

Domain Ontologies Integration into the Learning Objects Annotation Process

Zarraonandía, T.¹, Dodero J. M.¹, Díaz P.¹, Sarasa A.²

¹ Universidad Carlos III de Madrid
Laboratorio de Sistemas Interactivos
dei@inf.uc3m.es

² Universidad Complutense de Madrid
asarasa@sip.ucm.es

Abstract. This paper describes an extension of the LOM metadata elements used in the IMS learning object specifications, to incorporate new labels for the semantic classification of non-educational information to the set of LOM metadata. These labels have been obtained from specialized vocabularies represented by means of domain-specific ontologies described using RDF and transformed using a LOM to RDF binding. They are included as LOM standard metadata under the classification subcategory. In order to test its viability, the proposed extension has been implemented as an additional functionality of a learning object editing tool developed as part of the MD2 project.

1 Introduction

In the e-learning context we can distinguish two different areas where the ICTs could be applied: one of them is related with the educational process, while the main focus of the other one is the didactic material development. The convergence of the two areas has been lately revealed by the integration of concepts provided by educational model languages [13] with the packaging of e-learning [5,6] industry contents, which is reflected on IMS Learning Design Specification [7]. This work is focused on the second one of the topics: didactic material development.

Due to the domain-specific characteristics of the educational material, it is necessary to annotate them with domain-specific information, in order to improve their management and exploitation. This paper describes an authoring and annotation tool for learning objects, which has been developed to support the creation of extended didactic material in the framework provided by the MD2 project. Firstly, the context of the project is exposed. Secondly, the problematic of didactic material annotation with metadata extracted from domain-specific ontologies is drawn. Next, an edition and annotation tool and a practical example of LOM extension are presented. Finally, some conclusions are given and future works are described.

2 The Context: Didactic Material Creation

This work is an approach to solve one of the goals of the MD2 project. The aim of the project is to provide solutions to some of the problems related to the generation of learning material, which can be resumed in the following:

- The development of a method for the collaborative generation of learning contents, offering a framework for cooperative knowledge production with a view to improve efficiency and reduce conflicts and coordination issues.
- The extension of current learning object standards to incorporate concepts of instructional hypermedia such as learning links, and to achieve metadata cohesion using accepted and shared concepts (ontologies).

- To elaborate an evaluation framework for a priori testing of the usability and utility of educational products, which includes a method and a number of criteria, parameters and metrics, concerning the educative and interactive quality of applications.

These theoretical endeavors will be practically tested into a platform that will be developed with this purpose. The platform architecture is shown in Fig. 1. Next, a brief description of its modules and their functionalities is presented.

- *Edition + Annotation*: these modules provide basic functionality for edition and annotation of learning objects. Both of them are integrated on a unique authoring tool, similar to Reload¹ or Aloha², but enhanced with some capability extensions to consider traversal aspects served by other modules as ontology import and collaboration support.
- *D-Ontology Import*: this module allows the extension of the label set used for the annotation of learning objects by using ontologies described with RDF(S).
- *Assessment*: This module provides the means to perform a priori tests of the quality of the in-development learning objects [15].
- *Collaboration*: this module supports the collaboration mechanisms during the development of the learning objects, especially during the annotation and evaluation processes. It serves as the base mechanism for the collaborative generation method mentioned among the project goals.
- *Performance Analysis*: it carries out an analysis of the behavior of learning objects' users during the didactic process, in order to evaluate their performance in a given learning context. It takes into account the user model and the run-time engine provided by the LCMS (Learning Content Management System) where objects are executed. The results will revert into annotations to the learning objects regarding the performance of the users.

