Solomou, G., Pierrakeas, C., & Kameas, A. (2015). Characterization of Educational Resources in e-Learning Systems Using an Educational Metadata Profile. *Educational Technology & Society, 18* (4), 246–260.

Characterization of Educational Resources in e-Learning Systems Using an Educational Metadata Profile

Georgia Solomou[1*], Christos Pierrakeas[1,2] and Achilles Kameas[1]
[1]Hellenic Open University (HOU), Patras, Greece // [2]Technological Educational Institute (TEI) of Western Greece, Patras, Greece // solomou@eap.gr // pierrakeas@eap.gr // kameas@eap.gr
[*]Corresponding author

(Submitted March 27, 2014; Revised March 23, 2015; Accepted May 1, 2015)

ABSTRACT

The ability to effectively administrate educational resources in terms of accessibility, reusability and interoperability lies in the adoption of an appropriate metadata schema, able of adequately describing them. A considerable number of different educational metadata schemas can be found in literature, with the IEEE LOM being the most widely known; however, it is often the case where it cannot fully accommodate the characteristics of several types of educational resources, that's why application profiles have been proposed. Each metadata standard and application profile usually comes with a different (either less or more semantically enriched) binding, thus allowing the retrieval and dissemination of resources with varying degrees of effectiveness. In this work, we propose an application profile of the IEEE LOM standard having special focus on distance learning material, while being generic enough so as to be applicable to any educational material and application. We then present an ontology model for this profile that aims to improve the potential discovery and retrieval of educational resources within intelligent e-learning environments.

Keywords

Learning objects, Ontologies, Metadata application profile distance education, IEEE LOM

Introduction

Metadata are "*machine-readable information about electronic resources or other things*" (Berners-Lee, 1997) and are used to describe the features of a resource, thus making easier its management and retrieval. A set of metadata elements combined so as to serve a specific purpose, constitute a *metadata schema*.

Although the adoption of a single metadata standard would assure reusability of resources and interoperability among applications, there exists no metadata schema yet, appropriate to fulfil the requirements and needs of every application. Some schemas focus on technical metadata, other on educational metadata while some other on more specialized elements. When existing approaches are not sufficient enough to cover the special requirements of an institution or organization, the use of *application profiles* is suggested. According to Heery & Patel (2000), an application profile is *an aggregation of metadata elements selected from one or more different schemas and combined into a new compound schema.*

Particularly, in the case of educational recourses, the set of metadata used to describe their characteristics, should be able to capture their educational and pedagogical aspects. Therefore, apart from *author*, *title* or *type* – fields that are common in all metadata schemas – an educational metadata schema should also include information regarding the resource's particular learning type, its intended end users, the instructional context and many more.

A kind of educational resource that is increasingly used by Educational Institutions in recent years is the Learning Object (LO). According to Nikolopoulos, Solomou, Pierrakeas & Kameas (2012) LOs are pieces of educational material that directly correlate the knowledge they convey with specific objectives (learning outcomes) of the learning process. But although LOs constitute a common trend in organizing educational material and have been utilized by many modern e-learning systems (Schreurs & Al-Zoubi, 2007), they cannot be used effectively because there exists no metadata schema capable of capturing all of their characteristics.

This insufficiency becomes even greater in the case of LOs that are designed for use in the context of distance learning courses, where the proper handling and dissemination of LOs is crucial for the success of the learning process, because in most cases, contrary to what happens in a classroom, no human tutor would be continuously available to monitor students' path or progress through the educational process.

Hellenic Open University (HOU) is a Higher Education Institute specialized in distance and lifelong learning that the last two years seeks to re-organize its material and to provide its students with advanced services for delivering knowledge. Such services require the consumption of adequately characterized LOs, using a metadata schema that is capable of capturing as many as possible of their pedagogical aspects and especially those considered to be important according to distance learning principles. To the best of our knowledge, no such a schema or application profile exists, able to satisfy these requirements.

Consequently, through this work we propose an application profile of the IEEE LOM standard with special focus on the field of distance learning. After reviewing existing approaches for describing educational resources, as well as several binding methods (section *Background*), we move on to the presentation of our proposed Educational Metadata Profile (EMP), in section *EMP: New Elements and Modifications*. Its ontological binding is given in the subsequent section (*The EMP ontology*), whereas section *Evaluation of the EMP ontology* presents an evaluation of this ontology model, through its application for characterizing real LO instances. Conclusions follow, in our last section.

Background

In literature, several metadata standards and profiles have been proposed, each serving different purposes and needs. A few examples of them are mentioned in the following subsection.

Educational metadata standards and profiles

The IEEE Learning Technology Standards Committee (LTSC) has developed a standard for the description of learning material and learning resources, known as IEEE Learning Object Metadata (IEEE LOM) (Hodgins & Duval, 2002). IEEE LOM is without doubt a widespread standard for educational metadata and focuses mainly on the description of educational resources and especially LOs. It includes more than 60 elements classified into 9 categories (*General*, *Life Cycle*, *Meta-Metadata*, *Technical*, *Educational*, *Rights*, *Relation*, *Annotation*, *Classification*), each one of them containing metadata for various aspects of a LO, including its technical characteristics and rights, as well as educational and instructional features.

Several IEEE LOM profiles can be found in literature. Each of them is usually designed so as to accommodate very specific needs of its corresponding organization/creator. ARIADNE (see http://www.ariadne-eu.org/) is one such profile that intends to describe learning material used in secondary and post-secondary education. It is designed with the aim to solve two major problems: (a) the indexing of educational resources (i.e., the creation of the metadata by persons) and (b) the exploitation of metadata by users who look for relevant pedagogical material (that should be as easy and efficient as possible).

Another IEEE LOM profile is IMS Learning Resource Metadata (IMS LRM) (Nilsson, 2001), which constitutes a set of specifications for learning resources. It includes elements useful for the description of learning resources, while the specifications address issues like content packaging, question and test interoperability, learning design and simple sequencing. IMS LRM adopts all of the LOM’s categories and elements.

Two more LOM application profiles, that were created in order to describe resources locally, are CanCore (Canadian Core) (http://cancore.athabascau.ca/en/) and UK LOM Core (http://zope.cetis.ac.uk/profiles/uklomcore/). CanCore, used mainly in Canada, simplifies LOM and at the same time maximizes interoperability between different projects. UK LOM Core, designed for United Kingdom educational system, intends to provide guidelines to those who desire to create, use and apply metadata.

