A Semantic web approach for e-learning platforms

Miguel B. Alves[1]

[1]Laboratório de Sistemas de Informação, ESTG-IPVC
4900-348 Viana do Castelo.

mba@estg.ipvc.pt

Abstract. *When lecturers publish contents in an e-learning platform like a course degree or the bibliography for a given course, normally they do that either by uploading a document (MSWord, pdf) or by creating a resource that will be part of a webpage. In both cases, this kind of information is static and is only useful for human readers, the information is not open to other systems, and in general to the world. In this work, we propose to enrich an e-learning platform with semantic web content so that information is available to external systems. Concretely, we will develop our approach over the Moodle platform, the widest e-learning platform used. Moreover, we will focus on representing the course's degree and its bibliography.*

1. Introduction

Lets consider that a university wants to know which books are recommended in all the courses of its degrees. They need this information to know if its library is well-served of recommended books. The "well-served" conception has two dimensions: a) all the recommended books should exist in the library; b) all recommended books should exist in the library in the right quantity, which means, there are books that are widely used and there should be several copies available. The university policy forces teachers to publish the bibliography of each course in the e-learning platform. However, they concluded that all information is available to the users but it cannot be accessed in a structured way. It is not a desirable solution to develop another system for that particular purpose because of its cost. The ideal is to access the information that is published in e-learning platform, a web-based system, in a structured way. This relates immediately with the Semantic Web initiative where contents are machine interpretable. The Semantic Web is a proposal of the World Wide Web inventor Tim Berners-Lee and colleagues[Berners-Lee et al. 2001] that the Web as a whole can be made more intelligent and perhaps even intuitive about how to serve a user's needs. Berners-Lee observes that although search engines index much of the Web's content, they have little ability to select the pages that a user really wants or needs. He foresees a number of ways in which developers and authors, single or in collaboration, can use self-descriptions and other techniques so that context-understanding programs can selectively find what users want. However, despite the designation, Semantic Web is not only for web content but, in general, it is an interpretable machine approach and this has application in many areas. The global vision of the development of the Semantic Web is to make the contents of the Web machine interpretable. To achieve this overall goal, ontologies play an important role as they give the means for associating precisely defined semantics with the content that is provided by the web. Ontologies are defined as the representation of the semantics of terms and their relationships. They consist of concepts, concepts' atributes and relationships between concepts, all expressed in linguistic terms[Guarino et al. 1993].

This paper reports the project developed in a polytechnic institute which uses Moodle as web e-learning platform. The purpose of the project is the development of Moodle plug-ins to bring the bibliographic information out of the e-learning system. Moreover, the institute requests the teacher of a given course to do the mapping between course contents and the recommended bibliography. The purpose is to create the necessary infrastructure that allows the future development of a system where the student can consult the bibliography of a given subject in a course. He will be able to see if a book is available in library, connect to an on-line sales company if he desires to buy the book, and so on. This document is organized as follows. Section 2 describes our approach to enrich a e-learning platform with semantic information. Section 3 discusses future work, and the paper ends in Section 4 with conclusions.

2. Semantic Web Approach in E-Learning Systems

In [Diaconescu et al. 2008] are discussed the advantages of a semantic web approach in e-learning systems. Briefly, representing data in the Resource Description Framework (RDF) [Brickley and Guha 2004] [Tauberer 2006], instead of a traditional approach as relational databases, is a shift to the open world with many distributed resources, identified by URIs as a mechanism for referring to global entities on which there is some agreement among multiple data providers. Queries can be performed not only over a single database, but over the content of several distributed educational systems, including resources, which are externally available on the Web. RDF is a W3C [1] standard for modelling and sharing distributed knowledge based on a decentralized open-world assumption. RDF was designed as a metadata model and it has come to be used as a general method for conceptual description or modelling of information that is implemented in web resources. Knowledge is expressed by triples consisting of subject, predicate and object (like a short English sentence), also known as statement, forming RDF graphs. In the Web, we use RDF to make statements about resources. In particular, we can classify the resource. RDFa [W3C 2012] is a specification for attributes to express structured data in any markup language. RDFa, which means RDF in HTML *a*ttributes, adds a set of attribute-level extensions to XHTML for embedding rich metadata within Web documents. This allows web pages to be understandable by machines, give information to the browsers and search engines about the pages. A Semantic Web approach also allows reasoning over the contents. Rules can be defined over the RDF statemensts, and those rules can be extended any time. To describe a bibliography, we make use of the Bibliographic Ontology (BIBO)[D'Arcus and Giasson 2009], which is an ontology for the semantic Web to describe bibliographic things like books or magazines. To describe a bibliographic resource, BIBO makes use of the Dublin Core ontology. Dublin Core element set [DCMI 1998] is a flexible and usable metadata schema enabling information exchange and integration between digital sources. It is widely used by almost all digital libraries since it is simple, small and easily expandable, and provides qualifiers that enable the semantic expression. The significant role of DC in data exchange is obvious due to the fact that there are mappings from and to it by many widely used metadata schemas [Day 2002]. Dublin Core is widely used to describe digital materials such as video, sound, image, text and composite media like web pages. Although Dublin Core can be used to describe bibliography, we use BIBO Ontology because it has a higher degree of richness to describe

