Semantic Enhancement of Learning Object Repositories

Félix Fernández-Peña[1] Yariel Ramos-Moreno[2]
Yudit Ponce-Toste[1]

[1]Politechnical University of Havana, La Havana, Cuba.
[2]Coremain S.L, Santiago de Compostela, Spain.
felix@ceis.cujae.edu.cu, yponce@hispavista.com, yariel.ramos@coremain.com

Abstract

Trying to combine the easy-to-use tag-based annotations and the accurateness of structured formal annotations, a semantic enhancement of learning object repositories is proposed in this paper. Its potential impact to accessibility in the user experience is evaluated whilst similarities and differences against related work are discussed.

Keywords: learning object repository, semantic annotation, user experience enhancement.

1. Introduction

Group-based asynchronous e-learning refers to situations where groups of learners are working over an Intranet or the Internet where exchanges among participants occur with a time delay [1]. This scenario is been impacted by the growing development of social networks and related technologies [2].

The increasing availability of open data potentiates the creation of resource repositories [3]. Integrating learning objects (LO) into repositories is in fact a major advance for e-learning initiatives. Semantic Web technologies together with ontologies provide a rich environment for facilitating e-learning via the semantic annotated learning objects and shared repositories. Pathmeswaran and Vian proposed the use of enabling technologies for developing next-generation learning object repository in such a technological integration [4]. In the recent past, several Learning Object Repositories (LOR) have been delivered for free. Among others in a growing list, there are MERLOT (Multimedia Educational Resource for Learning and Online Teaching) and CAREO [1].

However, as users become more knowledgeable and comfortable with the use of LORs, they are demanding advanced features and functionalities, including better collaborative learning tools, and better content management capabilities. The next-generation of LORs will have to have improved its accessibility, customizability and flexibility [1]. An example of this is DELPHOS, a framework to assist users in the search for learning objects in repositories. It is based on a weighted hybrid recommender approach that uses different filtering or recommendation criteria, as Zapata indicates [5]. Cechinel has been focused in collaborative filtering used by recommendation systems to predict the utility of items for users. Different algorithms are proposed by Cechinel, based on the similarity among the preferences of users [6].

With a growing knowledge base, the demand arises to access resources in more structured ways that classical repositories do not support. Enriched metadata is crucial for depicting the content of each LO in the context it is used [7]. The flexibility of linking learning pieces in a semantic dimension allows readapting LOs to changing scenarios and to the actual fallibility of knowledge [8].

Semantic repositories promise to capture structured knowledge offering advanced querying capabilities [9]. However, the semantic web applications take off is being slower than expected [10]. In this paper, we propose the semantic enhancement of LORs by combining the easy-to-use tag-based annotations and the accurateness of structured formal annotations.

Our objective is to evaluate the potential impact of such enhancement into the accessibility of LORs. As a use case, the collaborative distributed system for retrieving educative documents, proposed by Fernández et. al. [11], has been considered as the LOR to enhance.

2. Semantic enhancement of LORs

We modeled LOs as a triple representing its integration to the repository:

$$\mathbf{LO}(C, M, S) \qquad (1)$$

where *C* stands for the LO content, *M* is the metadata describing the LO and *S* is a set of statements featuring the LO in the semantic dimension by conforming a domain ontology. The proposed enhancement of LORs focuses on the set *S* and its relationship with *C* and *M*. Our goal is to simplify the content adaptation of LOs by using a different interaction model for advancing accessibility.

2.1. Structured formal annotation

It has been proved that the formal description of a learning scenario helps to contextualize learning assets [12]. By defining a *domain ontology* and linking *C* and *M* to terms of this ontology, the e-learning environment in which learners and facilitators interact is enriched.

The fact is that valuable information is not easily harvested under the present standards for describing LOs [13]. Learning Object Metadata (LOM) and Dublin Core (DC), as outer-metadata, are not enough [12]. Rodríguez et. al. proposed to standardize specific inner-metadata taking into account the nature of the LO (text, animation, graphic, etc) [14]. The proposed enrichment of LOs simplifies the tasks of content adaptation, learner interaction trace capture, learner modeling, and adaptive navigation.

