

A RDF description model for manipulating learning objects

A. Bouzeghoub, B. Defude, S. Ammour, JF. Duitama, C. Lecocq
GET/INT 9, rue Charles Fourier 91011 EVRY Cedex - France
Firstname.Lastname@int-evry.fr

Abstract

Our claim is that semantic metadata are required to allow a real reusing and assembling of learning objects. Our system is based on three models used to describe the domain, users and learning objects. In order to automatically process these models and have basic reasoning capabilities, we have defined mappings to RDF. A prototype is being developed using Sesame, a RDF tool which supports a powerful query language: SeRQL. This paper focuses on RDF mappings definition and the use of SeRQL for different functions of our system.

1 Introduction

The development of technology-enhanced learning has been very high these last years. There are now numerous pedagogical materials available on the web (the so-called learning objects - LO). An important problem is to offer tools allowing users (learners and authors) to search for existing learning objects. Authors are interested in existing objects to reuse them directly or to combine them with other objects. Learners of course want to use these objects to improve their knowledge. To facilitate the search and reuse, some standards for meta-data have been proposed such as LOM [6] and SCORM [1]. Learning objects are stored into repositories such as Ariadne knowledge pool [5] or Educanext [8] which implement LOM like meta-data. In our opinion these proposals are not so powerful because the description of learning objects does not include their semantic. We propose to extend existing standard such as LOM with a semantic part. This will allow us to define powerful search mechanism, to propose formal composition operators to create new objects, to offer different learning strategies to learners and to adapt learning objects content to users (one to one delivering). In this paper we describe the different models we use to describe the knowledge domain, users and learning objects. In order to automatically process these models, we have defined mappings to

RDF [10]. RDF is the basis for the so-called semantic web. RDF allows describing meta-data about resources. Some inference layers are currently proposed such as DAML+OIL or OWL [7]. These inference layers are not currently implemented and we have chosen to use Sesame [3] an RDF tool which supports SeRQL ("Sesame RDF Query Language") a powerful query language. The paper is structured in the following way. In section 2 we present our approach based on a three-levels model (domain model, user model, resource model) and discuss its mapping into RDF-RDFS structures. In section 3 we present the basic inference capabilities of SeRQL and the way we use it to implement different learning strategies for learners. Finally we conclude and present some open issues we want to work on.

2 The three-levels model

2.1 Introduction

Our claim is that semantic metadata are required to allow a real reusing and assembling of learning objects. Our system is based on three models presented in details in [2]: the domain model which represents the concepts covered by the LO, the user model which keeps the profile of learners, the pedagogical model which describes LO content related to the domain model. Using this knowledge we can propose sophisticated tools for searching and browsing into the LO repository. Authors can reuse and compose existing LO using operators (such as sequence, alternative, parallel) to produce new LO. A LO may be automatically adapted to a specific learner. Finally, different learning strategies (LO-based or concept-based) are proposed to learners. [4] describes in details our vision of LO adaptation and learning strategies.

In this paper we focus on the implementation of such models. We need a common language to describe the learning objects, the user model, and the domain model. RDF is a general-purpose language for representing information in the web and describing relationships among resources in terms of named

properties and values. It is a well-adapted language with respect to our goals.

As a means to obtain our goals, we need a persistent RDF storage of large quantities of RDF-metadata and an expressive query language. Sesame [3] is an extensible architecture implementing these features, including SeRQL as basic inference layer.

In the following we present successively the domain model, the user model and the learning objects model. For each model, we give an abstract presentation followed by its mappings into RDF structures.

2.2 Domain Model

Our approach uses ontologies to describe the domain model (DM). The goal of this model is to define a normalized and common referential among all users of the system (administrator, authors and learners). The precision level of the model defines the precision of the system; i.e., if we choose a very precise domain model, the system will be able to provide a more sophisticated inference.

We define a terminological ontology for every specific knowledge domain; it is intended to describe its most relevant concepts. This ontology defines concepts and relationships among them. We use two types of relationships: a narrower/broader relationship to support hierarchical links between concepts and a set of rhetorical relationships such as *contrast* or *extend*. The RDFS model corresponding to this description is given in figure 1. One can noticed the use of *rdfs:subPropertyOf* to define the different rhetorical relationships.

