->>C2].	This rule is crucial for searching purposes. For example, and based on the facts that "*Collections hasTopic Array*" and "*DirectAccessCollections hasTopic Array*" then the fact "Collections hasTopic DirectAccessCollections" can be concluded. In other words, rule ensures that whenever a document with the content Collections is searched for, then the documents about "DirectAccessCollections" and "Array" are also found.
Context Related Rules	
For All D1, D2 D2: Document [prevDocument->>D1]-> Exists CX1, CX2, C C:Content and D2: Document [hasContext->>CX2] and CX2: Example and D1 [hasContext->>CX1] and CX1: Explanation and D1[hasContent->>C] and D2 [hasContent->>C]	This rule ensures that whenever two documents D1 and D2 are talking about the same content and the document D1 is in the presentation form of *Explanation* and the document D2 is in the presentation form of *Example* then the document D2 has the document D1 as its previous document. Indeed, this is true because the example comes after the explanation.
For All D1, D2 D1:Document [nextDocument->>D2]-> Exists CX1, CX2, C C:Content and D1: Document [hasContext->>CX1] and CX1: Introduction and D2 [hasContext->>CX2] and CX2: Discussion and D1[hasContent->>C] and D2[hasContent->>C]	This rule ensures that whenever two documents D1 and D2 are talking about the same content and the document D1 is in the presentation form of *Introduction* and the document D2 is in the presentation form of *Discussion* then the document D1 has the document D2 as its next document. Indeed, this is true because the discussion comes after the introduction of the topic.

is defined to be a subclass of the concept e-learning material.

Table 3 shows the OWL object properties that are used in our ontology. The domain and range of all these properties are also shown in the table. These properties play a main role in forming inference rules that are utilized by the description logic reasoner to infer new knowledge as will be shown in the next section.

The ontology inference rules: OWL is based on description logic and every OWL description can be mapped into description logic definition. Description logic reasoning techniques can be applied to infer new knowledge. Consequently, new hidden relationships between asserted concepts can be revealed.

In building up an eLearning course scenario, this enhances the chance to locate eLearning materials that are related to the learner's needs. Indeed, locating such learning materials would not been possible if considering only the asserted model but not the inferred one. A significant feature of our approach is that it employs description logic reasoning techniques when receiving a learner's query to broaden the chance of discovering e-learning materials that satisfy the user's needs and preferences.

Description logic supports four main reasoning tasks on concepts: consistency (satisfiability), taxonomy (subsumption), equivalence and disjointness. The consistency check is essential to make sure that defining a new concept does not contradict with other defined concepts. This is done by applying what is called a satisfiability inference check. For example, if the concepts desktop application and web based application are defined as disjoint concepts, then defining a concept as a subclass of both the concepts desktop applications and web based application is considered inconsistent. This is because such a concept cannot have any instances. Indeed, something cannot be instance of desktop application and web based application at the same time.

The taxonomy inference checks whether a certain concept denotes a more general concept than another one. While the automatically computed class hierarchy that results from this reasoning task is called an *inferred ontology class hierarchy*, the manually stated class hierarchy is called an *asserted ontology class hierarchy*. For example, if we define the concept *WWW* as a subclass of the concept *Internet*, then any individual that is a member of the concept *WWW* is implicitly an individual that is a member of the concept *Internet*. Reasoning tasks can also find equivalent and disjoint concepts that are not explicitly asserted in the ontology.

The following are some of inference rules of our ontology presented in F-Logic (Kifer *et al*., 1995). The statement D1::D2 could be read as concept D1 is a subconcept of the concept D2 and the statement *D*1[*somerelation*->>D2] could be read as the concept D1 is in the relation some relation with concept D2. Table 4 depicts some of the description logic inference rules defined in our ontology.

CONCLUSION AND RECOMMENDATIONS

Recently, we have seen many proposals that aim at augmenting eLearning material descriptions with semantics. The aim of the Semantic Web is to make the Web content machine understandable besides being processable. The use of the new emergent Semantic Web technologies in eLearning systems can offer more flexibility not only in terms of locating suitable eLearning materials that satisfy the learner needs and preferences but also in terms of navigating through eLearning course.