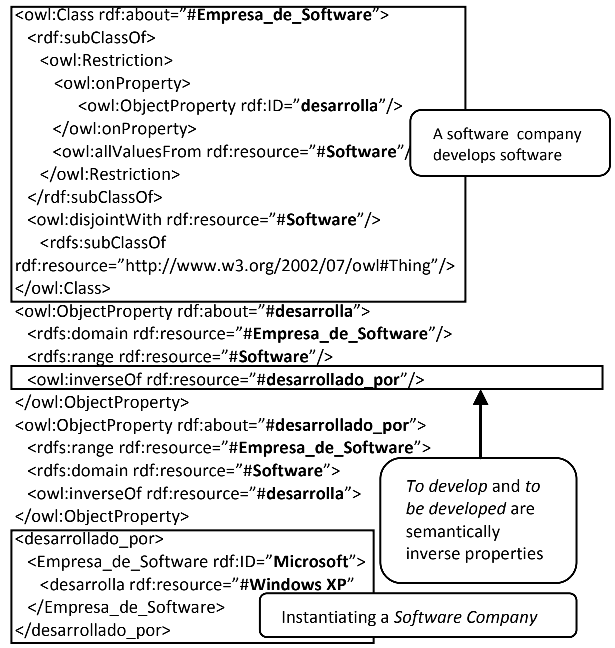
Later, Guzmán et. al. integrate outer-metadata with the description of learners and facilitators using the *Friend of a Friend* (FOAF) vocabulary, making easier the interoperability of LORs [13]. Actually, using FOAF for linking social networks has been validated by Golbeck in more general scenarios [15].

Other proposals simplify the structured data used for semantic annotation. Matthes et. al. have improved traditional wikis with the introduction of the so called hybrid wikis [9]. They modeled the dynamic enrichment of wiki pages by using a fixed database structure. The proposal impacts the user experience but once again metadata is “transparent” for software.

As a special syntax has to be used to add semantic annotations, editing structured content could be a difficult and cumbersome task with existing tools [9][10]. Nevertheless, instead of lowering the semantic enrichment by using a fixed schema (like in hybrid wikis), we decided to make e-learning facilitators responsible of the *domain ontology* updates. Otherwise, interoperability and reasoning capabilities would be affected [16].

Making facilitators responsible of the process is a natural decision taking into account that facilitators gather a global vision of the scenario and they are used to the learning process quality supervision. Their experience with concept maps makes easier for them to formalize concepts and to describe existing relationships among them in the graph structure of a *domain ontology*. By taking into account the learners’ interpretation of learning assets, a shared vocabulary has been defined in several specific scenarios of the use case.

Figure 1. Specific domain formalization example in OWL



The use of *upper ontologies* like EuroWordnet [17] is not discarded. However, our experience has proved that their general scope makes it hard to use them for modeling specific descriptions of LOs, as it is the case. The lack of availability of refined versions of upper ontologies in Spanish made them even more difficult to use in current scenario. Figure 1 shows a segment of a *domain ontology* created for the use case, encoded in OWL [18].

2.2. Custom content annotation

Due to the fallibility of knowledge, we recommend the learners annotation of LOs as feedback to facilitators. Flexible tag-based annotation mechanisms are easy to implement and easy to use, increasing collaboration and reinforcing social skills in learners [19].

However, freely created tags, not associated with a semantic formalism, are not naturally suited for collaborative processes due to linguistic and grammatical variations, as well as human typing errors [19]. Kim et. al. defined a tag ontology intended to be used in the representation of tagging data. The *tag ontology* describes the definition of a tag; it allows to establish the existing correspondence of tags in distributed environments; nevertheless it does not state how to describe the tag definition in a specific context (as we do in section 2.1).

Our concern with the accessibility issue made us propose the use of free-tag annotation techniques but supported by terms-recommendation techniques and auto-completing text-edit controls. This is similar to [9] in terms of the design of software components. The difference is that we encourage the use of terms conceptualized in the *domain ontology* over terms used just by specific users. Personal terms are differentiated in their look in the interface, actually discouraging their use.

Further, if a user decides not to use a term from the shared vocabulary but one of his own, s/he is able to link the term to equivalent concepts of the *domain ontology* (assuming that an equivalent relationship exists among term and concepts).