The Sharable Content Object Reference Model (SCORM) is a reference model which controls how the learning content is organized, described and linked with Learning Management Systems. SCORM allows the extension of LOM, thus enabling organizations to add new elements and enhance the existing controlled vocabularies.

On the other hand, the Dublin Core Metadata Initiative (DCMI) (DCMI, 2012) has been developed by organizations so as to aid the sharing of any kind of generic web resources. Its initial version, the Dublin Core Metadata Element Set (DCMES) (DCMI, 2012), known as Dublin Core (DC), consists of 15 elements. The subsequent Qualified Dublin Core (QDC) (DCMI Usage Board, 2012) extends DC with 7 new elements. However, even this enriched version of the DC schema is unable to capture the pedagogical aspects of an educational resource.

GEM (Gateway to Educational Materials) (http://dublincore.org/groups/education/GEM-Study.html) is a DCMI profile that is education-oriented. It is an RDF metadata vocabulary, designed for the description of educational resources and, apart from the DCMI elements, encompasses several additional educational-specific properties. GEM includes controlled vocabularies for the different levels of end users, evaluation methods and tools as well as the types of resources.

Although all of the aforementioned standards and profiles manage to capture some of the most important pedagogical characteristics of educational resources, they become weak when applied to distance education, because, in order to be effective, the educational material designed to serve distance learning courses has to conform to specific requirements regarding its content, layout, structure, technical properties, etc. (Lockwood, 2013). This material assumes the role of the tutor of a distance learner, so it has to be explicitly correlated to learning outcomes and be available in various formats, in order to be able to cover various learning styles. So, even the very promising IEEE LOM standard, despite its generality, still fails to specify some of these important educational aspects of learning resources, like for example the expected learning outcomes of a course, the direct correlation among learning outcomes and educational material, as well as the different (learning) types of the latter that are used within a distance learning course.

Binding of metadata elements

The elements of a metadata schema are usually handled as SQL tables, text files, HTML meta-tags and so on Nejdl et al. (2002). Such a technical realization of the abstract model of a metadata schema in a specific format is called a *binding*. Structured formats for binding the elements of a metadata schema are XML, XML Schema, RDF, RDF Schema and OWL.

XML provides a surface syntax for structured documents, but imposes no semantic constraints on the meaning of these documents. XML Schema is a language for restricting the structure of XML documents and also extends XML with datatypes. The Resource Description Framework (RDF) (RDF Working Group, 2014) is a datamodel for objects (“resources”) and relations between them and provides simple semantics which in turn can be represented using XML syntax. RDF Schema is a vocabulary for describing properties and classes of RDF resources, with the necessary semantics for generalization – hierarchies of such properties and classes. Even though RDF is intended for representing knowledge, it lacks reasoning abilities; RDF does not support making inferences or deductions. Therefore, a much more expressive framework is required, so that metadata can be meaningfully encoded.

Ontologies, expressed in OWL (OWL Working Group, 2014), are the pillar of the Semantic Web and provide the ability to represent any domain of interest in a more structured way. OWL provides the vocabulary for describing properties and classes (relations between classes, cardinality, equality, richer typing of properties, characteristics of properties and enumerated classes) and poses constraints about what statements the user can declare.

For IEEE LOM, XML and RDF bindings are available, implemented by the IMS Global Learning Consortium (Nilsson, 2001; Nilsson, Palmér & Brase, 2003). The usage of XML for the LO metadata expression facilitates the indexing process and the retrieval of annotated learning resources. However, this format seems to be not sufficient enough to address the limitation of text-based searching, since XML does not provide the meaning of the described structures. RDF attempts to overcome the problem by adding semantics to each metadata element. Therefore, the description of LOM elements via RDF facilitates their integration into e-learning systems which nowadays are dominated by Semantic Web technologies.
A great number of web-based educational systems exist that embed ontological models in their implementation (Al-Khalifa & Hugh, 2006). These ontological models can reflect various aspects of an e-learning system, such as

student profiles and knowledge domains. The integration of an educational resource's components to such systems requires semantically richer representation (http://dublincore.org/groups/education/GEM-Study.html). Therefore, many research groups have attempted to annotate semantically the educational metadata. Some representative examples are ALOCoM, SCORM, and an ontology based on the ACM Computer Classification System (ACM CCS).

More specifically, the ALOCoM ontology (http://hmdb.cs.kuleuven.be/alocom/) consists of three parts: the ALOCoM Core ontology, the ALOCoM Content Structure and the ALOCoM Content Type. It has been used in the LO repository of the ARIADNE Foundation as a format to store well-structured and easy-to-reuse LOs. It is a generic content model that defines a framework for LOs and their components. That is, it describes the structural elements of a LO, it focuses on potential pedagogical roles of content units and defines concepts that describe the components common to any type of a LO.

In Yang, Chen, Tsai and Chao (2005) a systematic approach called the Visualized Online Authoring Tool model (VOAT) is developed. Using this model, educators and instruction designers can construct SCORM 2004-compliant courses. Through the Ontology-Based Outline Authoring Tool (OBOAT) educators can be guided by the domain ontology and thus more accurately construct an outline of their course, based on their own individual teaching contexts (beliefs, preferences and student characteristics).

Finally, a well-known example of ontology for the computer science domain has been based upon the ACM Computer Classification System (http://www.acm.org/class) and is defined using RDFS (Brace & Nejdl, 2004). This ontology, which is specific to the educational resources purpose and structure, has been used in the Edutella system (Nejdl et al., 2002) to classify LOs in order to improve searching.

All of the aforementioned ontologies reveal an attempt by educational content providers to create richer representations of the metadata elements they use to characterize their resources. Such representations can have many advantages in the learning process, from alleviating the design of a course to improving the discovery of educational resources. In addition, ontologies are fundamental to improve interaction with users and intelligent agent systems and make systems open to external knowledge integration (Wray, Lisse & Beard, 2004).

EMP: New elements and modifications

In order to create an educational metadata schema suitable for the efficient characterization of distance learning material, the design principles of distance learning should be taken into account. Such a schema should be flexible and functional, so that it can capture a large number of educational aspects and pedagogical features of educational resources and make them a prominent means in the knowledge discovery and delivery process. In addition, it should be in accordance with the requirements of other structures that manage educational resources, like institutional libraries, which are usually based on a cataloguing standard.