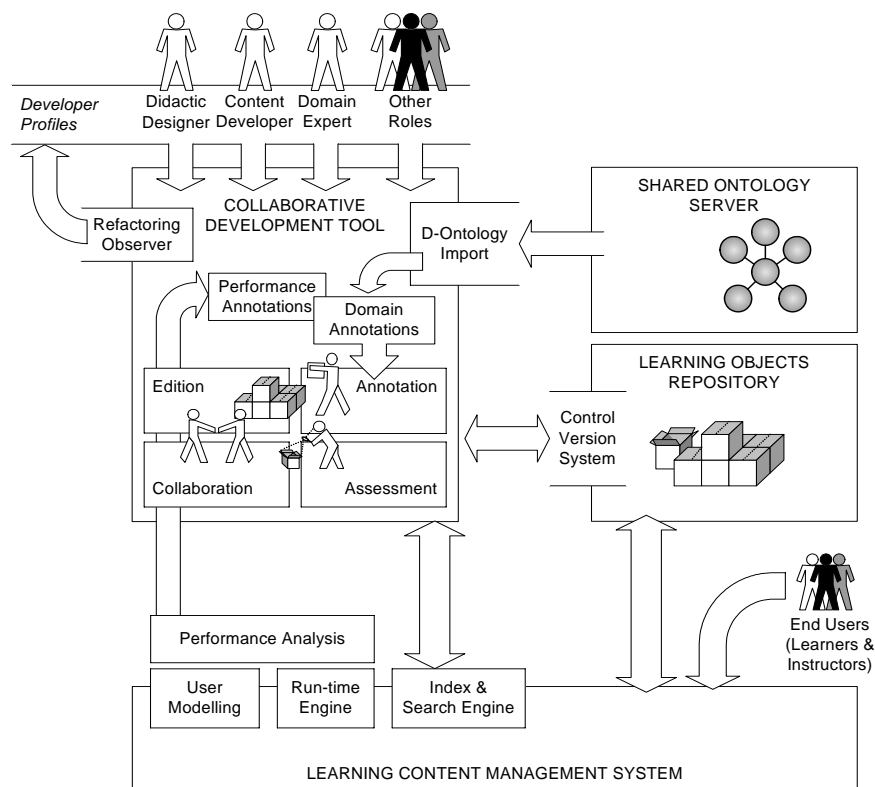


Fig. 1. General architecture of MD2 platform

- *Refactoring Observer*: this is an asynchronous system which, taking as input the values and annotations generated by the previous module, can generate proposals for re-designing the learning objects (i.e., refinement of their objectives and/or requisites, recommendations for

¹ <http://www.reload.ac.uk/>

² <http://aloha.netera.ca/>

new examples, splitting of merging contents, etc.) in order to obtain didactic improvements to adapt them to different learning contexts.

The rest of the components of the figure are external subsystems (LCMS, learning objects repositories and shared ontology servers) that should be adapted to the development platform in order to take advantage of their functionalities. To accomplish this, a Web Service architecture is being developed; a detailed description can be found in [10].

3 Learning Objects Annotation

A major task during the process of creation of a learning object is to generate annotations according to IEEE Learning Object Meta-Data (LOM) standards [6]. These specifications distinguish different metadata categories (i.e. general, technical, educative, etc.) to describe a learning object. Among them, the *classification* category is used to accommodate the annotations related to a particular classification scheme (e.g. the Dewey's decimal classification system [1], or the generalist taxonomies of the Open Directory Project [4]). In our work, the elements *taxon* and *taxonpath* from *classification* category are chosen for cataloguing resources with domain-specific information. It must be noted that this is a limited solution [3], since current LOM specification is designed for the use of simple taxonomies, but not for full-fledged ontologies that can be represented by description logics, as those proposed in OWL [8].