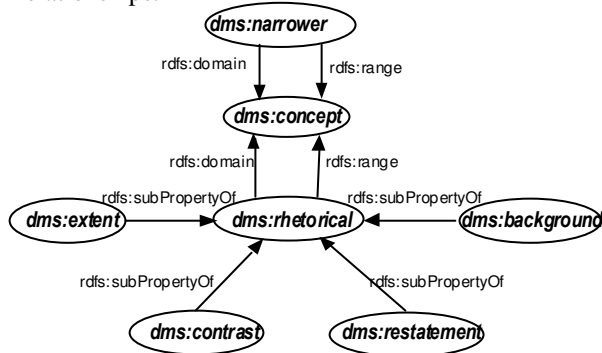


Figure 1: RDFS Domain Model

2.3 User model

An adaptive e-learning system may adapt contents depending on learners' background, preferences, and goals. Our approach considers the three aspects. It uses an overlay model to maintain an evaluation of learner, and allows to user selecting preferences as language,

format, and maximal learning time. Finally, user selects goals from concepts of domain model.

We describe a user under two facets. The first one called his preferences describes facts (name, email, language, colour,...) and is modelled with a set of couple (attribute-value). The second one called knowledge, describes concepts known by the user qualified with one or several roles (e.g. introduce, define, resume) and a weight (user level for this concept-role). A user knowledge grows dynamically when he acquires new concepts.

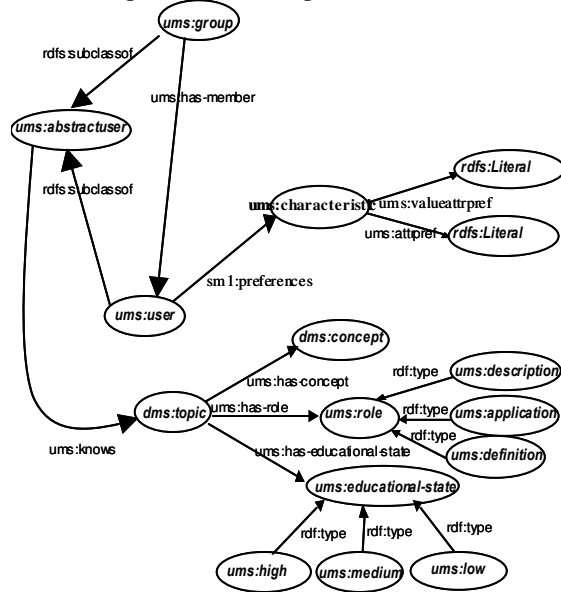


Figure 2: RDFS User Model

The RDF schema presented in figure 2 implements the user model. We introduce the notion of *abstract-user*. An *abstract-user* is either a group of users or an individual user. The user preferences are described with a dynamic but untyped property (*ums:characteristic*). This will allow adding new preferences without updating the RDFS schema. The user knowledge is represented by a *topic*. A topic is a class used to express a non-binary relation between a *dms:concept*, an *ums:role* and an *ums:educational-state*. The definition of namespaces enables us reusing for example *dms:concept* from DM schema.

2.4 Learning object model

In order to be found and re-used, a LO must be described by a set of metadata. In the LO model (LM), we distinguish two types of metadata: the first one describes LO general characteristics (e.g. author, title, language, media) using LOM standard and the second one describes the semantic of the LO. This semantic is structured in three parts: prerequisites are the LO

inputs whereas content and acquisition function are its exits.

The LO prerequisites are a set of triples (concept, role, level); the content is described with a set of couples (concept, role); the acquisition function indicates which triple (concept, role, level) will be added to the user model if a condition of validation is satisfied.

A LO can be a set of web pages, a file or a program (a simulator for example). We just suppose that it is a unit accessible via an URI. This unit can be used independently or for composition by third parties. We defined composition operators (SEQ, ALT and PAR) in order to compose recursively LOs.

A composed LO is an acyclic oriented graph where nodes are learning objects, or operators.

