Anita Petrinjak

Rodger Graham

Author
Anita Petrinjak, M.Sc., (anita.petrinjak@ualberta.net) is a Software Architect at Challenger Geomatics, Edmonton, AB, Canada. Correspondence concerning this article should be addressed to: Anita Petrinjak, Challenger Geomatics, 1400 10117 Jasper Avenue, Edmonton, AB, Canada, T5J 1W8. Rodger Graham (rodgerg@athabascau.ca) is an Instructional Media Analyst in the department of Educational Media Development at Athabasca University, Canada. Correspondence concerning this article can also be addressed to: Rodger Graham, Educational Media Development, Athabasca University, 1 University Drive, Athabasca, Alberta, Canada, T9S 3A3.

Abstract

This paper describes work that was done at Athabasca University as part of the EduSource Canada project. This work centered around learning object development based on pre-authored educational content. The major outcomes of the work were the development of an explicit semantic structure with strong educational focus for learning objects, and the implementation of that structure, using platform/software-independent XML technology. An explicit semantic structure for educational content has some significant advantages: it enables faster publishing of material in different formats using automated processes; it allows institutions to participate in seamless content exchange with other institutions; and it enables more accurate discovery and reuse of learning objects within learning object repositories.

Résumé: L'article est axé sur la description du travail exécuté à l'université Athabasca dans le cadre du projet EduSource Canada. Ce travail est le fruit de l'élaboration d'objets d'apprentissage basés sur du matériel éducatif existant. Nous avons élaboré une structure sémantique explicite en mettant un accent éducatif important sur les objets d'apprentissage et l'avons mis en oeuvre à l'aide d'une technologie XML indépendante des plate-formes et des logiciels. Une structure sémantique explicite du contenu éducatif comporte de nombreux avantages par rapport aux méthodes traditionnelles

car elle permet une publication plus rapide du matériel sous différents formats grâce à des processus automatisés. De plus, elle permet aux institutions d'effectuer des échanges continus de contenus avec d'autres institutions.

Introduction

As educational institutions, especially those specializing in open and distance learning, move toward digital storage and electronic delivery of course materials, there is a growing need for more sophisticated content management. Ideally, any given document should have only one master copy that can be published in a variety of formats (online, print, to a handheld device, etc.). This document should be constructed in such a way that different parts of the content can be reused for a variety of purposes, and it should use widely accepted standards for content markup and storage that guarantee that the content can be shared both within an organization and among institutions.

Developing a document management system that fulfils these criteria is a goal of the EduSource Canada project (McGreal, Anderson, Friesen, Sosteric, Hewitt, Ring et al., 2004) and it is mirrored by other efforts being carried out around the world. Considerable progress has been made in developing a common framework (e.g., Anderson & Downes, 2000). Much of the work being done now centers around two concepts: learning objects and learning object metadata.

Learning objects, while much discussed, have been defined only loosely, and as yet no single common definition has been agreed (Wiley, 2000; McGreal, 2004). For example, the Learning Technology Standards Committee (LTSC) definition gives only very general guidelines:

> Learning Objects are defined here as any entity, digital or non-digital, which can be used, re-used or referenced during technology supported learning. Examples of technology supported learning include computer-based training systems, interactive learning environments, intelligent computer-aided instruction systems, distance learning systems, and collaborative learning environments. Examples of Learning Objects include multimedia content, instructional content, learning objectives, instructional software and software tools, and persons, organizations, or events referenced during technology supported learning. (Institute of Electrical and Electronics Engineers [IEEE] Learning Technology Standards Committee [LTSC], Learning Object Metadata Working Group [LOMWG], 2000)

Based on the above definition, it is hard to draw a pragmatic line between learning objects and entities that are not learning objects. McGreal (2004) looks at this and other definitions and proposes a "practical" definition of learning object: "any reusable digital resource that is encapsulated in a lesson or assemblage of lessons grouped in units, modules, courses, and even programmes" (36). We adopted McGreal's definition in our work, and we focused primarily on a subclass of learning objects—textual learning objects.

