The Semantic Web for Learning Resources

Yolaine Bourda, Bich-Liên Doan
Ecole Supérieure d'Électricité
3 rue Joliot Curie, 91192 Gif-sur-Yvette, France
{Yolaine.Bourda, Bich-Lien.Doan}@supelec.fr

Abstract

This paper discusses a vision of the Semantic Web for learning resources and some problems found. We consider RDF as the first step to the Semantic Web. But using all the possibilities of RDF is not that simple. Constructing an RDF schema for a given metadata schema, translating RDF descriptions and querying a set of descriptions are the main problems we discuss here. The example of the LOM (Learning Object Metadata), the IEEE standard for education, is used. But even if some problems are not solved, it is important to use RDF now in a simple way to find more easily resources.

1. Introduction

The World Wide Web has become the largest database in the world. But finding relevant information is more and more difficult. To solve this problem, the W3C recommends the use of metadata and their encoding in RDF. The most common metadata schema is the Dublin Core one, soon an ISO standard. Its aim is the description of every digital resource. But it's use is not as widespread as it should be.

In the field of education, the IEEE LOM (learning Object Metadata) is the most common. But it's use is also lower than what it should be. It is used by consortiums and institutions to describe learning objects belonging to a pool. It is also widely used by some peer to peer networks like Edutella. These two kinds of use could be qualified as closed world use as data and metadata are not directly accessible on the Web.

How many available learning resources are there on the Web? How can we help learners and teachers to find them? By finding a simple way to index those resources. If everyone adds descriptions to his/her resources and if those descriptions were compatible then anyone could retrieve the resources needed using semantic criteria. The advantage is really important but the effort must be weak otherwise it will not be done.

2. Indexing Learning Resources

To be able to index learning resources (i.e. to associate to each resource a set of RDF descriptions encoding metadata) we need: an RDF Schema corresponding to the metadata schema; an RDF schema based editor.

While metadata standards define the structure of a metadata instance, they do not define how a system will represent or use a metadata instance for a resource. For example, the LOM standard defines a set of descriptors and says nothing about their implementation. For the LOM, the IMS consortium has developed an XML binding and an RDF binding. The construction of the XML binding has been done without problems. But finding an RDF schema is not that simple, and IMS doesn't manage most of the problems. For example, they consider, more or less, that all the metadata for one resource are contained in a single RDF document.

To give examples of correct RDF descriptions using the LOM is easy and many examples can be found. But to make an RDF schema for the LOM is a hard task. What we want is an RDF schema containing all the meanings of the metadata schema, to be able to use RDF tools. The validation of a distributed set of descriptions is a huge problem. Who could add a description to a resource? Is it necessary to impose that all the descriptions should be in one place? This is comprehensible in a closed world but we are in a web semantic approach in which all the descriptions of a resource are distributed over the web.

When one wants to construct an RDF schema for a metadata schema, the first idea that comes to mind is: find an RDF property for each descriptor. But when you try to do that, many problems arise. With RDF, the only possible constraints on properties are `rdfs:domain` and `rdfs:range`. There are not sufficient.

It is not possible to constraint one property to be mandatory. This is understandable because if the descriptions are dispatched all over the Web and if the node with that property is broken then the set of remaining descriptions doesn't validate the schema.

It is not possible to say that a property can only be used once for each resource. Unfortunately, some elements like Title in the LOM are concerned. We suggest using containers for properties which can be present several times. Those for which the order is important are encoded with `rdf:seq`, and `rdf:bag` is used for the others. This looks satisfactory but the validation of a set of distributed descriptions can be very complex. Is it necessary to search all the Web to be sure that a property is really present only once?

Sometimes there are constraints linking the values of the properties. In the LOM, one example is : if

Proceedings of the The 3rd IEEE International Conference on Advanced Learning Technologies (ICALT'03)
0-7695-1967-9/03 $17.00 © 2003 **IEEE**

`Structure = atomic` the `AggregationLevel = 1`. For this, we suggest defining subclasses.