[1] `http://www.w3.org/`

books, for example, isbn10 and isbn13 properties. One the of purposes of the project is to encourage the teachers to detail the information mapping course contents with bibliography, driving the student exactly to the contents that he should focus. For that, we need semantic information about the course, its contents and a way to map the contents with bibliography. We make use of the Academic Institution Internal Structure Ontology (AIISO) [Styles and Shabir 2008], that provides classes and properties to describe the internal organizational structure of an academic institution and TEACH, the Teaching Core Vocabulary [Kauppinen and Trame 2011], which is a lightweight vocabulary providing terms to enable teachers to relate things in their courses together.

2.1. Linked Open Data Project

The aim of Linking Open Data Project [Bizer et al. 2009] is using the web to create typed links between data from different sources via mapping of ontologies. In this way, instead of having isolated islands we have global interlinked datasets. Linked Data refers to data published on the Web in such a way that it is machine-readable, its meaning is explicitly defined, it is linked to other external data sets and can be accessed by them. Linked Data principles provide a basic recipe for publishing and connecting data using the infrastructure of the Web while adhering to its architecture and standards.

All ontologies used in this work belong to Linked Open Data, under `LinkedUniversities.org`. *Linked Universities* is an alliance of european universities engaged into exposing their public data as linked data. Using technologies such as RDF and SPARQL, it gives direct access to information such as their publications, courses, educational material, etc.

2.2. Semantic Annotation

In this subsection we will detail how the webpages can be annotated to be enriched with semantic information. For a better understanding of this work, next we summarize the namespaces used. Besides, to save space, in all of the HTML excerpts below we omit *namespaces* declaration.

```
xmlns:dc='http://purl.org/dc/terms/'
xmlns:aiiso= 'http://purl.org/vocab/aiiso/schema#'
xmlns:teach='http://linkedscience.org/teach/ns#'
xmlns:bibo='http://purl.org/ontology/bibo/'
```

Lets consider the following excerpt of HTML from a web page course in Moodle.

```
<div typeof="bibo:Book"resource="3642159699≫
<h2><span property="dc:title≫ A developer's guide to the semantic web</span></h2> <p>
Author:  <span property="dc:creator≫Liyang Yu</span><br>
ISBN-10:  <span property="bibo:isbn10≫3642159699</span> | ISBN-13:  <span
property="bibo:isbn13≫978-3642159695</span> <br> </p> </div>
```

The purpose is to enrich the web page with semantic information that can be used outside of Moodle. The resulting page might be:

```
<div typeof="bibo:Book"resource="ISBN:3642159699≫
<h2><span property="dc:title≫ A developer's guide to the semantic web</span></h2>
<p> Author:  <span property="dc:creator≫Liyang Yu</span><br>
ISBN-10:  <span property="bibo:isbn10≫3642159699</span> | ISBN-13:  <span
property="bibo:isbn13≫978-3642159695</span><br> </p> </div>
```

