Use of a Semantic Learning Repository to Facilitate the Creation of Modern e-Learning Systems

Xavier Ochoa  
Escuela Superior Politécnica del Litoral  
Km. 30.5 Vía Perimetral  
Guayaquil, Ecuador  
xavier@cti.espol.edu.ec

Gladys Carrillo  
Escuela Superior Politécnica del Litoral  
Km. 30.5 Vía Perimetral  
Guayaquil, Ecuador  
gcarrillo@cti.espol.edu.ec

Cristian Cechinel  
Federal University of Pelotas  
R. Felix da Cunha 630, Centro  
Pelotas (RS), Brasil  
contato@cristiancechinel.pro.br

ABSTRACT

Learning Object Repositories (LOR) has been usually implemented as traditional document stores. In this paper we explain the design of a Semantic Learning Repository that expand the concept of LORs to include linked information of entities not usually referred as Learning Objects, but necessary for the implementation of more advance e-learning systems. To demonstrate the usefulness of this new concept, this paper presents the design and evaluation of a Personalized Learning Path Recommender based on the Semantic Learning Repository. The main result obtained from the evaluation is that the inclusion of external information enables more accurate recommendations and these recommendations has a measurable impact on the student learning.

Categories and Subject Descriptors

K.3.1 |Computing Milieux|: Computers and Education-Computer Uses in Education

Keywords

Learning Object Repositories, Semantic Web, Recommender Systems

1. INTRODUCTION

Learning Object Repositories (LOR) has been the backbone for the construction of e-learning systems that provide access to a large amount of learning resources [9]. Traditionally, these LORs has been implemented as document repositories, that is, they are centered around only one entity, in this case, the Learning Object [20]. The information stored in a traditional LOR is the learning resource file and the metadata, in a predefined format, describing that resource [11]. In the case of the learning resource file, some LORs store only a reference to where the file is stored and these LORs are called “Referatories”.

Permission to make digital or hard copies of all or part of this work for personal or classroom use is granted without fee provided that copies are not made or distributed for profit or commercial advantage and that copies bear this notice and the full citation on the first page. To copy otherwise, to republish, to post on servers or to redistribute to lists, requires prior specific permission and/or a fee.

INTERACCIÓN 2014, September 10-12, 2014, Puerto de la Cruz, Tenerife, Spain  
Copyright 2014 ACM 978-1-4503-2880-7 ...$15.00.

The traditional design of LOR while useful for the direct retrieval of Learning Objects, present several shortcomings when used in real-life e-learning systems. First, e-learning systems should manage much more diverse entities that only the learning objects. The learner, the teacher, the lesson (sequence) should also need to be taken into account. E-learning systems usually solve this shortcoming having several repositories for different type of entities: one for the learning objects, other of the user profile and another for the lesson structure. While this let the e-learning system to store all the needed information, it adds complexity to the system and makes very difficult to maintain the very necessary relationships (links) between entities [12].

A second major shortcoming of traditional LORs is their reliance on a single metadata format to describe the learning resources. In the best case scenario this format will be a standard such as Learning Object Metadata (LOM) or Dublin Core (DC), otherwise it will be an ad-hoc structure. Due to this reliance on a single metadata format, a whole area of research on Learning Object Interoperability has been developed in order to be able to interchange information between several repositories [18]. These interoperability issues, again, add complexity to the design of e-learning system, specially if it is desired that their data remain open for others to be used.

Finally, being based on predefined formats for their metadata, traditional LORs are designed to operate with a static structure. If new elements or entities are added to the e-learning system, the LOR will be unable to accommodate them and a new repository, or a major re-design, will be needed to store their information. Rapid changing and adaptable e-learning systems could only communicate with LORs as a source of information, but not as a main component between the architecture of the system [3].

All these shortcomings demand a drastic redesign of the concept of Learning Object Repository to be the main persistence component of modern e-learning systems. This paper introduces the concept of a Semantic Learning Repository that try to solve the problems discussed above and propose a more flexible architecture adjustable to very different kind of e-learning systems. The structure of this paper is as follow: Section 2 describes other solutions to the traditional LOR problems. Section 3 proposes the concept of Semantic Learning Repository and the reasons why it is more flexible that other current solutions. Section 4 presents the design of a Personalized Learning Paths Recommender System based on a Semantic Learning Repository. Section
5 evaluates the workings and impact of the Recommender system. Finally, conclusions and recommendations are presented to the reader.