While the definition of a learning object is still being debated, the metadata used to describe learning objects has received more attention. There are widely accepted standards for defining metadata in detail; for example, the Dublin Core Metadata Element Set (Dublin Core Metadata Initiative [DCMI], 2003) and the IEEE Learning Object Metadata IEEE Learning Technology Standards Committee [IEEE LTSC LOMWG, 2002). Furthermore,

it is widely recognized that learning objects, in and of themselves, no matter how well they are described by metadata, cannot simply be slapped together like Lego bricks to produce learning events. Koper (2001) describes a formal approach to ensuring pedagogically sound use of learning objects in his discussion of an educational modeling language (EML) used to describe the learning process workflow including roles and activities of students and teachers. The IMS Learning Design specification (IMS Global Learning Consortium, 2003) is based on EML and embodies this approach.

It would appear, then, that while there have been significant achievements in defining metadata and learning design, there has been considerably less progress in developing specifications for learning objects themselves. In our research effort, we have focused on the analysis, definition and design of the semantic structure for textual learning objects as the basis for educational content development. Clearer definitions and specifications applied to learning objects would facilitate their reuse and interoperability between repositories that store such objects.

Friesen (2004) describes the key problems with the existing approach to learning objects, such as the ambiguity of the concept itself, and a focus on technical rather than educational aspects. Our efforts attempt to address these issues by taking a more pedagogical orientation and defining the educationally relevant structure of the learning object. We hope that by providing well-defined and educationally focused learning object specifications, we can reduce the ambiguity surrounding the questions of what the objects are, and how they can be used. In doing so, we are aiming for a medium-neutral, educationally relevant and flexible way to define and represent learning objects.

It was immediately obvious, after articulating our goal, that the Extensible Markup Language (XML) was a natural answer to our requirements. XML is transformable into different formats for different publishing media and therefore medium-neutral (Bray, Paoli, Sperberg-McQueen, & Maler, 2000). Moreover, documents created with XML can be made educationally relevant through the definition of individual elements within the XML Schema, or Document Type Definition (DTD). Overall, XML offers a flexible yet powerful means for representing learning objects.

The following sections describe our proposed semantic structure for learning objects in more detail, and show how we used it as part of Athabasca University's EduSource Project, to create learning objects from pre-authored course materials.

Learning Object Schema

Part of our mandate in developing learning objects was to find ways to adapt current course materials to create learning objects. Athabasca University possesses a wealth of digitized content, most of which has already been designed for open learning and distance delivery, based on well-founded pedagogical principles. This fact gave us a jump start on content development. A survey of representative courses, coupled with a review of the approach to educational content modeling used by the Open Learning Agency (OLA) of British Columbia (Bartz, 2002), was instrumental in helping us develop and refine our

schema. Once we had identified the required elements for the schema, we were able to define the semantic structure of the learning objects and implement that structure as an XML schema. The semantic structure is illustrated in Table 1.

Table 1.
Learning Object Schema (LO Schema.

#	Name	Explanation	Requirements, Examples
1	learning_object	This is the top category that encompasses the whole learning object content.	
1.1	identifier	A globally unique identifier that identifies this learning object. Same as General Identifier in the (IEEE LOM) metadata record associated with this learning object.	Required.
1.1.1	catalog	The name of the cataloging scheme used for the corresponding identifier entry. "A namespace scheme" (IEEE, 2002)	Optional. e.g., "URI"
1.1.2	entry	The value of the identifier, based on the cataloging scheme from "catalog" category, that unambiguously identifies this learning object. "A namespace specific string."(IEEE, 2002)	Required. e.g. , "http://www.atha bascau.ca/LOS"
1.2	title	The title of the learning object content.	Optional.
1.3	introduction	Introduction text to the learning object content.	Optional.
1.4	learning_outcomes	Expected learning outcomes of deploying this learning object in the learning environment.	Optional.
1.5	prerequisites	Prerequisites needed for a learner to use the learning object, such as knowledge of a certain area or taking a certain course.	Optional.
1.6		This section denotes the repetition of choices among the following three categories.	
1.6.1	content	General text of the learning content.	Optional.
1.6.2	practice	Learning content for learners' practice.	Optional.
1.6.3	assessment	Learning content for learners' assessment.	Optional.
1.7	metadata_reference	This category references a globally unique Meta-metadata identifier of the associated metadata record.	Optional
1.7.1	catalog	The name of the cataloging scheme used for the corresponding identifier entry. "A namespace scheme" (IEEE, 2002)	Optional
1.7.2	entry	The value of the identifier, based on the cataloging scheme from "catalog" category, that unambiguously identifies the metadata record associated with this learning object. "A namespace specific string."(IEEE, 2002)	Optional.

