MAPPING IEEE LOM TO WSML: AN ONTOLOGY OF LEARNING OBJECTS

Salvador Sánchez-Alonso, Miguel A. Sicilia and Maria Pareja

Information Engineering Research Unit, University of Alcalá, Spain
{salvador.sanchez,msicilia}@uah.es, maria.pareja@alu.uah.es

ABSTRACT

This paper introduces a mapping of the standard for learning object metadata IEEE LOM in the ontology language Web Service Modelling Language (WSML). The objective is both to provide a basis for translating existing IEEE LOM metadata records to WSML, and to serve as a basic learning object ontology from which richer ontological representations can be devised. The mapping provided is not solely aimed at changing form a language to another, but also to provide improved machine-oriented semantics for metadata records whenever possible.

KEYWORDS

Semantic Web Services, Ontologies, Learning Objects, IEEE LOM.

1. INTRODUCTION

The evolution of Web-based learning has fostered the search for methods and technologies that enable a degree of reuse of learning contents and learning activity designs. Such attempt is intended to allow for the reuse of quality resources and the development of automated resource-search tools, and it may eventually reduce the costs of devising learning activities. The concept of learning object [1] is the cornerstone in this new paradigm for instructional design of Web-based learning that emphasizes reuse as a quality characteristic of learning contents and activities. In practical terms, a learning object is a piece of Web content of arbitrary type and structure that is described by a metadata record that provides information about the object and its prospective educational usages. Learning object metadata thus becomes the key to reuse.

Learning object repositories are systems that provide access to collections of learning objects. The major reason for their existence is the fact that online web search engines are not effective (i.e. they return too many results) when searching for educational contents. Instructional materials like policy guides, assignments, simulations, websites, tutorials, matrices and other kinds of educational materials are easier to find from within a controlled collection. The main functionality of a learning object repository is that of searching for learning objects. This search can either be human-oriented (often implemented as interactive search/browsing interfaces) or machine-oriented (implemented as software interfaces providing search services that can be consulted by software agents, e.g. through Web services). The mission of the LUISA project (http://www.luisa-project.eu) is that of exploiting the advantages of a Semantic Web Service Architecture to make richer and more flexible the processes of query and specification of learning needs in the context of Learning Management Systems and Learning Object Repositories.

The LUISA approach to reusability is that of providing formal metadata expressed in terms of ontologies (i.e. expressed in formal ontology languages). The existence of such medatata, combined with the capabilities of a semantics-enabled service-oriented architecture, would eventually foster the discovery of learning object providers based on such formal expression of metadata. The Web Service Modelling Ontology (WSMO) [2] and the Internet Reasoning Service - IRS [3] are among the most important initiatives related to modelling and executing Semantic Web Services concerns. The LUISA project aims at integrating both approaches as part of an interoperable and decoupled reference Service Oriented Architecture. The real challenge for the architectural design of LUISA was to offer an expressive interface while assuring openness and autonomy at a syntactic and semantic level, given that clients need to interact with a number of heterogeneous services.

Semantic repositories of LO are justifiable as an improvement over existing representational mechanisms. This entails a number of principles for design that have resulted in essential commitments for the architecture of the LUISA Learning object metadata repository (LOMR). The aim is to provide an open source reference Service-Oriented Architecture which can be implemented for the flexible learning object discovery, selection, negotiation and composition through the use of Semantic Web Services. In the context of LUISA, this entails that the common practices of storing learning object metadata must be bridged to the ontology-based approach of Semantic Web services. The LUISA LOMR is responsible, among others, for the storage of ontologies that describe the learning resources. Storing ontologies is required for end user applications to expose parts of the ontology(-es) that are used to formulate learning goals of a diverse type. The interfaces that can be provided for that purpose are of two kinds: generic interfaces for querying ontology structure and domain or pedagogy-specific interfaces. This functional view of the LUISA LOMR is depicted in Figure 1.