The meta-metadata set of descriptors describes by whom and when the metadata was associated to the resource. It is not possible to imagine an instance for each RDF description. It thus should be supposed that the set of RDF descriptions concerning a resource is also a resource and is contained in one file or one database. This also means that we must divide the set of metadata in two: the first one includes metadata given by the author and the second one what we can call "annotation metadata". Another solution is to use the reification mechanism. But, do we want to say that "this learning object has this title" or "the creator of the meta-metadata statement says that this learning object has this title". This is not exactly the same meaning. Furthermore, what kind of requests should be made using an RDF query language? Reified triples are not semantically equivalent to non-reified ones. An assertion of a reification of a triple does not implicitly assert the triple itself.

2.2. Indexing learning resources

Associating RDF descriptions written as triples to a resource respecting a given schema is not really difficult. It could be done using a simple text editor but this means could be considered old fashioned by many users. We don't want to associate an XML DTD or schema to each RDF schema although this is easy and will allow the use of XML editors. The main reason is that we don't want to have an XML based approach but rather an RDF one. RDF has a semantic approach and not a structure one like XML. RDF descriptions could be assigned to resources in a modular way. Now, we can find some RDF editors generating RDF triples and so we recommend their use.

4. Finding learning resources

When files containing RDF descriptions are easily available, in fact as available as the resources they describe, one can imagine a web crawler looking for that kind of files. After that, we need a tool to find the resources corresponding to a request. But two problems remain :

- How can we query, in the same request, descriptions based on different schemas?
- How can we take advantage of RDF, deducing new properties from those given by the users.

4.1. Transforming RDF descriptions

Multiple RDF schemas can be found for the same domain (education for example) resulting in the lack of interoperability. So, translating RDF descriptions is mandatory if a user wants to request multiple descriptions using multiple schemas even if one doesn't know that different schemas exist. RDF descriptions are based on RDF schemas so, the translations of descriptions may be deduced from the translations of schemas. But, transforming one schema into another must be done while preserving the meaning of the schemas.

4.2. Searching RDF resources

It is easy to imagine queries like: find all the resources having a certain value for a certain property. But what kind of deductions could we make from a set of descriptions? If article is a sub-property of document, a request looking for all the documents written by a particular author must give all the articles by that author. So a logic based query language is needed.

5. Conclusion

Metadata may be considered as the first step to the semantic web and RDF is a foundation for processing Metadata. But many problems remain unsolved in the RDF binding of a metadata schema. Maybe the most important one is that many people forget that RDF descriptions can be distributed all over the web. So, as a first step to the generalization of RDF, we suggest building simple RDF schemas (to share the vocabulary) to help the indexation of distributed resources by "individual" users and more complex schemas for the indexation of "institutional" resources.

Developing learning resources is time consuming and they must be reuse to decrease their cost. But to reuse a resource means it has been found before. So everything allowing resources to be found easily (and with some semantic criteria) must be encouraged. It is a first step to a semantic web for learning resources.

6. References

[1] "Resource Description Framework (RDF), Model and Syntax Specification", W3C Recommendation, 02/ 1999 http://www.w3.org/TR/1999/REC-rdf-syntax-19990222

[2] "RDF Vocabulary Description Language 1.0: RDF Schema", W3C Working Draft 12 November 2002 http://www.w3.org/TR/rdf-schema

[3] Mikael Nilsson "Semantic issues with the LOM RDF binding",2003-01-15 http://kmr.nada.kth.se/el/ims/md-lom-semantics.html

[4] IMS Resource Description Framework(RDF) Bindings http://www.imsglobal.org/rdf/

[5] T. Kunze, J. Brase, W. Nejdl "Editing Learning Objects Metadata: Schema Driven Input of RDF Metadata with the OLR3-Editor", *Semantic Authoring, Annotation & Knowledge Markup Workshop.* 15th European Conf. on Artificial Intelligence, July 22-26, Lyon, France.