2. RELATED RESEARCH

One of the first identified problems and, arguably the one that draw more research attention, is the limitation of using LOM, DC or any other kind of static metadata format as the only source of information about the learning object. Usually metadata is created and stored alongside the object at publication time [13]. However, much more valuable information about the object is created through the lifetime of the objects as teachers and learners (re-)use it, rate it, comment it and share it. Existing metadata standards were not well suited to store this kind of information. One of the first works that try to capture and use this information was Contextual Attention Metadata (CAM) [22]. Martin proposed a metadata format to capture all the interactions that users have with learning objects in order to complement existing standards. A larger movement, supported by the Department of Education of USA, is the Learning Registry [1].

In the same spirit as CAM, the Learning Registry goal is to capture what they call “paradata”, social interactions where learning objects are involved.

Other line of research to improve the current design of LORs is the inclusion of Semantic Technologies to increase its capabilities and flexibility. Soto-Carrion et al. [19] proposed the idea of a Semantic Learning Object Repository. This repository enables the use of different metadata formats, expressed as Resource Description Framework (RDF) triplets, to describe a Learning Object. The use of several formats to store information about the object facilitates the interoperability with other repositories and the implementation of more advance search facilities.

While these works expand and improve the concept of Learning Object Repository, they do not solve all their current problems. Adding usage and social information only make it more visible the fact that the entities that generate that information (that is the learners and the instructors) are not present or described in the repository. Using Semantic Technologies could solve this issue, linking learning objects and social actions with new entities, but current implementations do not explore this opportunity.

In the following sections, a model that mixes and builds over these ideas, the Semantic Learning Repository, is presented. Its main contribution is to incorporate linked entities inside the repository in order to better support modern e-learning systems that should deal with more than only the search and retrieval of learning objects.

3. SEMANTIC LEARNING REPOSITORY

The idea of Learning Object was central to the development of e-learning. Learning Object Repositories that were able to describe, store, search and retrieve large amounts of learning resources were a disruptive and widely implemented idea [17]. During that time, the phrase “Content is king” [10] was used with reference of the importance of learning materials. With the advances in pedagogy in e-learning, it was clear that learning was not achieved only by providing good learning resources to the learner [16]. First these objects should be relevant to the current learning objective of the learner, second, they also should be well suited for the learner context [4]. To be able to provide the learner with the right material, at the right time, in the right sequence and in the right format more information that the one traditionally available in metadata about the object is needed.

To reach this level of performance, an e-learning system should have information about the learner, the learning objectives, the learning context and a deep understanding of the qualities of the learning resources. This need is based on what usually a human instructor does in order to select materials for their students. In other words, Learners, Lessons, Context, Objects, and any other entity involved in the learning processes "are kings" in the sense that need to be taken into account to provide an effective e-learning system. Traditional LORs are not able to fulfill the requirements of these systems precisely because of their focus on Learning Objects.

To modernize LORs, this work proposes to include other entities as first-class citizens into the repository. These Learning Object Repositories will be no more just about Learning Objects, but about any entity involved in the desired learning process. Learner profiles, course and lesson structure, learning objectives, learning activities, evaluation resources and, of course, learning objects can be described, stored, searched, selected and retrieved with equal ease. The "Learning Object" part of the name of the repository should be replaced by just "Learning" to indicate that any entity involved in the learning process could be included. Moreover, all the entities in the learning process are interlinked: the instructor select the learning object that is published into a LMS as part of a course sequence and then used and rated by several learners. Learners could comment on the object and recommend it to fellow students. This new Learning Repository should be able to meaningfully express these relations and interactions between entities. The more flexible way to represent these relationships between entities is to use Semantic Technologies. The paradata described in the Learning Registry [8] will not be extra information about the resources, but the information that is generated during the learning process and is stored as the relationships between its different entities. These Semantic Learning Repositories can, in this way, manage all the information needed by modern e-learning systems to be able to recommend and personalize diverse learning processes.

The concept of Semantic Learning Repositories solves the LOR shortcomings identified in the introduction of this work. First, if all needed information can be stored in a single repository, there is no need for the e-learning system to include or connect with other types of repositories. Different types of e-learning systems could include different description for the entities and even different entities depending on the learning process they are supporting. Second, the use of Semantic Technologies leads to a format-free repository. Any metadata standard could be used to describe the existing entities. Mapping between metadata standards or ad-hoc structures is greatly facilitated by the use of RDF triplets to store information. The interoperability issues are also reduced if the data is published as Open Linked Data [2]. In this way, it can be easily consumed by any other Semantic Learning Repository or e-learning application. Finally, changes in the metadata formats and/or stored entities can be easily incorporated into the Semantic Learning Repository without need to change its functionality. This is possible also thanks to the flexible nature of RDF triplets.
implementation that do not require a predefined schema for the data.