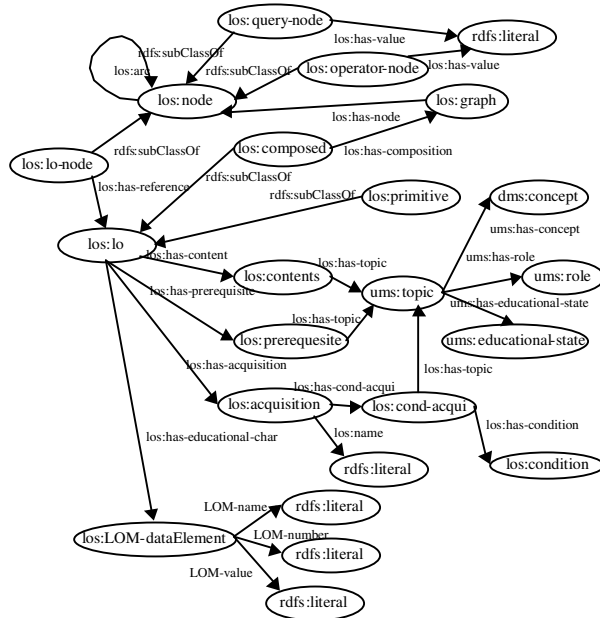


Figure 3 RDFS Learning Object Model

The RDF Schema, shown in Figure 3, reflects this model. The composition graph is described by *los:nodes* linked by *los:arcs*. There are three subclasses for *los:node*: LO-node which references a *los:lo*, *los:operator-node* which references a literal (ALT, PAR, SEQ) and *los:query-node* which references a query modelled by a literal (see intentional LO). For simplicity, we have chosen a generic representation for the LOM part and not the standard RDF binding.

2.5 Intentional learning object

An intentional learning object (ILO) is a composed LO whose composition graph has at least one node defined by a query instead of a specific object. In other words,

an ILO can have three kinds of nodes: an *operator-node* (ALT, PAR), a *LO-node*, and a *query-node*. A *query-node* is defined by an intentional query (IQ) specifying the condition to be satisfied by candidate LOs. An IQ is defined by:

$$IQ = Q_{content} \wedge (Q_{prerequisite} \vee Q_{educational})$$

$$Q_{content}|Q_{prerequisite} = (c_{1,1}, r_{1,1} \wedge \dots \wedge c_{1,k}, r_{1,k}) \vee \dots \vee (c_{n,1}, r_{n,1} \wedge \dots \wedge c_{n,m}, r_{n,m})$$

where *c* is a concept, *r* a role, *k*, *n* and *m* ≥ 1
Qeducational is a logical combination of attribute – value comparisons.

Qcontent cannot be empty but *Qprerequisite* and *Qeducational* can.

The semantic of a query-node is partly defined at authoring time: it has always a content (its *Qcontent*) but the other parts of its description may be undefined. At delivering time, all query-nodes of an ILO will be processed. If (at least) one query-node is empty (its corresponding query returns an empty set), the ILO is undefined and cannot be delivered. If all query-nodes return some LOs, these LOs will be composed by an ALT operator (the ILO is fully instantiated). A fully instantiated ILO can be adapted and delivered as a classical LO. ILOs are represented with RDFS as classical LO, except that it introduces a new type of node in graph composition (*query-node*) referencing a query described by a *rdfs:literal*. Intentional LOs are provided in order to support more generic and flexible LO for authors and to increase flexibility in the adaptation process. It allows learners to define virtual LO which can be considered as views on actual LO.

2.6 Learning objects and users properties

Our models of users and LO allow us to define several properties (in the following *L₁*, *L₂* are two learning objects and *U* a user):

- satisfaction: *U satisfies L₁* when his user model includes prerequisite of *L₁*;
- master: *U masters L₁* when his user model includes *L₁* content;
- substitution: *L₁ may be substituted to L₂* when *L₁* prerequisites are equals to *L₂* prerequisite;
- equivalent: *L₁ is equivalent to L₂* when *L₂* can be substituted by *L₁* and *L₁* content is equals to *L₂* content;
- weak precedence: *L₁ weakly precedes L₂*, if *L₁* content is included inside *L₂* prerequisites;
- strong precedence: *L₁ strongly precedes L₂*, if *L₁* content is equals to *L₂* prerequisites.