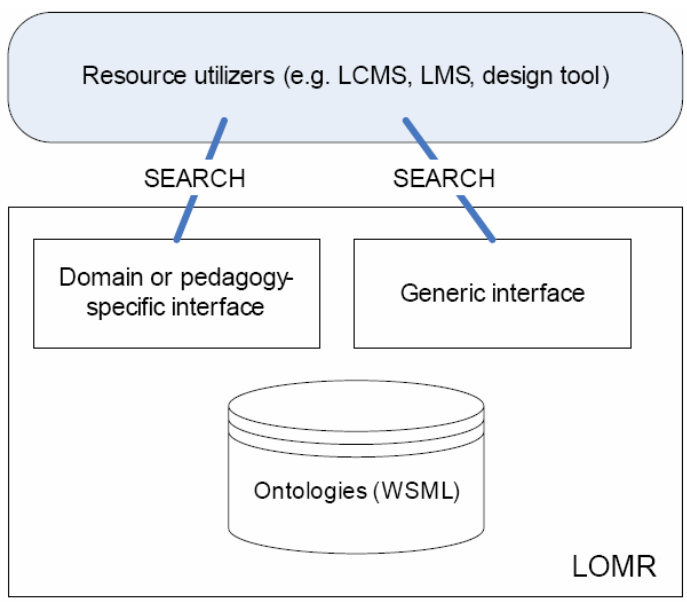

Figure 1. LUISA LOMR (Learning Object Metadata Repository) architecture

From the need to provide a basis for translating existing IEEE LOM metadata records expressed in XML to WSML, and secondly, from the need to serve as a basic learning object ontology from which richer ontological representations can be devised, a mapping from IEEE LOM [4] to

WSML v0.21 [5] was found to be necessary. This paper describes such mapping, formally engineered as an ontology called LOM2WSML.

The mapping provided is not solely aimed at changing form a language to another, but also to provide, whenever possible, improved machine semantics for metadata records. WSML has been preferred because, regardless other serializations of LOM (e.g. XML, RDF), it provides the semantic capabilities outlined in the introduction to this paper, as it is a full ontology language.

The guiding principles of the LOM to WSML mapping are the following:

- The mapping must be compliant with LOM in the sense that any LOM statement must be translatable to the WSML ontology.
- The mapping must retain LOM terminology to the extent possible.
- The mapping must be oriented to providing the richer computational semantics possible.

These principles entail that the mapping described here is restricted by the LOM conceptual schema, but it is not restricted in the way data is represented. For example, some values represented in LOM as character strings are represented in WSML through instances of a given relationship, as this is the approach chosen in the mapping.

To better understand the importance of including semantic information in the e-learning arena, let us focus on an example about the practical use of learning objects. When a user (a person or an application) needs to retrieve educational resources from a repository in the form of learning objects, probably due to prior pedagogical needs, the precise meaning of each particular learning object-related terms are extremely important. Retrieving the most adequate learning objects is rooted in this fact. If the user's needs lead to retrieving learning objects on "Java programming", only those objects that fulfil the conditions of the "contract" with the user should be selected. However, different users might need different levels of training because of their previous knowledge on the topic or due to their personal background. Thus, users with no previous knowledge on programming will not have the same needs as a senior C++ programmer investigating about the features that differ between Java and C++. This is the reason why it is so important to add semantic information to learning object metadata records.

If search criteria are to be based, as they are in this example, on the users' competencies, the specific competency addressed by each learning object, as well as the competency level measure (in a shared scale) that learners obtain as an outcome of the learning process, should be explicitly stated. But, do all the members of the community of contributors to the learning object repository think the same about what a competency is? Are there different points of view, different conceptualizations of this and other terms? To be ready to implement advanced search capabilities, a shared, unique and unambiguous definition must be agreed.

For this, different approaches to the representation of knowledge exist, such as controlled vocabularies, glossaries, thesauri, hierarchies or taxonomies. All provide us with different ways to define terms of a domain, as well as the relationships between those terms in a very expressive way. However, they do not provide extra semantic information from which automated reasoning can be inferred, as the internal representation of the knowledge is not formalized in a logics-based language. Non-semantic information is not sufficient for the semantic interoperability between different schemes: to reach interoperability between those schemes using the information, it is necessary to define the differences between different meanings of a term, in a clear and unique way. Only through the use of ontologies, which allow to built knowledge upon logics-based formalisms, terms in the domain (in the case of this example *competency*, *CompetencyMeasurement*, and all the rest of terms involved in the selection processes) can be semantically defined.

The rest of this paper is structured as follows. Section 2 sketches the ideas around the term "learning object", the central concept of the ontology and thus a key element to describe before the mapping can be understood. Sections 3 and 4 provide some detail on the implementation of the mapping itself. Finally, conclusions and further work are to be found in Section 5.