Having these in mind, we propose a new application profile of the IEEE LOM standard that, unlike all other existing profiles, makes provision for the particular features of the distance learning material. We opted for IEEE LOM, due to its relatively wide acceptance in the academic environment and its extensive usage by institutional repositories. To build this profile, we took into account the guidelines provided by CEN/ISSS (Smith, Van Coillie & Duval, 2006), according to which we had to accomplish the following steps:

- identify the specific characteristics of the distance learning material;
- identify which of these characteristics are reflected in the standard, existing elements of the base schema (IEEE LOM);
- modify the base metadata schema according to these specific requirements (extend it with additional, new elements, modify value space and/or data type of existing elements);
- provide a binding

The proposed Educational Metadata Profile (EMP) adopts a subset of the IEEE LOM element set and augments it with some new attributes in order to represent concepts commonly used in distance education. This new application profile has been designed to be rich enough, so as to effectively describe some of the most important aspects of an educational resource (pedagogical, technical, etc.), but not exceedingly analytic as to become difficult to use.

In what follows (subsection *Structure*) we give the list of EMP elements and summarize the modifications and additions that we made to the base schema. In the next subsection *Mapping to DCMI metadata elements*, we provide a mapping of the resulting schema to the widely accepted DC metadata standard.

Structure

Table 1 summarizes the structure elements of the proposed EMP. Those that have been directly taken from IEEE LOM are considered to come with associated semantics. All other additions and modification have been categorized into three groups, as follows:

- The first group (Group A) reflects elements that appear with differences in the value space of their controlled vocabulary, as compared to their counterparts in the original IEEE LOM schema.
- The second group (Group B) refers to elements that come with modifications in their definition and data type only.
- Finally, the third group (Group C) includes completely new elements, with no counterparts in the original IEEE LOM schema. These new entries address several important characteristics of the distance learning material.

Table 1. The elements of the proposed Educational Metadata Profile (EMP)

Name	# Repetitions	Datatype
General		
Identifier[1]	10	-
Identifier. Catalog	1	CharacterString
Identifier.Entry	1	CharacterString
Title	1	LangString
Language	10	CharacterString
Description	10	LangString
Keyword	10	LangString
Aggregation Level	1	Controlled Vocabulary
Life Cycle		
Contribute[1]	30	-
Contribute.Role	1	Controlled Vocabulary
Contribute.Entity	40	CharacterString
Contribute:Entity:Affiliation***	1	CharacterString
Contribute.Date	1	DateTime
Technical		
Format*	40	Controlled Vocabulary
Size	1	Integer
Requirement**	40	LangString
Duration	1	CharacterString
Educational		
Learning Resource Type*	8	Controlled Vocabulary
Intended End User Role	10	Controlled Vocabulary
Instructional Context***	10	Controlled Vocabulary
Typical Age Range	1	LangString
Difficulty	1	Controlled Vocabulary
Typical Learning Time	1	CharacterString
Learning Outcome[1]	1000	-
Learning Outcome.Identifier[1]	10	-
Learning Outcome.Identifier.Catalog***	1	CharacterString
Learning Outcome.Identifier.Entry***	1	CharacterString

Learning Outcome.Description***	1	LangString
Rights		
Copyright	1	LangString
Description	1	LangString
Relation		
Relation Kind	1	Controlled vocabulary
Resource[1]	1	-
Resource.Identifier[1]	10	-
Resource.Identifier.Catalog	1	CharacterString
Resource.Identifier.Entry	1	CharacterString

Note. *Element with modified value space (Group A); **Element with modified definition and data type (Group B); ***New element (Group C); [1]indicate subcategories of elements that are only populated through their corresponding refinements.

Group A summarizes the elements that come with modifications in their predefined set of accepted values. These modifications reflect our attempt to meet more accurately the specific characteristics of a resource. The elements that have changed in this way are the following: (1) *Format* of the *Technical* category, (2) *Learning Resource Type* of the *Educational* category and (3) *Kind* of the *Relation* category.

More specifically, *Format* has now a new set of allowable values, as summarized in Table 2. This set is based on the official IANA MIME media types (Freed, Klensin, & Hansen, 2013) and has emerged after considering the characteristics of the educational material that is already used by HOU. Consequently, it was carefully selected so as to be quite broad and able to cover a wide range of technical data types met in a distance education institution.

Table 2. Possible values of the format element (i.e., technical data types of the resource)

Element	Value space
Text	document hypertext
Image	photo map graph image presentation
Streaming Media	audio recording animation self-running presentation webcast video
Application	interactive software hypermedia application wiki presentation

Learning Resource Type comes with substantial modifications in the set of its accepted values. The main problem with its original value space – as given in the IEEE LOM specification – is that it consists of values which express both educational information (e.g., Exercise, Problem Statement, Simulation) and technical information (e.g., Diagram, Figure, Graph) of an educational resource. In our schema, the technical information (form) is captured by the *Format* element, so *Learning Resource Type* should normally address only the various types of educational resources according their educational content. In addition to this separation, some important types of educational resources, such as Example, Serious Game, Case Study and Project, which were absent from the original value space, were added.

To this end, we define a completely new list of acceptable values that reflect the most common types of educational material used within distance education courses and incorporate only information related to the instructional perspective of a resource. This list of values is, to a certain degree, based on the content object types provided by the ALOCoM generic content model. The complete list of the learning types we propose is presented in Table 3.

The value space of the *Kind* element consists of a controlled vocabulary expressing the various kinds of relationships among educational resources and especially LOs. But apart from the default relationship *has part* and its inverse *is part of*, two additional types of relationships are needed so as to capture all possible interconnections among LOs:

- *supports* and its inverse *is supported by* attempts to correlate a "supportive" LO that contains complementary or prerequisite knowledge, with an LO that has a key role in the learning process (a "core" LO).
- *is alternative type* correlates two or more LOs that have exactly the same educational content and differ only in their technical format. This is a highly significant relationship, especially in the case of personalized learning because it makes use of the learners' preferences.