3 SeRQL as a basic inference layer

SeRQL is a new RDF/RDFS query language, which has interesting capabilities. Besides its integration with RDF-RDFS, it supports graph transformation and expressive path expressions. A SeRQL query is defined using Select...From...Where statements. The FROM clause defines path expressions used into the query. The WHERE clause filters graphs denoted by the FROM clause whereas the SELECT clause defined the interesting nodes.

SeRQL can be used at numerous levels in our system. It is well adapted to search or browse the domain model (with path expressions), user model and LO model but also to express queries in intentional LO and even in internal processes. In this section, we will focus on the use of SeRQL to support three interaction modes for learners: concept-based, query-based and LO-based learning (see [4] for details). In the first one, a learner selects LOs with a condition based on a set of concept he/she wants to learn about. In the second one, a learner selects LOs with a condition based on LOM data elements. In the last one, he/she selects a specific LO. In all modes, this selection step is followed by an adaptation step where the system adapts the selected LO to the learner. In the rest of the section, we illustrate these modes with use cases.

3.1 Use case 1: concept mode

- 1) The learner selects the concepts he/she wants by browsing the domain model. We suppose he/she selects *SQL* concept with *description* role and *Relational algebra* concept with *illustration* role.
- 2) Using these inputs, the system builds a SeRQL query (see Query 1).

```

Query 1
SELECT LO
FROM {los:LO}<los:has-content>
{los:contents} <los:has-topic>
{ums:topic}; <ums:has-concept>
{dms:Concept}; <ums:has-role>
{ums:Role};
WHERE
(Concept="SQL" and
Role="description") or
(Concept="Relational algebra" and
Role="illustration")
    
```

The system processes this SeRQL query and produces the output represented in figure 4.

3.2 Use case 2: query mode (LOM)

- 1) The learner defines his/her search criteria based on LOM metadata. We suppose that he/she selects LOs written in English or authored by Jim. This search criteria is expressed with SeRQL query 2:

Query 2 retrieve all LOs written in English or authored by Jim

```

SELECT LO
FROM {los:LO} <los:has-educational-
char> {los:LOM-dataElement};
<los:LOM-number> {number}; <los:LOM-
value> {value};
WHERE (number="1.3" and value="en")
or (number="2.3.1" and value="author"
and number="2.3.2" and value="jim")
    
```

- 2) The system processes this SeRQL query and produces an output.

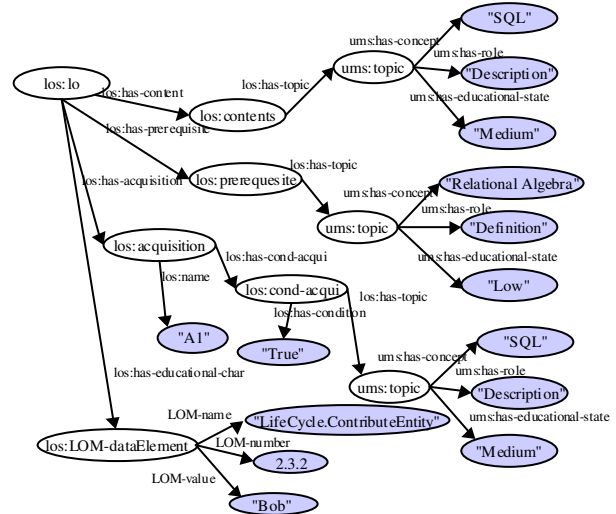


Figure 4 Result of query 1

3.3 Use case 3: LO mode

- 1) The learner (let say Bob) wants to follow an existing LO (let say the *DataBase* course).
- 2) In this mode an important function for adaptivity is to verify if the learner model satisfies the prerequisites of the LO (*satisfaction* property). For this, the system retrieves the prerequisites of the LO: <http://int-evry.fr/CM-exemple#DataBase>:

```

Query 3
SELECT Concept, Role,
EducationalState
FROM {http://int-evry.fr/CM-
exemple#DataBase} <los:has-
prerequisite> {prerequisite}
<los:has-topic> {topic} <los:has-
concept> {Concept};
<los:has-role> {Role};
<los:has-educational-state>
{EducationalState}
    
```