2. MODELLING LEARNING OBJECTS

IEEE LOM is a property–oriented specification. As such, it defines a number of properties that can be used to describe instances of an implicit class of entities that requires a representation in ontological terms. This leads to the definition of a `LearningObject` concept representing such class of entities[1]. This apparently simple decision has some important consequences:

- Learning objects require some identity condition. From the viewpoint of the IEEE LOM specification, this is a matter of using unique identifiers, so that it seems unproblematic. However, in practical settings this means that when importing metadata, information must be attached to the right object considering the identifiers.

- Metadata then becomes simply the property values associated to a given instance of `LearningObject`.

It should be noted that these consequences do not preclude having different and even conflicting metadata regarding the same object –as for example IMS LD and QTI, see [6]. To deal with these conflicts, solutions can be as varied as ensuring that conflicting metadata pieces are stored in different repositories or including additional mechanisms for dealing with such inconsistencies, e.g. creating some additional ontological definitions for different contexts in which conflicting metadata statements can be accommodated. Subclasses of the `LearningObject` concept can be used to further describe relevant subsets of the collection of learning objects as described in [7]. Figure 2 shows the properties (called *attributes* in the WSMO Studio editor which was the tool used for its engineering) of the `LearningObject` concept, the central term of the ontology. In this figure, the left hand side part of the editor shows the terms in the ontology (in green), while the right hand side panel shows the properties of the term selected (`LearningObject` in this particular case).

As a help to facilitate the mapping, but also as a powerful tool to attain interoperability, a complete general knowledge base and commonsense reasoning engine has been used. OpenCyc, the open source version of the Cyc technology [8], contains hundreds of thousands of terms, along with millions of assertions relating the terms, forming an upper ontology whose domain is all of human reality. In the ontology created as a result of mapping IEEE LOM to WSML, the prefix "oc" is used to represent concepts defined in the knowledge base OpenCyc. However, the names of those terms in OpenCyc do not have any prefix (e.g. the term "Person" in OpenCyc will be referenced in this ontology as `ocPerson`, meaning "The concept person as defined in Opencyc").

As we have just mentioned, metadata becomes in this mapping the property values associated to a given instance of `LearningObject`. Some metadata elements, however, are dually mapped. A *basic mapping* represents the information in WSML as it is represented in LOM, by making use of the datatype correspondences in Table 1. Remarkably, a second mapping called *semantics-oriented mapping* makes use of terms in Opencyc by linking concepts in the WMSL

[1] In what follows, the concepts of the ontology will be represented in `courier` font.

ontology of learning objects to concepts defined in Opencyc, the aim being to foster the semantic interoperability of IEEE LOM conformant metadata descriptions.