As we can see, the web page was enriched with semantic information that can be used by other systems. BIBO ontology is used to describe the book. Next, we list the RDF information extracted from the previous example using a RDFa parser service [2].

```
<ISBN:3642159699> <rdf:type> <bibo:Book> .
<ISBN:3642159699> <dc:title> "A developer's guide to the semantic web".
<ISBN:3642159699> <dc:creator> "Liyang Yu".
<ISBN:3642159699> <bibo:isbn10> "3642159699".
<ISBN:3642159699> <bibo:isbn13> "978-3642159695".
```

Now, consider a course named *Semantic Web*. One of the topics of this course is *The RDF language and its XML serialization*. Next, we list an excerpt of a HTML webpage to describe the course and its contents, enriched with semantic information.

```
<div typeof="aiiso:Course"resource="semweb≫
<h2><span property="dc:title≫Semantic Web</span></h2>
<p> </div>
<div typeof="aiiso:Programme"resource="semweb-topic2≫
<span property="dc:title≫The RDF language and its XML serialization</span>
<span rel="aiiso:knowledgeGrouping"resource="urn:semweb≫</span> </div> </p>
```

The course and its contents are modelled with AIISO ontology. A course is of the type of *Course* class. We modelled the topics of the course with the class *Program*. Both *Course* and *Program* are sub-classes of the class *KnowledgeGrouping*, which represents a collection of resources, learning objectives, timetables, and other material. A *KnowledgeGrouping* may be contained by another *KnowledgeGrouping* or an organizational Unit using the *knowledgeGrouping* property. We use this property to relate a course with its contents. Additionally, this property allows the definition of sub-topics, in a hierarchical view.

Now, lets consider that one recommended reading to the topic *The RDF language and its XML serialization* of the *Semantic Web* course is the chapter 2 of book *A developer's guide to the semantic web*, named *The Building for the Semantic Web:RDF*. This mapping can be done in two ways: a) when the teacher is editing the course content, he indicates the recommended reading for a given topic; b) when the teacher is editing the bibliography, he indicates the topics which are supported by that book or chapter. To model chapters of a given book, BIBO ontology has the class *Chapter*. We make use of Dublin Core property `isPartOf` to define that a chapter belongs to a given book. The following HTML webpage excerpt shows the semantic information associated with a book chapter:

```
<div typeof="bibo:Chapter"resource="ISBN:3642159699-chapter2≫
Chapter 2 - <span property="dc:title≫ The Building for the Semantic
Web:RDF</span>
<span property="bibo:chapter≫2</span>
<div rel="dc:isPartOf"resource="ISBN:3642159699≫</div></div>
```

Let us consider first the model where the teacher associates the recommended reading when he is editing the course content. For that, we make use of the property `reading` of TECH ontology. The HTML excerpt below shows how modelling is done.

```
<div typeof="aiiso:Course"resource="semweb≫
<h2><span property="dc:title≫Semantic Web</span></h2>
```

[2] `http://rdf-in-html.appspot.com/`

```
<p> <ul> <li typeof="aiiso:Programme"resource="semweb-topic2≫
<span property="dc:title≫The RDF language and its XML serialization</span>
<span rel="aiiso:knowledgeGrouping"resource="urn:semweb≫</span>
<span rel="teach:reading"resource="ISBN:3642159699≫</span>
</li> </ul> </p> </div>
```

The corresponding RDF triples are:

```
<urn:semweb> <rdf:type> <aiiso:Course> .
<urn:semweb> <dc:title> "Semantic Web".
<urn:semweb-topic2> <rdf:type> <aiiso:Programme> .
<urn:semweb-topic2> <dc:title> "The RDF language and its XML serialization".
<urn:semweb-topic2> <aiiso:knowledgeGrouping> <urn:semweb> .
<urn:semweb-topic2> <teach:reading> <ISBN:3642159699> .
```

Consider now that the teacher indicates the book or the chapters of books that support a given topic when he is editing the bibliography. For that, he make use of the property `isReferencedBy` of Dublin Core Ontology.

```
<div typeof="bibo:Chapter"resource="ISBN:3642159699-chapter2≫
Chapter 2 - <span property="dc:title≫ The Building for the Semantic Web:RDF</span>
<span property="bibo:chapter≫2</span>
<span rel="dc:isPartOf"resource="ISBN:3642159699≫</span>
<span rel="dc:isReferencedBy"resource="semweb-topic2≫</span> </div>
```