The only element of the IEEE LOM schema that has been adopted with different definition and data type (Group B of modification), is *Requirement*, which is used to describe any particular need in terms of software or hardware. The data type of *Requirement* has been altered to LangString. Therefore, the value of this element contains a short description of any special software or hardware requirements, e.g., "*The use of Adobe Acrobat Reader required, version 6.xx or newer.*" The corresponding element in the original LOM schema required a more detailed description and more fields to be filled. We thus simplify the process of describing the technical requirements of a resource, keeping at the same time all necessary information intact.

Table 3. Possible values of the learning resource type element

Learning resource type	
Guidelines	Self-Assessment
Presentation	Multiple Choice Questions
Demonstration	Open Type Question
Lecture	Problem Statement
Definition-Principle-Law	Experiment
Narrative Text	Serious Game
Analogy	Exercise
Example	Multiple Choice Questions
Activity	Open Type Question
Case Study	Problem Statement
Problem Solving	Project
Text Composition	
Question	
Simulation	
Interactive	
Non Interactive	

Finally, we have augmented our schema with some new elements (Group C) that are necessary in order to represent concepts commonly used in distance education.

- *Affiliation* keeps important information about the life cycle of a resource as it determines the status of the entity that has contributed to the creation and development of this particular resource.
- *Learning Outcome is* placed under the *Educational* category and expresses the correlation of an LO with one or more learning outcomes. In particular, for each learning outcome that an LO satisfies, one needs to give a natural language statement, via the *Description* element, as well as to assign to it an identifier (*Identifier:Entry* and *Identifier:Catalog* respectively), according to a specific identification system.
- *Instructional Context* implies the actual context where the learning process takes place, and can accept values like "distance education," "face to face learning" and "blended learning." This is a key element because it reveals the mode of learning for which the particular object is appropriate.

Elements in Group A and C mostly address modifications/additions that render our EMP suitable for adequately describing educational resources used in the context of a distance learning course. New elements have been introduced (*Learning Outcome, Instructional Context*), additional values for the *Learning Resource Type* have been provided – capturing the various types of distance learning material – as well as new relationships that directly correlate learning outcomes and LOs, have been declared.

Mapping to DCMI metadata elements

The EMP described above has particular orientation in education. Nevertheless, to facilitate the potential implementation of our EMP by applications based on the widely used DC schema, we provide a mapping of the EMP to DCMI metadata elements (see Table 4). Apart from those elements that had a direct correspondence to DC (e.g., title, language), the rest of EMP elements are mapped to those DC terms that are closer in meaning.

Table 4. Mapping of the EMP elements to the DC metadata terms

IEEE LOM (EMP)	DC metadata terms
General	
Identifier.Catalog	isPartOf
Iedntifier.Entry	identifier
Title	title
Language	language
Description	description
Keyword	subject
Aggregation Level	isPartOf
Life Cycle	
Contribute.Role	contributor, creator, publisher
Contribute.Entity	description
Contribute.Entity.Affiliation	description
Contribute.Date	date
Technical	
Format	format
Size	extent
Requirement	requires
Duration	extent
Educational	
Learning Resource Type	type
Intended End User Role	audience
Context	educationLevel
Typical Age Range	description
Difficulty	type
Typical Learning Time	temporal
LearningOutcome.Identifier.Catalog	isPartOf
LearningOutcome.Identifier.Entry	identifier
LearningOutcome.Identifier.Description	description
Rights	
Copyright	rights
Description	description
Relation	
Kind	relation
Resource.Identifier.Catalog	isPartOf
Resource.Identifier.Entry	identifier

This mapping is a means to assure interoperability with many applications and educational institutions worldwide that adopt the DC metadata. Besides, DC is supported by default by some of the most well-known digital repository systems, like DSpace, EPrints and Digital Commons (Castagné, 2013).

The EMP ontology

In this section, we define the EMP binding, by specifying how a learning technology system will represent or use a metadata instance of an educational resource that follows the EMP. Ontologies have already been used to this end (see subsection *Binding of Metadata Elements*) and seem a very promising approach, given that they manage to transform the textual information captured by a metadata instance into a machine-understandable format.

Here, we describe the OWL binding of EMP, that is the ontological model used to represent the structure and characteristics of an LO, as these have been defined in section *EMP: New Elements and Modifications*. In order to build our *EMP ontology* we followed a widely-adopted methodology, proposed in Noy & McGuiness (2001). For its formal representation, we were based on OWL 2 whereas for creating and managing our ontology, we used the Protégé editor.

Ontology structure

The "LearningObject" class is a top class used to capture the notion of an LO, or an educational resource in general. The various characteristics of an educational resource are expressed as either classes or properties in the ontology. The elements *Title*, *Language*, *Description*, and *Aggregation Level* of the *General* category of the EMP, are expressed via the datatype properties "title," "language," "description" and "aggregation level" respectively. We opted for datatype- and not object- properties given that all of these elements simply assign values to some of the resource's basic characteristics and express no correlations among other entities. More specifically, the "aggregation level" is an integer datatype property that can take values from 1 to 4. The "language" property can have as fillers any of the known language identifiers, like "en" for English or "el" for Greek.

The *Contribute* element of the *Life Cycle* category is represented by the "Contributor" class. Refinements of the *Contribute* element (*Contribute:Date*, *Contribute:Entity:Affiliation*, *Contribute:Role*) are expressed via the corresponding datatype properties ("contributeDate," "affiliation" and "contributorRole," respectively). "contributorRole" is an enumerated datatype property that can be filled with the values *publisher*, *creator*, and *reviewer*, all corresponding to several important roles in the life cycle of a learning resource. Instances of the class "Contributor" are related to instances of the "LearningObject" class via the "contributor" object property.