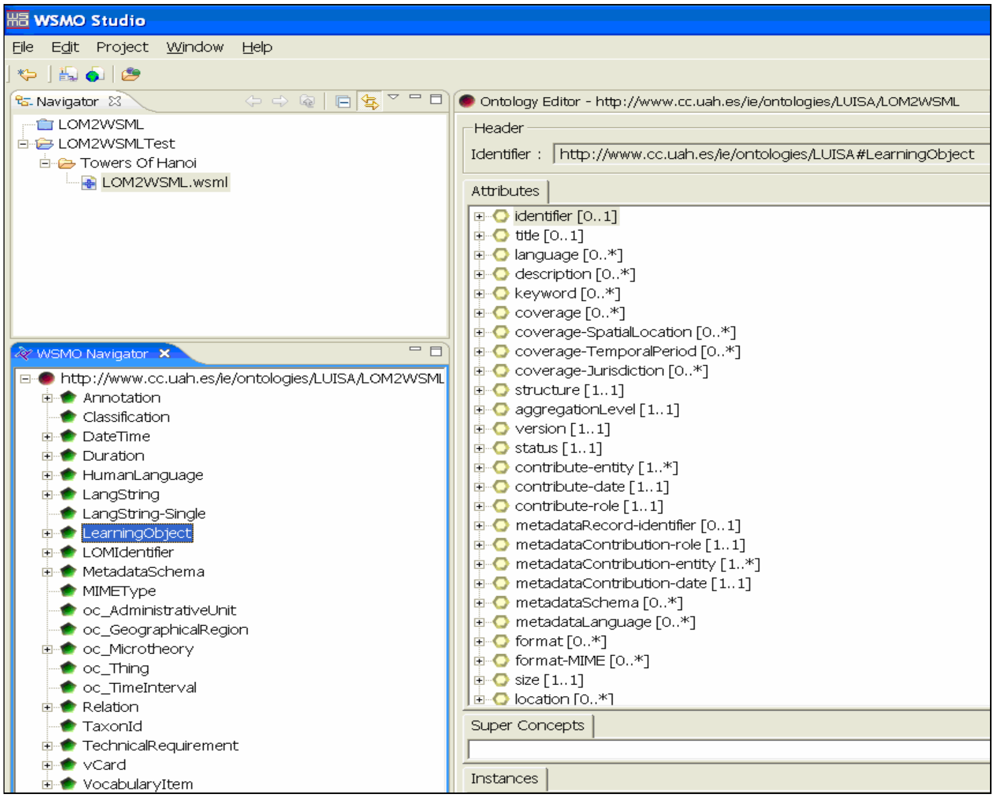

Figure 2. Editing the WSML ontology in WSMO Studio

3. MAPPING IEEE-LOM DATA TYPES

As a previous work, mapping IEEE LOM data types to WSML was not only necessary but of utmost importance. Data types are the basis of this mapping, as all the concepts in the ontology will include attributes or properties which will need to be described in terms of a datatype. Table 1 summarizes all the work in this area, but following discussion will go deeper into the details of each data type.

Table 1. Mapping of IEEE LOM data types to WSML.

IEEE LOM datatype	**WSML element**
CharacterString	WSML `_string`
LangString	Instance of the `langString` WSML concept that has a number of associated `langString-single` instances. Each actual string is connected to an instance of the termconcept `ocHumanLanguage`.
DateTime	Instance of the `dateTime` WSML concept in which both a `_dateTime` and a `LangString` value describing the date are provided.

Duration	Instance of the duration WSML concept in which both a _duration value and a LangString value describing the duration are provided.
Vocabulary item	Those LOM elements explicitly considered in the ontology, will be represented as concept instances of the corresponding sub-concept in the vocabularyItem hierarchy. For vocabularies other than those directly represented in the WSML ontology, the mapping is to an instance of the vocabularyItem concept.

The LOM standard states that "The LOMv1.0 Base Schema does not specify encodings for CharacterString". Thus, a straightforward mapping is that of mapping CharacterString to the built-in WSML datatype _string.

LOM LangStrings "may include multiple semantically equivalent character strings, such as translations or alternative descriptions". Thus, LangStrings are possibly multiple pairs in the form (language, CharacterString) with the same meaning.

Further, languages are represented through an additional concept ocHumanLanguage, defined in OpenCyc as "a specialization of Language, representing a language that is used by human communities for communication. This collection differs from NaturalLanguage in that a HumanLanguage may be purposefully created, while NaturalLanguage evolve without a purposeful creation process".

LOM DateTime values are defined as "a point in time with accuracy at least as small as one second". This fits the definition of the WSML type _dateTime. In a similar way, LOM Duration values, defined as "an interval in time with accuracy at least as small as one second", have been mapped to the WSML type _duration.

Vocabulary items in LOM are defined as pairs of CharacterStrings (isource,value), with source being typically a URI that identifies the vocabulary. A straightforward mapping may be that of creating a vocabularyItem concept and map the CharacterStrings to string values through the vocabularyItem properties itemValue and itemSource. However, vocabularies are often better represented as specific concepts in WSML, and the source can be expressed as a non-functional property. For compatibility with vocabularies not considered in LOM, this latter option is recommended.

4. MAPPING IEEE LOM CATEGORIES

This section shows how some elements of metadata were mapped to WSML. Reporting the full mapping is impossible due to the length restrictions but also uninteresting for most readers as the key ideas can be revealed without the need of an exhaustive description of all the categories of metadata in LOM. In what follows, the most relevant categories of metadata, from the point of view of the interest of the decisions taken to map it, will be discussed.

4.1. LOM Category "1.4. Description"

Descriptions are texts that can be used to give descriptive statements about the learning object. These texts are intended for human consumption, and consequently do not require any additional semantics. This has been mapped to WSML through an attribute called *description* of the LearningObject concept, whose range is of langString type.

Other textual and descriptive categories in LOM, such as LOM category *2.1.Version* (information referring the edition of the learning object) or *4.6.Other platform requirements* (Descriptive information about other software and hardware requirements not covered by other previous data elements), are represented in the same manner.

4.2. LOM Category "1.6. Coverage"

Coverage is represented in LOM as strings describing "the time, culture, geography or region to which the learning object applies". The mapping provided retains the possibility of storing simple strings, but allows also for the improved semantic representation of the elements explicitly mentioned in LOM. This latter possibility is rooted in an affirmation in the LOM standard that explicitly states that "Coverage will typically include spatial location (a place name or geographic coordinates), temporal period (a period label, date, or date range) or jurisdiction (such as a named administrative entity)".