As mentioned in the Section "Related Research" this concept has the potential to be a better and more adequate solution for the implementation of modern e-learning systems. While it is not the first to suggest the storage of more information than just the Learning Object metadata or the use of semantic technologies to store that information, it is the first to model the LOR as a multi-entity semantic repository.

To study the soundness of the concept, the next Section will present the implementation of what could be considered a modern e-learning systems using the an underlying Semantic Learning Repository.

4. PERSONALIZED LEARNING PATH RECOMMENDER SYSTEM

One of the main claims of this work is that the Semantic Learning Repository facilitates the design and implementation of modern e-learning systems. To this this assertion a Personalized Learning Path Recommender system will be designed and implemented from scratch based on this concept. This section describes the different steps of this process for the reader evaluation.

4.1 High Level Description

The main purpose of the system will be to recommend customized learning paths to students based on their preferences. These learning paths will be based on lessons published by teachers. In order to specify the desired system an object-oriented description is proposed:

First, the entity Author (in this case, the teacher or instructor) defines the entity Lesson. Several Lessons could be assembled into an entity Course. A Lesson is no more than a sequence of entities Learning Objectives in a particular order. This Learning Objective contains a topic and a verb. For example, in the objective: "The student should be able to program nested loops", the topic is "nested loops" and the verb is "to use". The Author also define the entity Learning Activity that fulfill a Learning Objective. The Learning Activity could contain one or more entities Learning Objects. All Learning Objects are related to a topic. When the entity Student use the system to require a Lesson, an entity Adaptive Learning Path is created as a sequence of Learning Activities that fulfill the Learning Objectives of the Lesson. The Learning Activities are selected based on the characteristics of the Students. The characteristics of the students are captured first by a This description is represented by a graph in Figure 1.

Given that the Semantic Learning Repository is based on entities and their relationships, there is a one-to-one translation of the object-oriented description and the structure of the repository. The entities that need to be stored in the repository are: Author, Student, Course, Learning Objective, Adaptive Learning Path, Learning Activity and Learning Object.

4.2 Entity Specification

In the next step, each of the identified entities is described in more detail and their relationships with other entities are modeled. The more interesting are described in the following subsections. All these details (characteristics and links) can be, again, easily mapped into the implementation of the Semantic Learning Repository.

4.2.1 Student

Each Student should have basic information related to her account (user, password) and basic identification information (name, institution, etc.). The most important information for the system, however, is the description of her preferences. These preferences will be expressed as the learning styles levels described by Felder [5]. The main relations that the Student has with other entities are: Students rate, comment and recommend Learning Activities (Social Actions), Students start, follow and complete Adaptive Learning Paths, Students follow and complete Learning Activities. The preferences of the students are updated based on the Social Actions performed with different kind of Learning Activities.

4.2.2 Learning Object

Learning Objects are described using the desired fields presented in the Learning Object Metadata standard. Apart from this information, Learning Objects are related to one or more Topics.

4.2.3 Learning Activity
Learning Activities has information about its title and description. They can be followed and completed by a Student. One or more Learning Objects are included in the Learning Activity. The Learning Activity is related to a Learning Objective that they fulfill. This relation can be created by an Author or automatically by the system based on the topics to which their Learning Objects are related. Each Learning Activity has also a description of the learning styles to which is more suitable. This information can be added by the Author or automatically based on the technical format of the Learning Objects that it contains.

4.2.4 Adaptive Learning Path

An Adaptive Learning Path has only basic information on when it was created. The rest belong to relationships that maintain with several other entities. The Adaptive Learning Path belongs to a Student. It can be started, followed or completed by a Student. The Adaptive Learning Path contain a sequence of Learning Activities selected based on the comparison between the preferences of the user and the description of the learning styles favored by the activity.

4.3 Low-Level Repository Implementation

In order to implement the Semantic Learning Repository, the natural choice, given the internal structure of the data, is to use an RDF store. This type of repositories is able to store and retrieve RDF triplets. In this concrete example, 4Store [7] was used. To facilitate the implementation of the e-learning system in languages that do not support RDF natively, a transformation service between RDF and XML is provided.