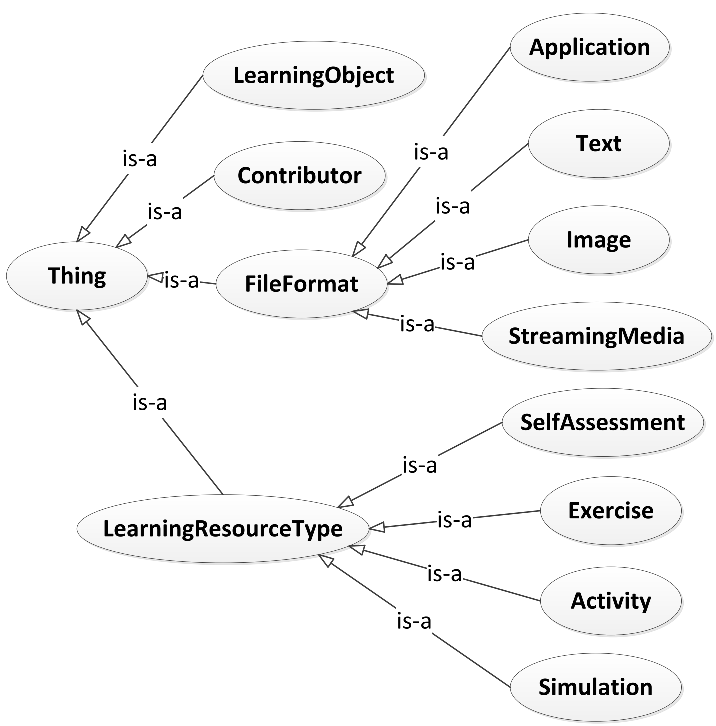

Figure 1. The class hierarchy in the LO ontology

To express the format of an educational resource (captured by the element *Format* of the *Technical* category), the “FileFormat” class is introduced in our ontological schema. In order to capture its top-four possible values (*Text*, *Image*, *Streaming Media* and *Application*) the corresponding subclasses were placed under the “FileFormat” class (see *Figure 1*). Their refinements (as presented in Table 2) became instances of these subclasses. The remaining elements of the *Technical* category – expressing the physical size, special software or hardware requirements, as well as the time duration for streaming media – are described by the datatype property “size,” “requirement” and “duration,” respectively.

The *Learning Resource Type* element, which is used to specify the different educational types of LOs, is captured by the “LearningResourceType” class. In the EMP, this element is associated with a predefined list of terms (see Table 3). Each such term is represented as an instance in the “LearningResourceType” class, with the exception of *Self-Assessment*, *Activity*, *Simulation* and *Exercise*, which have been placed as subclasses of the main “LearningResourceType” class. The ontological structure of the *Learning Resource Type* element is indicated in Figure 2.

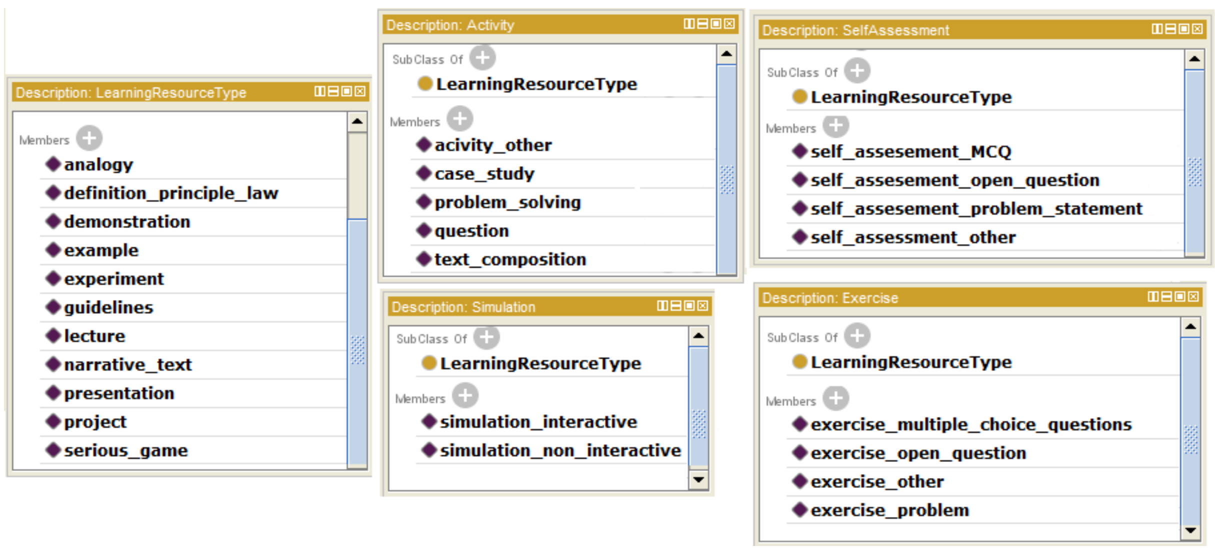

Figure 2. The Learning Resource Type category as class in the EMP ontology

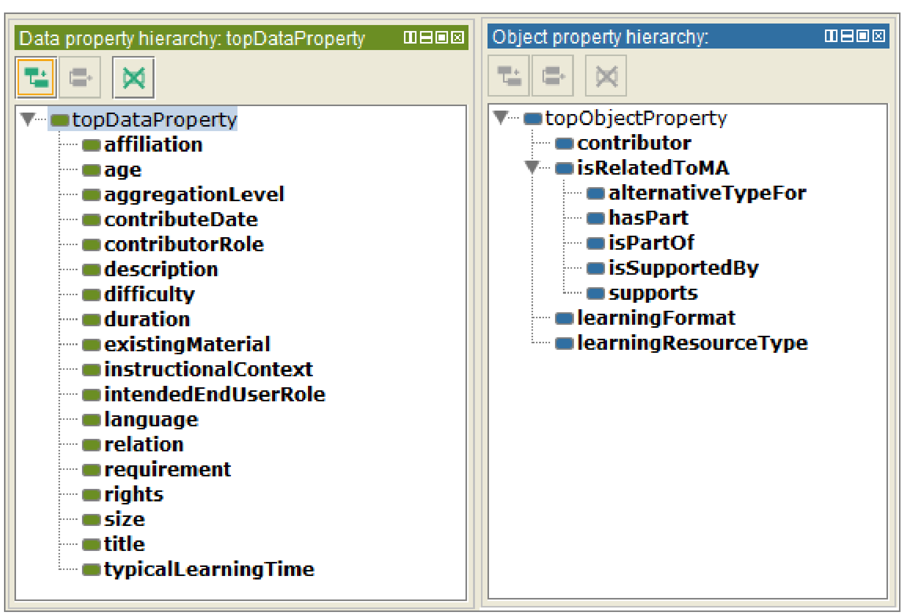

Figure 3. Object and Datatype Properties in the LO ontology

Self-Assessment, *Activity*, *Simulation* and *Exercise* are subclasses of the *LearningResourceType* class, whereas their refinements are captured as instances.