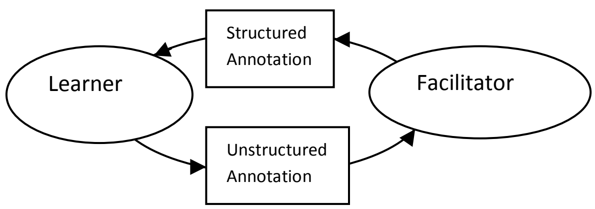


Figure 2. Learning object annotation lyfe cycle

2.3. Learning object annotations merge

As shown in figure 2, the interaction of learners with the repository is used as feedback for facilitators. Hence, "their actual interpretation of real world" is taken into account for a frequently refinement of the *domain ontology*. A closed life cycle for annotating LOs is achieved.

As a result, semi-structured and structured formal annotations of LOs are merged into a novel model, following Gruber vision of collective knowledge systems [8]. Existing relationships among components in the resulting *domain ontology* are illustrated in figure 3.

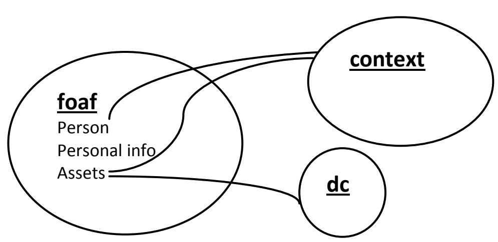


Figure 3. Modeling structural formal annotation of LOs

The formal description of the *domain ontology* supports the interoperability of repositories [13]. The overall impact over accessibility is evaluated next.

3. Accessibility improvement assessment

Learning is viewed as an active, constructive process whereby the learner strategically manages the available cognitive resources to create new knowledge by extracting information from the environment and integrating it with information stored in memory [1].

The improvement of accessibility is evaluated from three different points of view: the processing of user data and behavior (section 3.1), the deployment of novel mechanisms for navigation and content recovery (section 3.2) and the extension of search capabilities (section 3.3).

3.1. User data and behavior

Users can share their semantically enriched personal annotations of LOs with others. Restricted by the policy of to whom this information is available, and based on annotations, the system may suggest which users have a similar profile or the same interests.

Using FOAF as a normalized vocabulary for describing human relations makes easier to complement available data in the repository with open data available on Internet.

Related to user behavior, once semantic annotation has been normalized, learner interaction tracing processes are enriched in the semantic dimension of language. This enrichment allows seeing a primitive interaction event as a LO-specific semantic event interpreted in context [14].

We propose to data-mine the storage of user interaction semantic trace in order to identify patterns. This meta-information allows facilitators to validate the correct integration of LOs in an academic course and to refine the metrics and rules that decide when assets are related or not.

3.2. Navigation and content recovery

Contrary to the opinion of Matthes [9], we consider that translating the expressivity of classic established semantic web formalisms (like RDF or OWL) and the user by means of new user interfaces is actually possible. In fact, García et. al. have integrated their *Rhizomer* platform combined with a wiki in order to create a semantic portal [20].

The object-action interaction paradigm, applied in the *Rhizomer* Semantic Content Management System [21], seems natural to take advantage of the semantic enhancement of LORs towards a better user experience. García et. al. generates HTML table-based views using an XSL transformation of an appropriate ontology fragment [21] [10].

We propose to use a similar mechanism for generating semantic views. Nevertheless, instead of an automatic fragmentation of the ontology, the backend for facilitators provide appropriate tools for depicting subsets of the ontology as semantic views. As a result, an XML file is generated. The document type definition of this XML was specifically designed to be able to generate a tree-based HTML/JavaScript GUI for learners interacting with the LOR.

Figure 4 Outlook of a sample semantic view for navigating in the LOR.