The mappings provided for this specific category are detailed below (see Table 2). However, these mappings are not exclusive to other additional mappings with more details on the coverage of the learning object.

Table 2. Coverage dual mapping

LOM element	WSML elements
coverage	*Basic mapping:* Attribute `coverage` of the `LearningObject` concept, with `langString` as range. *Semantics-oriented mapping:* Attribute `semanticCoverage` of the `coverage` concept, which in turn includes 3 properties: - For spatial locations: attribute `coverage-SpatialLocation` with `ocGeographicalRegion` as range. - For Temporal periods: attribute `coverage-TemporalPeriod` with `ocTimeInterval` as range. - For jurisdictions: attribute `coverage-Jurisdiction` with `ocAdministrativeUnit` as range.

In Table 2 several terms from Opencyc were used as part of the semantics-oriented mapping. The definitions of those Opencyc terms are the following:

- `ocGeographicalRegion`: "A tangible spatial region that includes some piece of the surface of a planet (usually PlanetEarth), and may be represented on a map of the planet. This includes purely topographical regions like mountains and underwater spaces, places defined by demographics (e.g. language areas) and territory otherwise demarcated"

- `ocTimeInterval`: "An intangible temporal thing that is characterized fully by its temporal extent. In this way, time intervals differ from Situations such as Events. For example, the year 1969 C.E. is a `TimeInterval`; although many interesting things happened during that year, the year itself is completely defined by its temporal extent."

- `ocAdministrativeUnit`: Each instance is a unit with administrative responsibilities.

4.3. LOM Category "2.3. Contribute"

In LOM, a contribution is modelled as information on the 2.3 category *contribute*. This term refers to those entities (i.e., people, organizations) that have contributed to the state of this learning object during its life cycle (e.g., creation, edits, publication). Contributions should be considered in a very broad sense here, as all actions that affect the state of the learning object.

Mapping contributions to WSML was similar to the previously explained information on coverage (see Table 3). As in the previous section, some Opencyc terms were used in the semantics-oriented mapping, whose definitions follow:

- `ocOrganization`: "Each instance of Organization is a group whose group-members are instances of `ocIntelligentAgent`. In each instance of `ocOrganization`, certain relationships and obligations exist between the members of the organization, or between the organization and its members. Instances of `ocOrganization` include both informal and legally constituted organizations, each being capable of undertaking projects, entering into agreements, owning property, and doing other tasks characteristic of agents."

- `ocRole`: "A specialization of `ocObjectPredicate` whose instances relate situations to individuals that are involved in them in various ways."

Table 3. Contribute dual mapping

LOM element	WSML element
contribute	*Basic mapping:* Attribute `contributeLifeCycle` of the `LearningObject` concept, an instance of `contribute`, term which in turn includes the following properties: - Role: property `role` with `lomRolesVocabularyItem` as range. - Entity: attribute `entity` with `vCard` as range. - Date: attribute `date` with `dateTime` as range. *Semantics-oriented mapping*: Attribute `semanticContributeLifeCycle` of the `LearningObject` concept, including 3 properties: - Role: property `role` with `ocRole` as range. - Entity: attribute `entity` with `ocOrganization` as range. - Date: atribute `date` with `dateTime` as range.

4.4. LOM Category "5.6.Context"

This category includes information about the principal environment within which the learning and use of this learning object is intended to take place. It has been modelled through two properties (`context` / `semanticContext`) of the `LearningObject` concept. The range of these properties is `lomContextVocabularyItem` (for the basic mapping), a concept that belongs to the hierarchy of the concept `VocabularyItem`, and `ocMicrotheory` (for the semantics-oriented mapping).

Table 4. Mapping IEEE LOM Context information

LOM element	WSML element
context	*Basic mapping*: Attribute `context` of the `LearningObject` concept, with `lomContextVocabularyItem` as range. *Semantics-oriented mapping*: Attribute `semanticContext` with `ocMicrotheory` as range.

In the mapping of this information, the Opencyc term `ocMicrotheory` has been used in the semantics-oriented mapping. An `ocMicrotheory`: an atemporal abstract informational thing that represents a context in Cyc. Each microtheory serves to group a set of assertions together that share some common assumptions; the assertions in a microtheory constitute the content of that microtheory. In the Cyc knowledge base, each assertion must be explicitly stated to be true in at least one microtheory.