We suppose that the result R_{Q3} is equal to {<"relational algebra", "definition", "low">} Then the system retrieves Bob's knowledge:

```

Query 4
SELECT Concept, Role,
EducationalState
    
```

```
FROM {http://int-evry.fr/U-  
example#Bob} <ums:knows> {topic}  
<ums:has-concept> {Concept};  
<ums:has-role> {Role};  
<ums:has-educational-state>  
{EducationalState}
```

The result R_{Q_4} is equal to {<"relational algebra", "definition", "medium">, <"SQL", "description", "low">, <"GIS", "-", "very low">}

R_{Q_3} is included in R_{Q_4} and the *satisfaction* property is satisfied

4 Conclusion

In this paper, we present the RDF implementation of a description model for reusable learning objects. This model is based on a semantic description of each learning object. Our claim is that semantic metadata are required to allow a real reusing and assembling of learning objects.

This semantic allows describing domain model, user model, and LO model and provides authors and learners with powerful mechanisms to manage learning objects, concepts, and users (e.g. browsing, querying, composing, classifying, etc). RDF allows us to easily support these models. RDF schema allows us to interpret learning object metadata within or across several communities. It serves as a language for facilitating the worldwide exchange of learning objects and it proposes several reasoning layers.

We are currently implementing a prototype using Sesame [3] that will allow us to validate our approach. Sesame offers a storage layer for RDF statements and RDFS. Its only reasoning layer is the SeRQL query language. This language has interesting capabilities but lacks of expressivity to handle all type of queries we are interested in.

Several points are still open. For the moment we consider that all LO are stored in the same computer but we want to study different distributed architectures for our system. RDFS is not so powerful to express constraints on RDF statements and we will address the problem of typing LO. For example it should be interesting to constraint the size of a LO in terms of number of presented concepts.

5 References

- [1] ADVANCED DISTRIBUTED LEARNING INITIATIVE, *Sharable Content Object Reference Model. The SCORM Content Aggregation Model. Version 1.2.* Available from: <<http://www.adlnet.org/>>. October 1, 2001
- [2] A. Bouzeghoub, C. Carpentier, B. Defude, JF. Duitama "A Model of Reusable Educational Components for the Generation of Adaptive Courses", Proc. First International Workshop on Semantic Web for Web-Based Learning in conjunction with CAISE'03 Conference, Klagenfurt, Austria, june 2003
- [3] Broekstra, J., Kampman A., and Van Harmelen F. "Sesame: An Architecture for Storing and Querying RDF Data and Schema Information", Semantics for the WWW, MIT Press, 2001, D. Fensel, J. Hendler, H. Lieberman and W. Wahlster
- [4] JF. Duitama, B. Defude, A. Bouzeghoub, C. Carpentier "A Framework for the Generation of Adaptive Courses based on Semantic Metadata", Proc. Workshop on Metadata and Adaptability in Web-based Information Systems in conjunction with OOIS'03 Conference, Genève, september 2003
- [5] E. Duval et al. "The Ariadne Knowledge Pool System", Communication of the ACM, 44(5), 72-78, 2001
- [6] IEEE, "Draft Standard for Learning Object Metadata", (IEEE P1484.12.1.), 2002 Available from <<http://ltsc.ieee.org/>>
- [7] I. Horrocks, F. van Harmelen, Peter Patel-Schneider, T. Berners-Lee, D. Brickley, D. Connolly, M. Dean, S. Decker, D.Fensel, P. Hayes, J. Heflin, J. Hendler, O. Lassila, D. MCGuinness, and L.A. Stein. "DAML+OIL." Available on <http://www.daml.org/2001/03/daml+oil-index.html>, March 2001.
- [8] E. Law et al. "EducaNext: A Service for Knowledge Sharing", Proc. of the Ariadne Conference, Leuven, Belgium, Nov 2003
- [9] S. Ranwez, T. Leidig and M. Crampes, "Pedagogical ontology and Teaching Strategies: A New Formalization to improve Life-long Learning", *Journal of Interactive Learning Research*, Vol. 11, Number 3-4, AACE, 2000.
- [10] F. Manola and E. Miller "RDF Primer", W3C Working Draft 23 January 2003, Available on <http://www.w3.org/TR/2003/WD-rdf-primer-20030123/>