An ontology is as a formal, explicit specification of a shared conceptualization [16]. In this work we focus on the shared character of ontologies. Considering that the purpose of initiatives like the Knowledge Sharing Effort [11] is the development of conventions to support the sharing and reuse of knowledge among systems, it seems reasonable to think in ontologies as an appropriate basis to perform the annotation of learning objects. Despite the fact that currently we can only consider their taxonomic character, in order to not constraint the future evolution of annotations towards full-fledged ontologies, the selected languages have been the ones proposed by the W3C for the development of the Semantic Web: RDF(S) and OWL. On a first stage, it has been considered the importation of domain ontologies expressed in RDF(S), taking into account the strong taxonomic character of the annotations imposed by LOM specification. This has been achieved following the proposed recommendations from the LOM to RDF binding [9].

To increase the reusability and the quality description of a learning object, it is necessary to use more specialized metadata than those proposed by the LOM specifications. These are possibly of non-educational character, in order to adapt the objects to domain-specific, multidisciplinary learning contexts. We take advantage of the LOM infrastructure to introduce these metadata extensions. The development has been carried out on three different stages:

- Development of an authoring tool that provides basic edition functionalities.
- Development of a functionality to obtain metadata classification by using arbitrary hierarchies not directly linked with a specific domain.
- Development of a functionality to integrate hierarchical collections of labels based on domain-specific ontologies.

3.1. Learning Objects Edition & Annotation Tool

The learning objects edition & annotation tool is IMS content packaging [5] and metadata [6] standards compliant. The first one describes the structure of a learning object as a zip content package, composed by a manifest file *imsmanifest.xml* that is divided into four different sections (i.e. *organizations*, *metadata*, *resources* and *sub manifests*) and a set of referenced resource files.

A major objective of the tool was to provide the appropriate labels to classify the domain-specific information contained in the learning object, and for which the LOM specifications does not supply concrete metadata. Two different approaches can be undertaken to overcome this issue:

1. To create specific resources associated to the manifest file, by directly creating the taxonomy. In order to achieve this, an option has been implemented that allows the user to create generic taxonomies as deep and complex as needed. These taxonomies can be managed using visual

and interactive hierarchical components. Then the generated taxonomy can be stored in a different XML document, which will be referred to from the manifest file as another resource file.

2. To import specific domain ontologies: another approach to solve the problem is to extend the label set provided by LOM specification to generate new taxonomies. This is done taking *taxon* and *taxonpath* metadata elements as starting point. The objective of these metadata is to catalog the learning object information using specific classification systems. Based on them, LOM offers the possibility to introduce new labels relative to other namespaces, which share the same classification objectives. Our tool has taken advantage of this LOM extension facility to import specific domain ontologies by importing an RDFS description of the domain ontology and transforming it into LOM metadata [9]. The imported labels are internally stored in order to give a personalized label set for domain-specific classification, which can be reused in future annotations. The reverse process is also possible if the user opens a learning object for editing, which contains *taxon* metadata descriptions related to the specific domain ontology. In order to not constraint the semantic interoperability of the object in such situations, the namespaces used on the domain ontology have been maintained.

3.2. Practical Example of LOM Extension

A concrete example is provided, which consists in the annotation of a learning object about music specialties. D-Ontology Import module support will be used to obtain the required labels to annotate resources according to a domain-specific taxonomy, which is represented separately in RDF (underneath the xml namespace http://www.mysite.com/dmoz_music) and stored on the Shared Ontology Server. The reference taxonomy was built inspired from *dmoz.org* Open Directory project categories.

According to [9] an *rdfs:subPropertyOf lom-cls:classification* should be used to perform the adequate annotation in the manifest file, pointing always to *taxon* values in a taxonomy hierarchy which should be an instance of *lom-cls:Taxonomy*, as the example shown at **Fig. 2**. There are three different possibilities depending on the value of *Purpose* subcategory from the *Classification* category:

1. When the value is “Discipline” or “Idea”, use *dc:subject*, which is an *rdfs:subPropertyOf lom-cls:classification*.
2. If the value is part of the LOM restricted vocabulary for the *Purpose* subcategory, use one of the following, which are *rdfs:subPropertyOf lom-cls:classification* *lom-cls:prerequisite*, *lom-cls:educationalObjective*, *lom-cls:accessibilityRestrictions*, *lom-cls:educationalLevel*, *lom-cls:skillLevel*, *lom-cls:securityLevel* or *lom-cls:competency*.
3. Define our own *rdfs:subPropertyOf lom-cls:classification* to use for local purposes.