In this study, we have proposed an ontology-based eLearning material framework that governs the interaction of the main three component roles of the eLearning architecture: the author; the learner and the repository. We have also presented a semantic-based upper ontology to describe eLearning materials. The proposed ontology is considered as a step towards standardizing the semantic annotation of eLearning material descriptions. The ontology considers semantically annotating the descriptions the three dimensions of the semantic metadata: content, context and structure. While the upper ontology is implemented using OWL, the description logic RACER is used to infer new knowledge according to predefined set of inference rules.

One dimension of future work is to extend the set of inference rules that are existed in our knowledge base to reveal more hidden knowledge on the structure, content and context levels. Another dimension of future work is to extend the proposed semantic-based eLearning portal architecture to enable ranking the existed eLearning material descriptions in the repository according to their relevance to the learner's query.

REFERENCES

Adelsberger, H., M. Bick, F. Korner and J.M. Pawlowski, 2002. Virtual education in business information systems (VAWI) facilitating collaborative development processes using the Essen Learning Model. Proceeding of the 20th JCDE World Conference on Open Learning and Distance Education, Dusseldorf, Germany, April 2002, pp: 37-46.

Adrich, C., 2004. Simulations and the Future of Learning: An Innovative (and Perhaps Revolutionary) Approach to E-Learning. John Wiley and Sons, Hoboken.

ARIADNE Foundation, 2013. Retrieved form: http:// www.ariadne-eu.org/, (Accessed on: February 2013).

Apache Software Foundation, 2009. Jena Semantic Web Framework. Retrieved form: http://jena .apache.org/index.

Berners-Lee, T., 2000. What the Semantic Web can Represent. Retrieved form: http://www. w3.org/ DesignIssues/RDFnot.html, (Accessed on: February 2013).

Bray, T., J. Paoli, C.M. Sperberg-McQueen and E. Maler, 2008. Extensible Markup Language (XML) 1.0 5th Edn., W3C Recommendation 26 November 2008. Retrieved form: Available online: http:// www. w3.org/TR/xml/, (Accessed on: January 2013).

Brickley, D. and R.V. Guha, 2004. RDF Vocabulary Description Language 1.0: RDF Schema. W3C Recommendation 10 February 2004. Retrieved form: http://www.w3.org/TR/rdf-schema/.

Chen, W., Y. Hayashi, L. Jin, I. Mitsuru and R. Mizoguchi, 1998. An ontology-based intelligent authoring tool. Proceeding of the 6th International Conference on Computers in Education, pp: 41-49.

Crampes, M. and S. Ranwez, 2000. Ontology-supported and ontology-driven conceptual navigation on the World Wide Web. Proceeding of the 11th ACM Hypertext Conference, pp: 191-199.

Defense Advanced Research Projects Agency, (DARPA) 2013. Retrieved form: http://www.darpa.mil, (Accessed on: February 2013).

Dixon, N.M., 1999. The Organizational Learning Cycle: How We Can Learn Collectively. Gower Publishing Ltd., Hampshire, England.

Draft Standard for Learning Object Metadata, 2002. IEEE 1484.12.1-2002. Retrieved form: http://ltsc.ieee.org/wg12/files/LOM_1484_12_1_v1_Final_Draft.pdf, (Accessed on: January 2013).

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

Dublin Core Metadata Initiative, 2013. Available online: http://dublincore.org/documents/dcmi-type-vocabulary/, (Accessed on: March 2013).

Engelhardt, M., A. Hildebrand, D. Lange and T.C. Schmidt, 2006. Semantic overlays in educational content networks-the hylos approach. Campus-Wide Inform. Syst., 23(4): 254-267.

Erdmann, M. and R. Studer, 2001. How to structure and access XML documents with ontologies. Data Knowl. Eng., 36(3): 317-335.

Henze, N., P. Dolog and W. Nejdl, 2004. Reasoning and ontologies for personalized e-learning in the semantic web. Educ. Technol. Soc., 7(4): 82-97.