The semantic categories presented in Table 1 define a very flexible structure for learning objects. All the sections are optional except for the identifier of the learning object. The title, introduction, learning outcomes, and prerequisites can appear, at most, once, while the content, assessment, and practice sections can appear any number of times and can be intermingled. As a result, this schema can be used for the creation of a wide variety of learning objects. For example, instructional text can be represented by the introduction and content category, possibly with prerequisites and learning outcomes; a quiz can be

developed using introduction, prerequisites, and assessment; and an annotated index could use just title and introduction.

Within each of the semantic sections, marked-up text can be inserted; the only requirement is that it be XML compliant. Depending on the requirements, various formats can be used. XHTML (Extensible HyperText Markup Language) can be used for formatting learning object content. XHTML is a new, extensible, version of HTML (**H**yper **T**ext **M**arkup **L**anguage) and is appropriate for materials whose primary purpose is to be published online. XHTML is written in XML, and as such is an XML application. XHTML can be transformed using an XSLT (E**x**tensible **S**tylesheet **L**anguage **T**ransformations) into any delivery format: Web, print, a handheld personal digital assistant (such as a Palm Pilot), or an automated Web reader for reading aloud.

This proposed structure and format of learning objects has some significant advantages over the existing situation, in which there are many different learning object formats, many of them proprietary and lacking much semantic structure. The proprietary formats of textual resources necessitate the purchase of appropriate software and foster dependence on a particular vendor, which tends to block the free flow of the learning objects. If educational institutions can agree on open standards, such as those proposed in this paper, learning objects could be exchanged easily within an institution and among different institutions, with less expense and fewer time-consuming adjustments to format. Currently, learning resource developers are preoccupied with a particular way of publishing (e.g., Web or print), and it is expensive and time consuming to use one master copy for different delivery formats. Our proposed structure and format can be transformed automatically using XSLT into different formats and it circumvents problems arising from multiple master copies of the same content. Moreover, the explicit semantic structure can be recognized and processed automatically by computer programs. In many cases this feature relieves the course author from repeated edits and adjustments of the content. For example, an "assessment" section within the schema includes self-tests, quizzes, and other forms of assessment. This section can be extracted automatically from the course learning object and randomized to create examinations. Another example is the automatic extraction of the "title," "prerequisites," and "introduction" sections from the object to create or update a course syllabus.

Overall, adoption of a schema such as that presented in this paper can reduce the expenses normally associated with traditional content management, while increasing the usability and reusability of the content itself.

Interoperability With Existing Standards and Developments

An overriding imperative in developing any schema is to adhere to standards whenever possible. Fortunately, new standards for many kinds of shareable content are emerging. The flexibility of our schema should allow the adoption of standards as they are developed. For example, the IMS Question and Test Interoperability standard (IMS QTI) is a standard that allows the sharing of content among various kinds of assessment tools, such as

multiple-choice, short-answer, fill-in-the-blank, and other quizzes. Similarly, the IMS Reusable Definition of Competency or Educational Objective (IMS RDCEO) has been developed to facilitate common understanding and exchange of competencies, learning pre-requisites and learning outcomes. We can easily accommodate IMS QTI and RDCEO by including the XML code within the <assessment> and <learning_outcomes> semantic sections. Appendix 1 illustrates how QTI and RDCEO may be used in our schema.

Moreover, our intention was to make the learning object schema interoperable with other standards, such as the Institute of Electrical and Electronics Engineers Learning Object Metadata standard (IEEE LOM) and the IMS Learning Design specifications (IMS LD). There is now a significant effort in many parts of the world to develop and standardize the infrastructure for online education, including the repository architecture and interoperability protocols (e.g., IMS Digital Repositories Interoperability, Sharable Content Object Reference Model), metadata standards (e.g., IEEE LOM, Dublin Core), and learning process environments and workflow specifications (e.g., IMS LD). The learning objects based on the LO Schema can be associated with their respective IEEE LOM metadata records. In the schema, we incorporated an identifier structure that is equivalent to the IEEE LOM identifier structure to enforce the compatibility with this internationally accepted standard.