4.5. LOM Category "9. Classification"

The information in this category serves to describe where this learning object falls within a particular classification system. As a classification is understood in IEEE LOM as a whole containing several sub-elements, this mapping has shapes classifications as instances of an entity called "Classification". Following the semantic-oriented mappings previously described in categories such as *content* and *contribute*, we have explored the possibility of providing enhanced semantic capabilities to learning objects through the modelling of semantic-oriented competence and discipline classification points.

Regarding competency classifications, the semantic-oriented data have been modelled like this:

- *Purpose*: competency (implicit when using the learningObject concept attribute semanticCompetencyClassification [0..*])
- *TaxonPath*: the property semanticCompetencyClassification of the concept learningObject has `semanticCompetencyClassification` as range. This latter concept has properties which allow linking the LOM2WSML ontology to specific competence ontologies. In the LUISA community, a specific ontology called GCO (which stands for General Competency Ontology) has been engineered. The existence of GCO forces somewhat the way in which this first version of LOM2WSML deals with competencies. In this particular case:
    - Two sub-concepts of `semanticCompetencyClassification` (namely `semanticCompetencyElementClassification` and `semanticCompetencyElementDefinitionClassification` are used to link to the respective concepts in GCO (namely `CompetencyElement` and `CompetencyElementClassification`).
    - *Description* and *keyword*: properties of `semanticCompetencyClassification` whose range is langString.

5. CONCLUSIONS

This paper describes the overall structure and design guidelines of a mapping from the IEEE LOM to the ontology language WSML. The mapping reported, which is version 0.1 of the LOM2WSML ontology, will be used as a means to store learning object metadata records in a formal ontology-based format, allowing components and Web services of the LUISA project to properly manage learning object descriptions.

Further work should validate the dual description of most components (basic and semantics oriented), investigating whether other categories of metadata can be described in this form. As a test bench, the ontology should now be populated with instances from those partners of the LUISA consortium whose contribution to the project consist in providing educational materials as use cases (such as EADS[2] and the Henri Pointcaré University[3]).

We also need to link (and eventually merge both ontologies) with the general competencies ontology engineered by other team in the LUISA project, especially insisting in referencing OpenCyc elements and using more Opencyc properties.

ACKNOWLEDGEMENTS

The work reported in this paper is funded by the European Commission under the project LUISA, reference FP6-027149. It is also part of the Project SERE-OC (Resource selection in e-learning based on competency ontologies), code CCG06-UAH/TIC-0606 funded by UAH and the CAM (Comunidad Autónoma de Madrid).

REFERENCES

[1] Polsani, P. R. (2003). "Use and Abuse of Reusable Learning Objects". Journal of Digital information, Vol. 3, No. 4.

[2] Roman, D., Keller, U., Lausen, H., de Bruijn, J., Lara, R., Stollberg, M., Polleres, A., Feier, C., Bussler, C., and Fensel, D. (2005). "Web Service Modeling Ontology". *Applied Ontology*, Vol. 1, No. 1.

[3] Domingue, J., Cabral, L., Hakimpour, F., Sell, D., and Motta, E. (2004) "IRSIII: A platform and infrastructure for creating WSMO-based Semantic Web Services". In Bussler, C. et al (Eds.) *Proceedings of the WIW 2004 Workshop on WSMO Implementations*, CEUR. Online at: http://sunsite.informatik.rwth-aachen.de/Publications/CEUR-WS/Vol-113/

[4] IEEE Learning Technology Standards Committee (2002) *Draft standard for Learning Object Metadata*. Last retrieved from http://ltsc.ieee.org/wg12/files/ 10th January 2006.

[5] De Bruijn, J. (ed.) (2005) *The Web Service Modeling Language WSML, Final Draft*. Last retrieved from http://www.wsmo.org/TR/d16/d16.1/v0.21/ 10th January 2006.

[6] Downes, S. (2004). "Resource Profiles". *Journal of Interactive Media in Education*, Vol. 2004, No. 5.

[7] Sicilia, M.A., García, E., Sánchez-Alonso, S., and Rodríguez, E. (2004). "Describing learning object types in ontological structures: towards specialized pedagogical selection". In *Proceedings of ED-MEDIA 2004 - World conference on educational multimedia, hypermedia and telecommunications*, pp. 2093-2097.

[8] Lenat, D. B. (1995). "Cyc: A Large-Scale Investment in Knowledge Infrastructure". *Communications of the ACM*, Vol. 38, No. 11, pp. 33-38.

---

[2] http://www.eads.com/
[3] http://www.uhp-nancy.fr/