The elements of the EMP *Education* category, which express concepts like the intended end users, the instructional context, the age range of end users, as well as difficulty level and average learning time, are represented by object properties ("intendedEndUserRole," "instructionalContext," "age," "difficulty," "typicalLearningTime," respectively – see Figure 3). Finally, the datatype property "rights" has been defined in order to express the copyright data that apply.

The potential relationships among LOs, as these have been described in section *EMP: New Elements and Modifications*, are represented as object properties; these properties can be used to correlate instances of the "LearningObject" class. All of them are sub-properties of the main property "isRelatedtoMA" (see Figure 3).

Combination with other ontologies

Some elements of the EMP have been captured as object properties in the EMP ontology, meaning that they correlate a resource with concepts (rather than literal values) originating from other supporting ontologies. Such supporting ontologies can be the LearningOutcome ontology presented in Kalou, Solomou, Pierrakeas & Kameas (2012), the well-known FOAF ontology (http://xmlns.com/foaf/spec/) and ontologies that represent knowledge domains.

In particular, the element *Keyword* of the EMP *General* category is expressed via the object property "subject." This property links instances of the "LearningObject" class with class instances representing specific knowledge domain concepts. By this way, a network of interrelations among LOs and knowledge domains is created, thus allowing for the effective management and reusability of resources. Besides, it is crucial in a distance learning course to efficiently discover the appropriate educational material and provide it to the learners in the right order.

The "Contributor" class of our EMP ontology has been set as equivalent to the "Agent" class of the FOAF ontology. We, therefore, take advantage of the FOAF's built-in object properties "foaf:name" and "foaf:surname" so as to formulate the *Life Cycle* element *Contribute:Entity*. In this way, the contributor of a resource is represented in an interoperable way and can be recognized in the context of other applications that also adopt the FOAF ontology.

Finally, in our EMP ontology we have introduced properties that combine LOs with concepts from the LearningOutcome ontology, to be used by LOs that have been designed to satisfy specific learning outcomes. More specifically, we determined the object properties "satisfies" and "satisfiesInd," both of which associate instances of the "LearningObject" class with instances from "LearningOutcome" class, directly and indirectly, respectively. The indirect correlation of LOs can be only elaborated by restriction rules, expressed in the Semantic Web Rule Language (SWRL), and by exploiting existing reasoning mechanisms. By using both of these properties, we can better handle LOs within a distance learning course, given that we can retrieve them based on their association to the goals of learning.

Evaluation of the EMP ontology

To evaluate the potential of the proposed EMP, we exploited it for describing LOs designed to serve several distance learning courses of HOU. In particular, we have taken advantage of the EMP's ontological binding and used it to characterize LOs for the following courses: Object Oriented Programming (Java), C Programming Language, Introduction to Computer Science, Programming Techniques and Data Structures.

In this section, we present an example use of the EMP ontology in the knowledge domain of Object Oriented Programming (Java). Our main goal is to illustrate how such an ontological binding of an educational metadata schema could lead to a better organization and discovery of resources within intelligent e-learning applications and allow for their reusability among different educational contexts.

Population of the EMP ontology

The Java programming language is a domain of knowledge covered by the HOU study course module of "Software Engineering." An ontological representation for this knowledge domain is already available and is described in Kouneli, Solomou, Pierrakeas & Kameas (2012).

With the aid of the Protégé editor, we combined the EMP ontology with the Java ontology and the LearningOutcome ontology, presented in Kalou et al. (2012). The resulting (combined) ontology constitutes a rich knowledge representation, capable of effectively characterizing, organizing and correlating LOs and learning outcomes for the Java course, via semantic relationships.

The resulting combined ontology was populated with real individuals and in particular with LOs that have been developed for the course of Java. For every LO we created an instance in the combined ontology, by making the corresponding *object*- and *datatype*- assertions for it. We finally got 24 instances in total. An example LO instance, as this is captured in Protégé, is depicted in Figure 4.

Using the "subject" property we correlated a LO with concepts from the knowledge domain of the Java programming language (Figure 4). Nevertheless, in the place of Java, any other knowledge domain ontology could be imported, and thus render our combined ontological schema able to handle LOs of a different course.

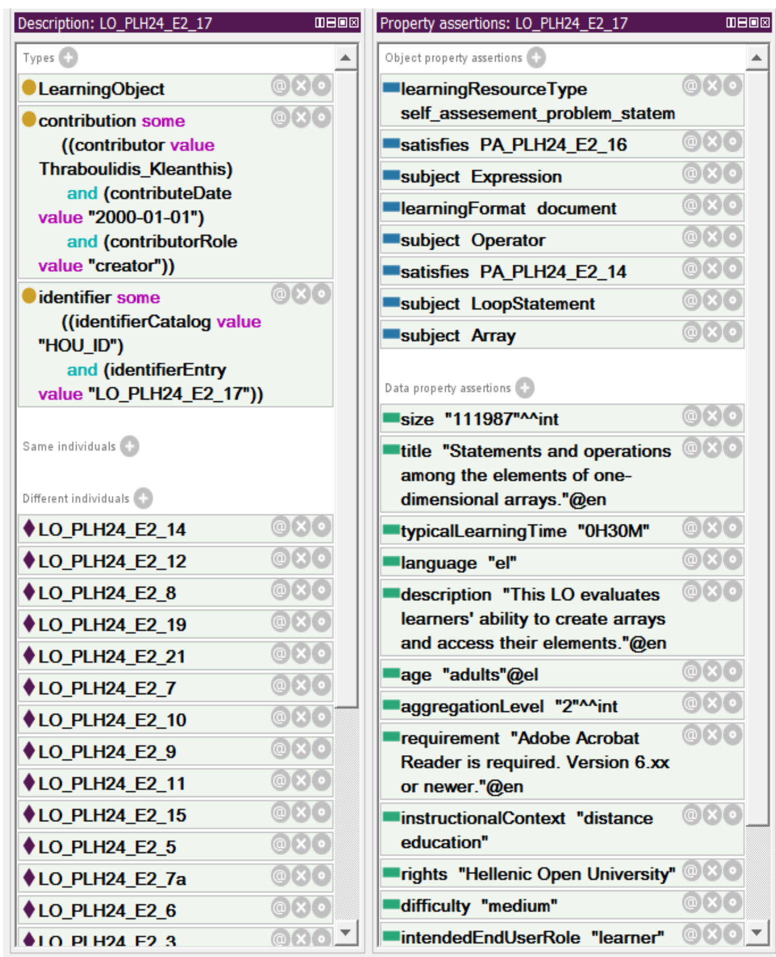

Figure 4. Example instance of LO with ID "LO_PLH_24_E2_17" in the combined ontology

The property "satisfies" links an LO with instances of the LearningOutcome ontology. Consequently, the various semantic relationships that have been declared among the learning outcome instances can yield implicit – and very useful in terms of discovery – relationships for the corresponding LOs.