The process to store and retrieve entities is briefly described for illustrative purposes. The repository receives a request to save an entity in XML format. This request should contain all the information required to specify the entity:

```
<course>
  <courseID>1003</courseID>
  <title>Some course</title>
  <description>A brief description of the course</description>
  <language>en</language>
  <author>xochoa</author>
</course>
```

The metadata is then translated into triplets:

<table>
<thead>
<tr>
<th>resource</th>
<th>property</th>
<th>value</th>
</tr>
</thead>
<tbody>
<tr>
<td>1003</td>
<td>title</td>
<td>Some course</td>
</tr>
<tr>
<td>1003</td>
<td>description</td>
<td>A brief description of the course</td>
</tr>
<tr>
<td>1003</td>
<td>language</td>
<td>en</td>
</tr>
<tr>
<td>1003</td>
<td>author</td>
<td>xochoa</td>
</tr>
</tbody>
</table>

When the repository received a query request, this is translated to SPARQL (SPARQL Protocol and RDF Query Language) [15]. The query is then executed in the RDF database. For example, to get the information of the course 1003:

```
SELECT ?p ?o FROM <http://domain/COURSE> WHERE 
{ <http://domain/COURSE/ID\#1003> ?p ?o }
```

The database response is then internally translated in a friendly response in XML format, equal to the one used for storing it.
4.4 System Components

Apart from the repository, the system should implement other components to provide the end user (learner) with the Learning Path recommendation. Not being the main focus of this work, these components are briefly described in the following subsections. The whole system architecture is presented in Figure 2.

4.4.1 Recommender Logic

The personalization of the Learning Paths is conducted by this component. When a Student wants to follow a Lesson, this component start to assemble a new Learning Path based on the Student’s preferences and the Learning Activities recommended learning styles. The Social Actions information is also used to refine the Learning Path when more than 2 objects fit the eligibility criteria. The component is implemented using Apache Mahout [14] to implement a collaborative-based filtering [6] that use the Social Actions information and a basic rule-base algorithm to conduct the selection of the Learning Activities based on the learning style information. The resulting hybrid recommender is then use to construct the final Adaptive Learning Path.

4.4.2 End-User Interface

A PHP-based Web application was created to provide a final user interface to Authors and Students. Authors are provided with tools to publish Courses, Lessons, Learning Objectives, Learning Activities and Learning Objects. Students are provided with a search facility to find relevant Courses or Lessons and to follow them. The system also provides the Students with tools to perform Social Actions such as rate, comment and recommend to a friend. This system is connected to the Recommender and the Semantic Learning Repository through Application Programmable Interfaces (API). An screenshot of this system is presented in Figure 3.

5. EVALUATION OF THE SYSTEM

To gain insight on the performance of the implemented Personalized Learning Path Recommender, two tests were conducted. The first test measured the quality of the recommendation of learning objects compared with the selection made by human tutors. The second measure the different in performance in students that used the recommender compared to the ones that did not use it. The setting of these experiments was the Programming Fundamentals course conducted in a mid-sized University during the years 2010 - 2011. This course provides an introduction to the concept of Programming in the C language and it should be the first programming course taken by students. This course is mandatory for all the students of Computer Science, Telecommunications and Telematics majors. Students from other majors can take the course as optional credits. Currently, 270 students are taking this course, divided in nine groups. Historically, this course present a large failure rate (>50%), a reason why it was selected as the impact of the intervention could be easily measured.

5.1 Recommender Evaluation

The objective of this experiment was to establish if the recommender system based on the Semantic Learning Repository could perform its job. This job is to recommend objects present in the repository to different kind of students in a similar way in which a human tutor would do it. The experiment consists in comparing the list of Learning Activities recommended by the hybrid algorithm with the ones suggested by an subject-expert subject. One professor of the Programming Fundamentals course received a list of 10 programming study topics. He had to select from the repository those Learning Activities related to each topic, ordering them differently for students that prefer Visual material and for students that prefer Verbal material. At the same time, the recommender component was used to recommend from the same pool of materials those that were in the same 10 topics and for artificially created students profiles with visual and verbal preferences.