Haarslev, V. and R. Moller, 2001. RACER system description. Proceedings of the 1st International Joint Conference on Automated Reasoning, pp: 701-706.

Horrocks, I., F. Harmelen, P. Patel-Schneider, T. Berners-Lee, D. Brickly, D. Connolly, M. Dean, S. Decker, D. Fensel, R. Fikes, P. Hayes, J. Heflin, J. Hendler, O. Lassila, D. McGuinness and L. Stein, 2001. DAML-S+OIL. Retrieved form: http://www.daml.org/2001/03/daml+oil-index.html.

IMS Global Learning Consortium, 2006. IMS Meta-data Best Practice Guide for IEEE 1484.12.1-2002 Standard for Learning Object Metadata, Version 1.3 Final Specification. Retrieved form: http://www.imsglobal.org/metadata/mdv1p3/imsmd_bestv1p3.html.

Kashyap, V., C. Bussler and M. Park, 2008. The Semantic Web: Semantics for Data and Services on the Web. Springer-Verlag, Berlin, Heidelberg.

Kifer, M., G. Lausen and J. Wu, 1995. Logical foundations of object-oriented and frame-based languages. J. ACM, 42: 741-843.

Manola, F. and E. Miller, 2004. RDF Primer W3C Recommendation 10 February 2004. Retrieved form: http://www.w3.org/TR/rdf-primer/.

Martin, D., M. Burstein, J. Hobbs, O. Lassila, D. McDermott, S. Mcllraith, S. Narayanan, M. Paolucci, B. Parsia, T. Payne, E. Sirin, N. Srinivasan and K. Sycara, 2004. OWL-S: Semantic Markup for Web Services, W3C Member Submission 22 November 2004, Retrieved form: http://www.w3.org/Submission/OWL-S/.

McGuinness, D. and F. Harmelen, 2004. OWL Web Ontology Language Overview. W3C Recommendation 10 February 2004, Retrieved form: http://www.w3.org/TR/2004/REC-owl-features-20040210/.

Naeve, A., M. Nilsson and M. Palmer, 2001. E-learning in the semantic age. Proceeding of the 2nd European Web-based Learning, Environments Conference, WBLE.

National Information Standards Organization, 2004. Understanding Metadata. (NISO), ISBN 1-880124-62-9, NISO Press, Bethesda, USA.

Noy, N.F. and D.L. McGuinness, 2001. Ontology development 101: A guide to creating your first ontology. Stanford Knowledge Systems Laboratory Technical Report KSL-01-05 and Stanford Medical Informatics Technical Report SMI-2001-0880.

Qu, C., J. Gamper and W. Neidl, 2001. A collaborative courseware generating system based on Web DAV, XML and JSP. Proceeding of IEEE International Conference on Advanced Learning Technologies-ICALT, pp: 197-198.

Sancho, P., I. Martínez and B. Fernández-Manjón, 2005. Semantic web technologies applied to e-learning personalization. J. Univer. Comput. Sci., (9): 1470-1481.

Semantic Web Activity Statement, 2013. W3C Retrieved form: http://www.w3.org/2001/sw/.

Sirin, E., B. Parsia, B.C. Grau, A. Kalyanpur and Y. Katz, 2007. Pellet: A practical OWL-DL reasoner. Web Semantic: Sci. Serv. Agents World Wide Web Softw. Eng. Semantic Web, 5(2): 51-53.

Stojanovic, L., S. Staab and R. Studer, 2001. E-Learning based on the Semantic Web. Paper presented at the World Conference on the WWW and Internet (WEBNET 2001), Orlando, Florida, USA, pp: 23-27,

Tudorache, T., C. Nyulas, N. Noy and M. Musen, 2011. Web Protégé: A Collaborative Ontology Editor and Knowledge Acquisition Tool for the Web. Semantic Web J., IOS Press, pp: 11-165.

Ullrich, C., 2004. Description of an instructional ontology and its application in web services for education. Proceedings of the Workshop on Applications of Semantic Web Technologies for E-Learning SW-WL04, pp: 17-23.