Example queries

To determine the capabilities of the proposed ontological model we ran semantic queries, and evaluated them against the populated ontology. The aim was to examine our model's capability to infer knowledge. These queries are to be used in the context of an intelligent tutoring system, able to take advantage of this model, so they are expressed in a formal semantic query language. More specifically, they are expressed in the Manchester OWL Syntax (Horridge & Patel-Schneider, 2008) and for the time being they have been tested against the DL query tab of Protégé. We actually want to demonstrate that apart from running simple lookup queries, based on matching mere literal values, we can request more complex answers, based on the semantic relationships that an ontology allows.

Consider, for example, that a Learning Management System operating on behalf of a tutor or, as would be the most frequent case in Distance Learning, by a student, has to obtain all LOs that are *difficult* to learn, satisfy learning outcomes that concern *Java operators* and fall into the *Cognitive Domain* of the Bloom Taxonomy. This is expressed by the query#1 of Table 5. With query#2, we can retrieve those LOs having as learning resource type *narrative text*. Finally, with query#3, we can retrieve all LOs that either explicitly (through "satisfies" property) or implicitly (through "satisfiesInd" property) lead to the completion of learning outcomes related to Java operators. The results of these queries, as evaluated through the query tab of Protégé, can be seen in Figure 5.

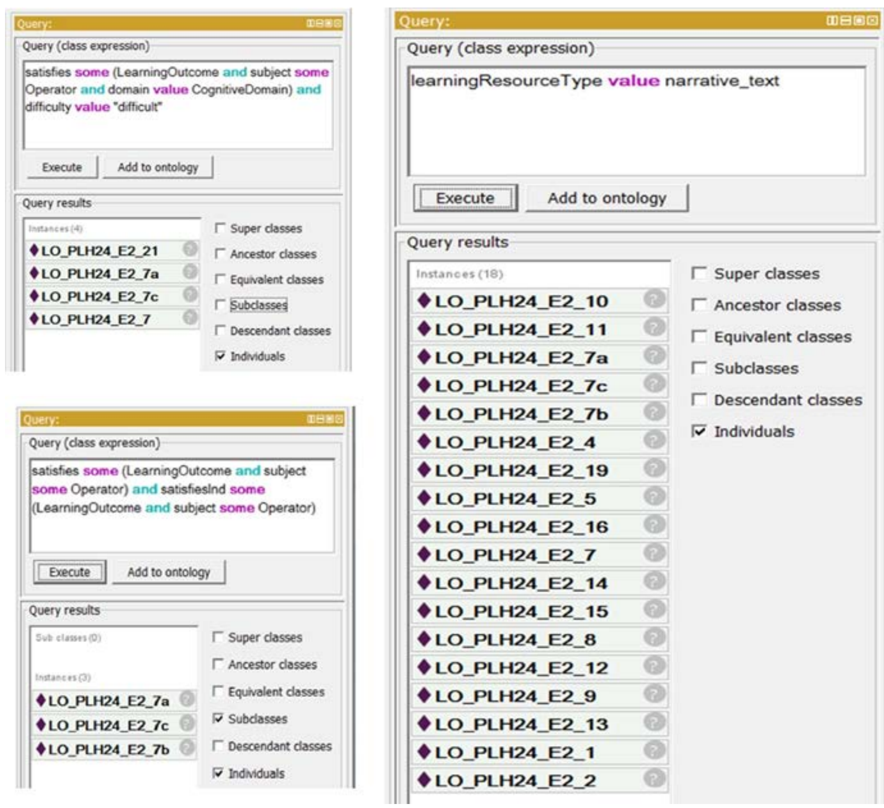

Figure 5. Results retrieved after evaluating queries 1, 2 & 3 of Table 5

Table 5. Some Example Queries in Manchester OWL Syntax

#	Query
1	satisfies some (LearningOutcome and subject some Operator and domain value CognitiveDomain) and difficulty value "difficult"
2	learningResourceType value narrative_text
3	satisfies some (LearningOutcome and subject some Operator) and satisfiesInd some (LearningOutcome and subject some Operator)

Of course similar requests can be made for different domain subjects, different levels of knowledge and any kind of relationship modeled in the ontology as a property. These semantic queries are actually examples of competency questions for the proposed ontology. Competency questions are a commonly used technique for evaluating ontology based formalisms (Grüninger & Fox, 1995).

Conclusions

In this work we propose an educational metadata profile (EMP) that can be used to characterize digital educational resources indented for use in distance learning courses. The proposed EMP either modifies some of the existing elements of the IEEE LOM schema, in order to reflect important aspects of the educational material, such as its technical data type or learning resource type, or augments it with new elements, such as the expected learning outcomes of a course, thus expressing concepts which are essential in the learning process, especially if it is supported by e-learning systems. EMP is rich enough to effectively describe both educational and technical aspects of an educational resource and especially a LO, but not exceedingly analytic so as to become difficult in use. We model our EMP as an ontology, in order to capture and process the semantic relationships among learning resources. Ontologies come with many applications in the field of education and their usage for representing the structure of an educational resource could lead to the development of advanced, intelligent applications. The ontological representation of our EMP was accomplished by "translating" its structural elements to classes, properties and instances in the ontology.

We applied EMP to characterize real LOs, designed for the HOU distance course on Object Oriented Programming (Java). With the aid of Protégé we show that the ontological representation of our EMP could lead to the discovery of LOs according to a specific concept of the knowledge domain – even if this is not explicitly stated - or according to very specific characteristics of a course, i.e., learning outcomes. This knowledge-based discovery could facilitate the process of designing a course and reusing LOs in various educational contexts, as described in Pierrakeas, Solomou and Kameas (2012).