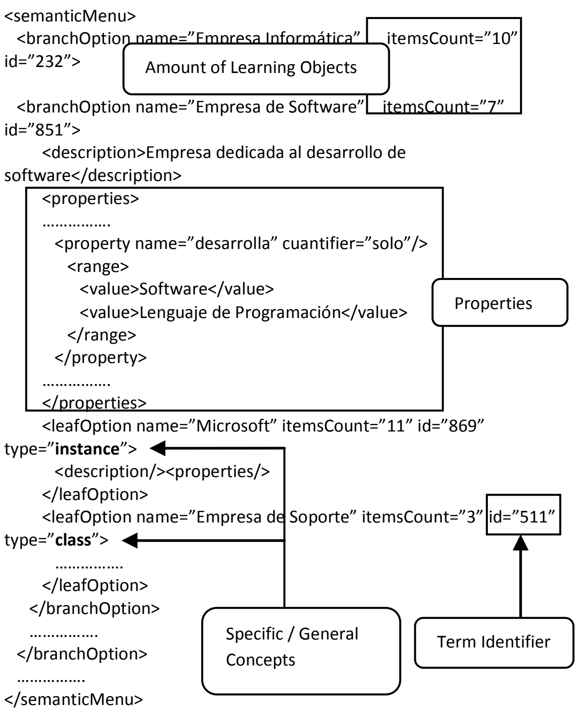


Figure 5. A segment of a semantic view XML encoding.

By using Web 2.0 available technologies, tree branches content is retrieved on demand (minimizing network traffic) whilst the dynamic rendering of the interface allows to ask the LOR for the amount of available LOs per category. When the user requests further information (by positioning the mouse over a concept) an asynchronous request is sent to the server and, in an overlay layer, properties and conceptual definition (if available) is shown. Figure 4 illustrates these functionalities in action whilst figure 5 shows a segment of the corresponding semantic view XML file.

3.3. Learning objects search

People looking for available assets on a local repository are largely limited to keyword search and categories filtering [22]. Meanwhile, more elaborated facets-based interfaces, as proposed by Smith [22], do not avoid the need of knowing what you want in order to look for it.

An advantage that local repositories currently have over massive search services is local or community-based knowledge. The proposed semantic enhancement leverages this advantage for impacting the accessibility on LOs search as well.

In order to demonstrate so, a search user interface was designed by using three select HTML components. Each component is related to properties, domains and objects, respectively, assuming the first order logic-based semantic conceptualization of the *domain ontology*. The resulting view is shown in figure 6.

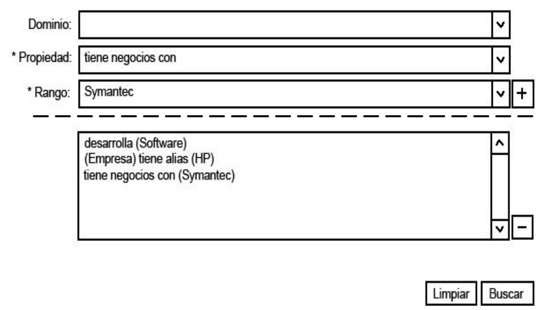


Figure 6. Semantic seach implementation.

An API was implemented in order to feed those components. It means that when a user clicks on one of those controls, associated terms are retrieved from the structural formal annotation model. Once the user has chosen a value, possible values for the other two controls are recalculated. At the end, when the user sets a value for at least one of the components, a semantic statement is ready for narrowing the search output.

As figure 6 shows, existential and universal logic operators are applied when specifying more than one statement. The results of reasoning over the *domain ontology* can be improved depending on the software capabilities. For demonstration purposes, Jena library [23] was used for deciding which LOs to retrieve. However, further research in this area is in progress.

4. Conclusions

By combining tag-based and structured formal annotations, learners collaborate and reinforce social skills. Learners' interaction with the repository is used as feedback to facilitators for structuring knowledge.

The proposed semantic enhancement of LORs combines structured and semi-structured knowledge representation techniques, achieving a balance among accurateness and easy-to-use user interfaces. The user is beneficiated with the improvement of LORs accessibility.

5. References

[1] S. Naidu, *E-Learning. A Guidebook of Principles, Procedures and Practices*, CEMCA, India, 2006.

[2] M. Litzler, J. García, A. Halbach, "Creation and Use of a Learning Object Repository for EFL", Procedia - Social and Behavioral Sciences , Volume 46 , pp. 5905–5909, Barcelona, Spain, February 2012.

[3] L. Carr, A. Swan, and S. Harnad, "Creación y mantenimiento del conocimiento compartido: contribución de la University of Southampton", *El profesional de la información*, 2011, January-February, v.20, n.1, pp. 101-110.