From these annotation approaches the third one is the less adequate as it could compromise the semantic interoperability of the object, since third party systems are obliged to know the defined set of labels in order to understand domain annotations. If we compare the first and second approaches, the former is more suitable as it is less specific and more easily adaptable to existing classification systems. Taking into account those issues, a common annotation will look like:

```
<dc:subject rdf:resource="http://www.mysite.com/dmoz_music#blues"/>
```

```

<?xml version="1.0" encoding="UTF-8"?>
<rdf:RDF xmlns:rdf="http://www.w3.org/1999/02/22-rdf-syntax-ns#"
  xmlns:rdfs="http://www.w3.org/2000/01/rdf-schema#"
  xmlns:dcterms="http://purl.org/dc/terms/"
  xmlns:dc="http://purl.org/dc/elements/1.1/"
  xmlns:lom-cls="http://ltsc.ieee.org/2002/09/lom-classification#">
  <lom-cls:Taxonomy rdf:ID="DMOZ">
    <rdfs:label>Music Specialties</rdfs:label>
    <lom-cls:taxon>
      <dcterms:DMOZ rdf:ID="blues">
        <rdf:value>blues</rdf:value>
        <rdfs:label>Blues</rdfs:label>
      </dcterms:DMOZ>
      <dcterms:DMOZ rdf:ID="flam">
        <rdf:value>flam</rdf:value>
        <rdfs:label>Flamenco</rdfs:label>
      </dcterms:DMOZ>
      <dcterms:DMOZ rdf:ID="jazz">
        <rdf:value>jazz</rdf:value>
        <rdfs:label>Jazz</rdfs:label>
      <lom-cls:taxon>
        <dcterms:DMOZ rdf:ID="jazz-fus">
          <rdf:value>jazz-fus</rdf:value>
          <rdfs:label>Jazz Fusion</rdfs:label>
        </dcterms:DMOZ>
      </lom-cls:taxon>
    </dcterms:DMOZ>
    <dcterms:DMOZ rdf:ID="eth">
      <rdf:value>eth</rdf:value>
      <rdfs:label>Regional & Ethnic</rdfs:label>
    <lom-cls:taxon>
      <dcterms:DMOZ rdf:ID="eth-lat">
        <rdf:value>eth-lat</rdf:value>
        <rdfs:label>Latin</rdfs:label>
      </dcterms:DMOZ>
    </lom-cls:taxon>
    </dcterms:DMOZ>
    <dcterms:DMOZ rdf:ID="rnb">
      <rdf:value>rnb</rdf:value>
      <rdfs:label>Ryhtm & Blues, Funk, and Soul</rdfs:label>
    </dcterms:DMOZ>
  </lom-cls:taxon>
</lom-cls:Taxonomy>
</rdf:RDF>

```

Fig. 2. A simple RDF taxonomy for learning resources about music specialties.

4 Conclusions and Future Work

We have described our experience during the development of a tool for edition and annotation of learning objects using domain-specific ontologies. We found that LOM metadata elements for the generation of specific domain taxonomies are rather limited. This inconvenient can be overcome by extending LOM. Three possible forms of annotations have been considered to perform this extension. After their evaluation, we suggest using a *dc:subject* tag as we found it is the less constraining with the semantic interoperability of the learning objects. Other possibility of LOM extension, which partially solves the taxonomic classification problem, is to use ontologies to perform the annotation. But since ontologies offer a wider spectrum of knowledge representation possibilities, more complex issues like the integration of description logics into LOM-based annotations are still open.