The proposed EMP is already being used for the characterization of the HOU educational resources (including textbooks, digital supplementary material, presentations, theses, etc.). In addition, it one of the core elements of the new LO-based approach adopted by HOU for the delivery of educational content to its students. Currently HOU is developing a Personalized Learning System that will use this EMP and its ontological representation in order to offer advanced services to distance learners. At the same time, HOU is investing in the development of knowledge domain and learning outcome ontologies for its courses; these will be gradually integrated to the EMP ontologies that are being developed for the corresponding LOs. Then, we plan to utilize an instructional design methodology to organize e-learning courses using LOs that are described by EMP, thus providing advanced services for learners that could support efficiently the handling and dissemination of educational material according to their specific needs.

References

Al-Khalifa, H. S., & Hugh, D. (2006). The Evolution of metadata from standards to semantics in e-learning applications. In *Proceedings of Seventh Conference Hypertext and Hypermedia* (pp. 69-72). New York, NY: ACM. doi:10.1145/1149941.1149956

Berners-Lee, T. (1997). *Web architecture: Metadata*. Retrieved from http://www.w3.org/DesignIssues/Metadata.html

Brace, J. & Nejdl, W. (2004). Ontologies and metadata for e-learning. In S. Staab & R. Studer (Eds.), *Handbook on Ontologies* (pp. 555-573). Heidelberg, Germany: Springer-Verlag.

Castagné, M (2013). *Institutional repository software comparison: DSpace, EPrints, Digital Commons, Islandora and Hydra* (Report No. LIBR 596). University of British Columbia. Retrieved from http://circle.ubc.ca/handle/2429/44812

DCMI (2012). *Dublin core metadata element set, version 1.1.* Retrieved from http://www.dublincore.org/documents/dces/

DCMI Usage Board (2012). *DCMI metadata terms*. Retrieved from http://dublincore.org/documents/dcmi-terms/

Grüninger, M. & Fox, M. S. (1995). Methodology for the design and evaluation of ontologies. In *Proceedings of the Workshop on Basic Ontological Issues in Knowledge Sharing, International Joint Conference on Artificial Intelligence (IJCAI-95)*. Montreal, Canada.

Heery, R., & Patel, M. (2000). Application profiles: Mixing and matching metadata schemas. *Ariadne, 25*. Retrieved from http://www.ariadne.ac.uk/issue25/app-profiles/

Hodgins, W., & Duval, E. (2002). *Draft standard for learning object metadata*. New York: Institute of Electrical and Electronics Engineers. Retrieved from http://129.115.100.158/txlor/docs/IEEE_LOM_1484_12_1_v1_Final_Draft.pdf

Horridge, M., & Patel-Schneider, P. (2008, April). *Manchester syntax for OWL 1.1*. Paper presented at the 4th International Workshop OWL: Experiences and Directions, Karlsruhe, Germany.

Freed, N., Klensin, O. J., & Hansen, T. (2013). Media type specifications and registration procedures. Retrieved from http://tools.ietf.org/html/rfc6838

Kalou, A., Solomou, G., Pierrakeas, C., & Kameas (2012). An Ontology model for building, classifying and using learning outcomes. In *Proceedings of IEEE 12th International Conference on Advanced Learning Technologies (ICALT)* (pp. 61-65). doi:10.1109/ICALT.2012.45

Kouneli, A., Solomou, G., Pierrakeas, C., & Kameas, A. (2012). Modeling the knowledge domain of the Java Programming language as an ontology. In E. Popescu (Eds.), *Advances in Web-Based Learning-ICWL 2012* (pp. 152-159). Springer Berlin Heidelberg. doi:10.1007/978-3-642-33642-3_16

Lockwood, F. (2013). *Open and distance learning today*. Oxon, OX: Routledge.

Nejdl, W., Wolf, B., Qu, C., Decker, S., Stintek, M., Naeve, A., Nilsson, M., Palmer, M., & Risch, T. (2002). EDUTELLA: A P2P networking infrastructure based on RDF. In *Proceedings of the 11th international conference on World Wide Web* (pp. 604-615). New York, NY: ACM.

Nikolopoulos, G., Solomou, G., Pierrakeas, C., & Kameas, A. (2012). Modeling the characteristics of a learning object for use within e-learning applications. In *Proceedings of the Fifth Balkan Conference in Informatics* (pp. 112-117). New York, NY: ACM.

Nilsson, M., Palmér, M., & Brase, J. (2003, November). *The LOM RDF binding - Principles and implementation*. Paper presented at the 3rd Annual Ariadne Conference, Katholieke Universiteit Leuven, Belgium.

Nilsson, M. (2001). IMS Metadata RDF binding guide. Retrieved from http://kmr.nada.kth.se/static/ims/md-ims.html

Noy, N., & McGuiness, D. (2001). *Ontology development 101: A Guide to creating your first ontology* (Technical Report KSL-01-05). Retrieved from http://liris.cnrs.fr/~amille/enseignements/Ecole_Centrale/What%20is%20an%20ontology%20and%20why%20we%20need%20it.htm

Pierrakeas, C., Solomou, G., & Kameas, A. (2012, October). *An Ontological approach in learning programming languages*. Paper presented at the 16th Panhellenic Conference on Informatics. Piraeus, Greece.

RDF Working Group (2014). *Resource Description Framework (RDF)* [Webpage]. Retrieved from http://www.w3.org/RDF/

Schreurs, J., & Al-Zoubi, A. Y. (2007). Converting content to reusable learning objects adaptable to user preferences and infrastructure. In T. Bastiaens & S. Carliner (Eds.), *Proceedings of World Conference on E-Learning in Corporate, Government, Healthcare, and Higher Education 2007* (pp. 6537-6544). Chesapeake, VA: AACE.

Smith, N., Van Coillie, M., & Duval, E. (2006). Guidelines and support for building application profiles in e-learning. C*EN/ISSS WS/LT Learning Technologies*.

Wray R. E., Lisse S. A., & Beard J. T. (2004). Ontology infrastructure for execution–oriented autonomous agents. *Robotics and Autonomous Systems, 49*(1), 113-122.

Yang, J. T. D., Chen, W. C., Tsai, C. Y., & Chao, M. S. (2005). An Ontological model for SCORM-Compliant authoring tools. *Journal of Information Science and Engineering*, *21*(5), 891-909.

OWL Working Group (2014). *Web Ontology Language (OWL)* [Webpage]. Retrieved from http://www.w3.org/2001/sw/wiki/OWL