To evaluate the similarity of the expert list and the one generated by the system two measurements were obtained: the precision and the Kendall Tau metric. The precision measure the percentage of the relevant documents selected by the expert that were also selected by the recommender algorithm. The Kendall Tau metric measures the difference in the order of the two lists. Results of this analysis could be seen in Table 5.1

<table>
<thead>
<tr>
<th>Expert</th>
<th>Recommender</th>
</tr>
</thead>
<tbody>
<tr>
<td>Visual</td>
<td>Visual</td>
</tr>
<tr>
<td>8</td>
<td>10</td>
</tr>
<tr>
<td>9</td>
<td>9</td>
</tr>
<tr>
<td>10</td>
<td>8</td>
</tr>
<tr>
<td>11</td>
<td>11</td>
</tr>
<tr>
<td>20</td>
<td>20</td>
</tr>
<tr>
<td>21</td>
<td>21</td>
</tr>
<tr>
<td>41</td>
<td>44</td>
</tr>
<tr>
<td>44</td>
<td>41</td>
</tr>
<tr>
<td>Precision</td>
<td>100%</td>
</tr>
<tr>
<td>Kendall Tau Distance</td>
<td>0.18</td>
</tr>
</tbody>
</table>

Table 1: Recommender evaluation results

The results suggests that the performance of the evaluation (100% precision and 0.18-0.39 Kendall Tau distance) is very good compared with the state of the art of learning Object Recommendation [21]. This result validate the claim that viable modern e-learning systems could be easily implemented on top of a Semantic Learning Repository.

5.2 Learning Impact Evaluation

The objective of this experiment was to have a real-world estimation of the impact that systems based on the Semantic Learning Repository could have in the learning performance of students. Two sections (67 students) of the course of Programming Fundamentals at a mid-sized University were involved in this experiment. This intervention took place during the second semester of the academic year 2013-2014. To identify the impact that the use of the recommender system had in their performance, the most problematic Lesson identified by professors of the subject was selected. This Lesson was “Know how conditionals work”. During a 3-question test conducted during the previous year, only 1% of the students got 2 correct answers, while 97% were able only to obtain 1 point. During the 2 weeks corresponding to the study of that concept during the course, the students were pointed to the corresponding lesson in the Personalized Learning Path Recommender. After the 2 week period, the same questions asked during the previous-year test were presented to the
Sixty-seven students responded to the test to measure their performance in the Lesson “How conditionals work”. As it can be seen in Figure 4, most of the students were able to solve the test (right side) in contrast to almost none in the base-line group (left side). Applying a Welch two sample t-test, it was found that there is an statistically significant improvement in the pilot group. From these results is clear that using the system had an impact on the learning of the students. While the Semantic Learning Repository is not directly responsible for these results, a system that was easily created on top of this repository has the potential to be useful in the real world.

5.3 Evaluation Conclusion

The evaluations made to the Personalized Learning Path Recommender system provide an indication that effective modern e-learning systems could be easily built on top of the Semantic Learning Repository. The characteristics of this new repository: multi-entity, metadata-format agnostic and relation-capture ready made it easy to implement advanced recommender algorithms that could use a variety of sources of information without the need to connect to diverse non-compatible repositories.

6. CONCLUSIONS

Current Learning Object Repositories (LOR) have serious shortcomings that reduce their usefulness to implement e-learning systems that go beyond retrieving Learning Objects. Even recent advances, such as the inclusion of para-data and contextual metadata or the use of semantic information to lower their dependence on metadata standards only alleviate part of the problem. Counter-intuitive as it may appear, the only way to improve the usefulness of Learning Object Repositories is to reduce its focus on Learning Objects. As more e-learning systems are learner-centered instead of content-centered, the Learning Object Repository should adapt to include the information about the learner, the learning and the learning context as first-class citizens inside its structure.

The concept of a Semantic Learning Repository that replace the role of the LOR as part of e-learning solutions is presented and described in this work. This concept integrates the ideas of multi-entity storage (storing information not only about learning objects but all the aspects involved in the learning process) with the idea of semantic links between those entities. The combination of these two ideas not only eliminate current LOR shortcomings, but also facilitate the design of e-learning solutions based on top of this new kind of repository.
To illustrate the process of the design and implementation of a modern e-learning system on top of the Semantic Learning Repository, this paper summarize the process from high-level description to low-level implementation. Given the more natural entity-relation based design that a semantic repository provides, the transitions between design and implementation is effortless. The semantic capabilities make very easy to build complex algorithms that exploit the relationships between entities.

To validate the adequate performance of the built system, it was evaluated in two test, one oriented to the technical performance, the other to the overall impact of the system. The result of both test suggest that the built system was up-to-par with existing, more complex, e-learning solutions.

The main conclusion of this work is that simple changes on the 25 years old idea of Learning Object Repository could lead to great improvements on its relevance in the field of Technology Enhanced Learning.

7. ACKNOWLEDGMENTS

The authors of this paper want to thanks the support received from the The authors want to acknowledge the support of the VLIR-UOS project ZEIN2010RIP09 and the IGUAL (Innovation for Equality in Latin American Universities) Project funded by the ALFA III Program of the European Commission.

8. REFERENCES