The corresponding RDF triples are:

```
<ISBN:3642159699-chapter2> <dc:title> "The Building for the Semantic Web:RDF".
<ISBN:3642159699-chapter2> <dc:isPartOf> <ISBN:3642159699> .
<ISBN:3642159699-chapter2> <dc:isReferencedBy> <urn:semweb-topic2> .
```

3. Future Work

The next step is to develop plugins to Moodle to permit inserting semantic information in webpages without needing technical knowledge. The user does not need to have knowledge about semantic web and is not expected to fill in the webpages with semantic information. That semantic information should be inserted by user-friendly tools, incorporated in Moodle as plugins. These user-friendly tools also can use information from ontologies. For example, to help the user in bibliography editing, an ontology like RDF Book Mashup[Bizer et al. 2007] can be used, helping filling in of data fields.

Another important task that is not directly related with this work but it is important to other projects is extracting the semantic information from webpages and keeping it in a database in order to be used. Jena GRDDL (Gleaning Resource Descriptions from Dialects of Languages) Reader[3] can be used to extract RDF data from HTML pages. This information can be kept in any RDF triple database or even in a relational database (however, this last option can result in drawbacks in using semantic reasoning). In [Diaconescu et al. 2008] is introduced how we can deal with semantic information.

In the future, it should be interesting to extend the semantic information to other fields of education, towards a completly linked university, using approaches and ontologies present in literature for that purpose, which we introduce in this work.

[3]http://jena.sourceforge.net/grddl/

4. Conclusion

In this work, we presented a semantic web approach in Moodle, the widest e-learning platform used. The purpose is to enrich web contents with semantic information, opening the contents to the open world. Adopting the linked open data principles, the information contained in webpages is available to outside systems, which can read and interpret the information without any kind of specifications or protocols. This work is an on-going project to make the information in one polytechnic institute accessible to other systems that can be developed in the future. However, as these systems are not planned yet, the purpose is make the information available in a standard and open way. This work focuses on modelling courses and their contents and bibliography, allowing mappings between them.

Referências

Berners-Lee, T., Hendler, J., and Lassila, O. (2001). The semantic web a new form of web content that is meaningful to computers will unleash a revolution of new possibilities. *Scientific American*, 284:28–37.

Bizer, C., Cyganiak, R., and Gauss, T. (2007). The rdf book mashup: from web apis to a web of data. In *3rd Workshop on Scripting for the Semantic Web, ESWC, Innsbruck, Austria, June*, volume 6, pages 1613–0073.

Bizer, C., Heath, T., and Berners-Lee, T. (2009). Linked data - the story so far. *International Journal on Semantic Web and Information Systems*, 5(3):1–22.

Brickley, D. and Guha, R. V. (2004). Rdf vocabulary description language 1.0: Rdf schema. *W3C Recommendation, http://www.w3.org/TR/2004/REC-rdf-schema-20040210/*.

D'Arcus, B. and Giasson, F. (2009). Bibliographic ontology specification. http://bibliontology.com/specification.

Day, M. (2002). Mapping between metadata formats. *UKOLN UK Office for Library and Information Networking*.

DCMI (1998). Dublin core metadata element set, version 1.0: Reference description. http://dublincore.org/documents/1998/09/dces/.

Diaconescu, I.-M., Lukichev, S., and Giurca, A. (2008). Semantic web and rule reasoning inside of e-learning systems. In Badica, C. and Paprzycki, M., editors, *Advances in Intelligent and Distributed Computing*, volume 78 of *Studies in Computational Intelligence*, pages 251–256. Springer Berlin / Heidelberg.

Guarino, N., Poli, R., and Gruber, T. R. (1993). Toward principles for the design of ontologies used for knowledge sharing.

Kauppinen, T. and Trame, J. (2011). Teaching core vocabulary specification. http://linkedscience.org/teach/ns/.

Styles, R. and Shabir, N. (2008). Academic institution internal structure ontology (aiiso). http://vocab.org/aiiso/schema.

Tauberer, J. (2006). What is rdf. http://www.xml.com/pub/a/2001/01/24/rdf.html.

W3C (2012). Rdfa 1.1 primer. http://www.w3.org/TR/xhtml-rdfa-primer/.