Upcoming research work is in the context of MD2 project and will involve the study of the extension of the LOM specifications to provide a more complete support to OWL ontologies that can be defined in any of its three sublanguages (*Lite*, *DL* and *Full*) [8].

Apart from that, we found that little attention has been paid to the developers of educational material and the multidisciplinary nature of their tasks (e.g., teaching, didactic advice, usability, or any other

specialized field from the knowledge domain). From that point of view we consider the educational material development as a collaborative and multidisciplinary process [12]. Augmenting the edition tool capabilities to integrate support for the collaborative development of learning objects will conform another line of future research. A learning object quality model is being also developed and integrated within the tool, to allow the definition of appropriate metrics to evaluate and annotate the learning objects during their development process.

Acknowledgements

This work is part of the MD2 project (TIC2003-03654), funded by the Spanish Ministry of Science and Technology.

References

1. M. Dewey, *The Man and the Classification*, Gordon Stevenson & Judith Kramer-Greene (eds.), 1983.
2. J. M. Dodero, M. A. Sicilia and E. García, "A Knowledge Production Protocol for Cooperative Development of Learning Objects", In *Proceedings of the 2nd Workshop on Agent-Supported Cooperative Work, International Conference on Autonomous Agents*, Montreal, 2001.
3. H. B. Enderton, *A Mathematical Introduction to Logic*, Academia Press, 1972.
4. A. Gerhart, "Open Directory Project Search Results and ODP Status", *Search Engine Optimization*, April 5, 2002.
5. IMS Global Learning Consortium, "IMS Content Packaging Information Model", Versión 1.1.3 Final Specification, 2003.
6. IMS Global Learning Consortium, "IMS Learning Resource Meta-data Information Model", Versión 1.2 Final Specification, 2001.
7. IMS Global Learning Consortium, "IMS Learning Design Best Practice and Implementation Guide", Versión 1.0 Final Specification, 2003.
8. D. L. McGuinness and F. Van Harmelen (eds.) "OWL Web Ontology Language Overview", W3C Recommendation, February 2004.
9. M. Nilsson, M. Palmer, and J. Brase, "The LOM RDF binding – principles and implementation", *3rd. Annual Ariadne Conference*, Katholieke Universiteit leuven, Belgium, November 2003.
10. C. L. Padrón, J. Torres, J. M. Dodero, P. Díaz, I. Aedo, "Learning Web Services Composition and Learner Communities Support for the Deployment of Complex Learning Processes", In *Proceedings of the 4th Int. Conf. on Advanced Learning Technologies (ICALT)*, Joensuu, Finland, August 2004.
11. R. S. Patil, R. E. Fikes, P. F. Patel-Schneider, D. Mckay, T. Finin, T. Gruber, and R. Neches, "The DARPA Knowledge Sharing Effort: Progress Report", *Principles of Knowledge Representation and Reasoning: Proceedings of the Third International Conference*, pp. 777-788, 1992.
12. P. Polsani, "Use and Abuse or Reusable Learning Objects", *Journal of Digital Information*, 3(4), Article No. 164, 2003.
13. A. Rawlings, P. van Rosmalen, E. J. R. Koper, M. Rodríguez-Artacho, and P. Lefrere, *Survey of Educational Modelling Languages*, CEN/ISSS WS/LT Learning Technologies Workshop, Brussels, 2002.
14. D. Rehak and R. Mason, "Keeping the Learning in Learning Objects", In A. Littlejohn (ed.), *Reusing Online Resources: A Sustainable Approach to eLearning*, Kogan Page, London, 2003.
15. A. Sarasa, and J.M. Dodero, "Towards a Model of Quality for Learning Objects", In *Proceedings of the 4th Int. Conf. on Advanced Learning Technologies (ICALT)*, Joensuu, Finland, August 2004.
16. R. Studer, V. R. Benjamins, and D. Fensel, "Knowledge Engineering: Principles and Methods", In *Data & Knowledge Engineering*, 25(1-2), pp. 161-197.