[4] R. Pathmeswaran, A. Vian, "Enabling technologies for developing next-generation learning object repository for construction", Automation in Construction, Conference in Planning Future Cities,Volume 22, pp. 247–257, March 2012.

[5] A. Zapata, V.H. Menéndez, M.E. Prieto, C. Romero, "A framework for recommendation in learning object repositories: An example of application in civil engineering", Advances in Engineering Software ,Volume 56, pp. 1–14, February 2013.

[6] C. Cechinel, M. Sicilia, S. Sánchez , E. García, "Evaluating collaborative filtering recommendations inside large learning object repositories", Information Processing & Management , Volume 49, Issue 1, pp. 34–50, January 2013.

[7] S. Sánchez, M. Sicilia, E. García, and C. Pagés, "Social models in open learning object repositories: A simulation approach for sustainable collections ", Simulation Modelling Practice and Theory , Volume 19, Issue 1, pp. 110–120, January 2011.

[8] T. Gruber, "Collective Knowledge Systems: Where the Social Web meets the Semantic Web". *Web Semantics: Science, Services and Agents on the World Wide Web*, 2007.

[9] F. Matthes, Ch. Neubert, and A. Steinhoff, "Hybrid Wikis: Empowering Users to Collaboratively Structure Information", *6th Conference on Software Paradigm Trends*, Sevilla, 2011.

[10] R. García, J. Gimeno, F. Perdrix, R. Gil, M. Oliva, J. López, A. Pascual, and M. Sendín, "Building a Usable and Accessible Semantic Web Interaction Platform", *Database Management Information Retrieval Journal*, Vol. 13, No. 1-2, pp. 143-167. Springer, 2010.

[11] F. Fernández-Peña, S. Willmott, A. Rosete. "Un Modelo de Sistema Colaborativo Supervisado para la Gestión Documental". *Revista de Investigación Operacional*. Vol. 29. Núm. 1, 2008.

[12] F. Fernández-Peña, N. Pina-Riesco, and Y. Ramos-Moreno, "Procesos de Enseñanza basados en la Descripción Semántica Formal y Explícita de Documentación Educativa", *I Simposio Internacional de Documentación Educativa*, Febrero, 2007.

[13] J. Guzmán, R. Motz, and A. Rodrigues, "Hacia la Publicación Abierta de Objetos de Aprendizaje". *Cadernos de informática*, May, v. 6, n. 1, 2011.

[14] O. Rodríguez, S. Chen, H. Shi, and Y. Shang, "Open Learning Objects: the case for inner metadata", Eleventh International World Wide Web Conference, Hawaii, 2002.

[15] J. Golbeck, M. Rothstein, "Linking Social Networks on the Web with FOAF", *Twenty-Third Conference on Artificial Intelligence*, 2008.

[16] H. Mason, *Integrating institutional repositories into the semantic web*, Ph. D. Thesis. University of Southampton, School of Electronics and Computer Science, June, 2008.

[17] EuroWorNet, "Building a multilingual database with wordnets for several European Languages", http://www.illc.uva.nl/EuroWordNet/

[18] W3C, Web Ontology Language (OWL), www.w3.org/2004/OWL/

[19] H. Kim, S. Scerri, J. Breslin, S. Decker, and H. Kim, "The State of the Art in Tag Ontologies: A Semantic Model for Tagging and Folksonomies", *International Conference on Dublin Core and Metadata Applications*, pp. 128-137, 2008.

[20] R. García, R. Gil, "Building a Semantic IntraWeb with Rhizomer and a Wiki", *IntraWebs Workshop, World Wide Web Conference*, Edinburgh, UK, 2006.

[21] R. García, J. Gimeno, F. Perdrix, R. Gil, and M. Oliva, "The Rhizomer Semantic Content Management System", *1st World Summit on the Knowledge Society*, Lecture Notes in Artificial Intelligence, Vol. 5288, pp. 385-394, Springer, 2008.

[22] D. Smith, J. Lambert, and M. Schraefel, "Rich Tags: Cross-Repository Browsing", *Open Repositories Conference 2008*, 1st-4th April, Southampton, United Kingdom, 2008.

[23] Ph. McCarthy, "Introduction to Jena", IBM DeveloperWorks, http://www.ibm.com/developerworks/xml/library/j-jena/.