

Semantic Web Aided Authoring of Educational Material¹

Carlos Delgado Kloos¹, Miguel Rodríguez Artacho²

¹ Universidad Carlos III de Madrid, Av. Universidad, 30, 28911 Leganés (Madrid), Spain, <cdk@it.uc3m.es>
² Univ. Nacional de Educación a Distancia, C/Juan del Rosal, 16, 28040 Madrid, Spain, <miguel@lsi.uned.es>

Abstract. A lot of attention has been drawn upon the teaching and learning process using web technologies. Also the interplay between e-learning and Semantic Web has been studied. But less attention has been devoted to the application of Semantic Web techniques to the authoring process of educational material. We present some ideas of applying semantic web technology for authoring e-learning material in order to harness the wealth of existing content. We present some ideas with the use of educational ontologies, annotation with RDF, and reasoning with CWM in the context of distributed repositories of educational material. The use of inference rules to automatically generate metadata provides an important assistance for the authoring and reuse of educational material.

1 Introduction

The creation of learning content has always been a big effort, because often the goal of authors is to develop a complete course. If the main goal of a learning object (LO) is for teaching and learning, the second goal should be its reuse. Extensive research has been carried out lately to standardize learning content components, to make them usable in interoperable and maintainable content repositories. To organize and help in the retrieval of LOs, metadata labels have been defined and standardized. But this has introduced an additional burden, namely that of annotating LOs following these metadata.

The killer application of the Internet is the people. It is with people and some organization that impressive results are achieved at low cost (eg. del.icio.us, Wikipedia, Flickr, or the syndication of blogs). Wouldn't it be nice to draw upon from the vast collection of educational material that exists on the web to easily create new courses? Several repository efforts have been launched, but they have not taken off. Often the problem is one of granularity: the presentation of the learning resources is too broad as to ease the reuse of specific content. But if the granularity is too small, it is more difficult to have a good overview. Dicheva et al. [4] propose the creation of a semantic layer based on the conceptualization of a subject domain. They suggest the use of Topic Maps to organize information in the repositories. Our approach is similar in the respect of creating a semantic layer, except that we use RDF, which provides a simpler, more flexible, and less specific model than Topic Maps. The possibility of reasoning with tools such as CWM [1] provides us with a tremendous power, allowing for the inference of new metadata from existing ones by the use of rules, so partially relieving from the burden of manual generation.

2 Supporting the Authoring of Learning Content

The authoring of learning content is similar to the COTS software building model, as it combines creation from scratch and reuse and modification of existing content, available in repositories or on the net. The reuse process requires a meaningful way to search and retrieve the appropriate content shaped as LOs. In our opinion, current approaches based on an available classification of LOs by means of metadata labels lack a way to search, retrieve, and reuse LOs from an instructional perspective. The reason for this drawback is based on the fact that (1) from an instructional perspective, retrieval of LOs by matching metadata attributes does not replicate the way a teacher operates when creating material, and (2) from the authoring perspective, the construction of metadata is very costly and not as precise and consistent as desirable.

¹ The first author acknowledges support from project MOSAIC (TSI-2005-08225-C07) of the Ministerio de Educación y Ciencia. He also thanks the Secr. Estado de Univ. e Inv. (MEC) of Spain for support during his sabbatical stay and MIT for hosting him during this year. The second author wishes to thank the SIG-AIED group in the KALEIDOSKOPE NoE and the *Programa Propio de Investigación* of UNED.

Concerning the first point, we have already explored in [6] how to access topic maps from a learning content specification to create learning material. Thus from the teachers' point of view, authoring of educational content is created using instructional queries like i.e. *'show examples to illustrate concept C'* or *'exercise concept C solving this problem'* or *'give me a hint'*. We can be even more precise by querying *'insert some easy examples to illustrate this concept'* etc. In the current way of retrieving LOs, we look for terms that match the metadata attributes, but without asking for a precise instructional relationship. It has been outlined also in [6] that authoring can be facilitated using these mechanisms from an educational modelling language, combining both abstraction and reusability.

Current approaches of educational modeling languages reflect the evolution of learning content authoring from a clip-art model based on a linear aggregation of LOs to a complex description of units of learning [5]. Some proposals differ in the instructional approach, regarding whether the content description paradigm is activity-driven (i.e. EML or IMS-LD) or structure-driven (i.e. PALO), or whether activities are individual (PALO) or collaborative (IMS-LD), but all of them provide mechanisms to integrate or refer to reusable components. In this sense, the use of external LO repositories structured according to an ontology can facilitate the way to create semantically labeled material. Moreover, with the mechanisms of PALO, preserving the instructional relationships and meaning of the retrieved LOs is straightforward by means of an equivalent RDF description on the resulting HTML page. Inference and reasoning over these metadata is possible using a shared ontology and the semantic labels of the educational web content.

3 Reasoning and Keeping the Overview

CWM is a general-purpose data processor for the semantic web with RDF as its core language, either in its XML version or in Notation3 [2]. This core language has been extended to include rules, allowing to reason for querying, checking, transforming and filtering information. The need to annotate with metadata is thus drastically reduced, since rules can infer new metadata information from the one existing. The problem of having to annotate LOs with metadata receives with this reasoning capability a big relief.

It is important not to lose the overview about what material has been finalized, what is still temporary and what is missing at all. Several tools can be used to help in keeping this overview. Tools like the Tabulator [3] can help visualizing the status of the work, in a similar way as iTunes helps to keep track of the music library. It is also possible to interface with graphical tools to show a representation as a graph. This allows to construct cooperatively educational material in a closed domain (in the realm of a particular department or group), or even better, in an open context of freely accessible LOs. When the author finds an LO that is useful for his/hers course, he/she can incorporate it (1) without changing the object, (2) without changing the already existing LOs, and (3) without having to worry about all the metadata that have to decorate it. Authoring should be a collaborative, distributed activity. It is wise to follow the following principles

- have LOs distributed
- keep them as neutral as possible (with respect to presentation, etc.)
- annotate them externally with RDF, including instructional information and relation of elements
- try to infer as many metadata as possible using rules
- the author should maintain the overview with graph representation tools, to see what is missing

References

1. Berners-Lee, T.: Cwm. <http://www.w3.org/2000/10/swap/doc/cwm> (checked March 2006)
2. Berners-Lee, T. (ed.): Notation3. <http://www.w3.org/DesignIssues/Notation3.html> (checked March 2006)
3. Berners-Lee, T. (ed.): Tabulator. <http://www.w3.org/2005/ajar/tab> (checked March 2006)
4. Dicheva, D., Dichev, C., Sun, Y., Nao, S.: Authoring Topic Maps-based Digital Course Libraries. Workshop on Applications of Semantic Web Technologies for E-Learning, Eindhoven, The Netherlands, 23-25 August (2004)
5. Koper, R., Rodríguez-Artacho, M., Rawlings, A., Lefrere, P., van Rosmalen, P. (2002) "Survey of Educational Modeling Languages" Technical Report of the CEN/ISSS Learning Technologies Workshop
6. Rodríguez-Artacho, M., Verdejo, M.F.: Modeling Educational Content: The Cognitive Approach of the PALO Language. In: Journal of Educational Technology & Society, Vol. 7 # 3 (2004)