Our approach is fully interoperable with existing standards and specifications. The schema defines a structure for the learning resource while relying on the above-mentioned standards to provide metadata records and learning process workflow that will use learning objects as modular, reusable, plug-in components.

Our work builds upon the work described by Bartz (2002), who discusses the Open Learning Agency's approach to developing a structured content model. The OLA developed a DTD (Document Type Definition) that defines the structure of a complete course, including metadata for the course and its sub-units. This is a valuable effort in introducing semantic structure into the learning content, but is a proprietary format that is neither modular nor fully compatible with the relevant educational standards such as IEEE LOM and IMS LD.

Authoring Learning Objects

The learning objects we created for this project were all the result of converting existing digital course materials that had been originally designed for distance delivery in print form. The documents resided in a proprietary content management system and required considerable editing once exported to remove extraneous markup code and to correct errors introduced by the export process. Export options from the proprietary system were limited to rich text format (RTF) or HTML. We chose to use the HTML export option because it meant that some of the mark up code we would need (mainly heading tags) was already present.

After cleaning up the files, we used XMLSpyÃ‚Â© to parse the documents into unit-sized files. The documents originally comprised anywhere from ten to twelve units. Each unit

covered an average of three topics and included learning objectives, an introduction, instruction, supporting graphics, and, in some cases, self assessment. We chose the unit as our standard of granularity because it afforded the most convenient and efficient method open to us to create a significant number of learning objects in the time we had available. In keeping with the scope of the EduSource project, we generated about 200 learning objects based on material from three separate courses, including supporting graphics and multimedia objects already in existence.

Learning objects based on our learning object schema can be created using any XML editor, such as XMLSpy or JEdit, but these editors require knowledge of XML and manual marking-up processes. To enable users who do not possess significant XML and mark-up knowledge to author semantically tagged learning objects, we have developed a prototype learning-object authoring tool. We have also modeled and developed the processes for converting existing, text-based course materials into semantically marked-up learning objects.

Figure 1 shows our learning object authoring tool. The left-hand side of the screen shows the original text, (the raw educational material in, for example, XHTML, IMS QTI or IMS RDCEO format), and the right-hand side of the screen shows the editable semantic sections of the learning object being created. Mark up buttons associated with each semantic section enable the user to add text and other elements.

Figure 1. Tagging tool for learning objects.

Different types of learning objects, such as narrative text on a particular topic, quizzes and practice exercises, can be developed using our schema. Appendix 2 shows an example of a learning object that we authored based on our learning object schema. Note that the object has been significantly shortened for inclusion in this paper.

The learning object in Appendix 2 contains mainly narrative text contained in the content sections, intermingled with practice sections for the purpose of enhancing students' understanding of the topic. The document also contains the following sections: title, introduction, and learning_outcomes. The complete text of this learning object could be published online or in print as a unit in an economics course. Furthermore, semantic tagging allows other automated processing of the document. Search engines can search the document based on semantic sections. Individual elements can be extracted and published separately, possibly producing an index or syllabus, by pulling out only the title and introduction sections. Or an instructor may wish to view only the learning outcomes to find an appropriate object.

Conclusion and Discussion

E-learning, once a futuristic vision, is becoming a real presence in everyday life. Educational institutions are building competitive frameworks for e-learning that will follow the demands of the market and exploit the power of technology. Standardization efforts

are trying to keep up with the growing need for interoperability among educationally oriented applications and technologies, but there are still areas that require more research. Due to the lack of definition for semantic structures of text-based learning objects, we propose an explicit semantic structure in combination with a platform-and-software-independent approach formatted in XML. An explicit semantic structure for educational content has a significant advantage compared to traditional approaches, because it enables faster publishing of material in different formats using an automated process. Another advantage is that institutions can participate in seamless content exchange with other institutions.

We used the semantic structure presented above for Athabasca University course content, and we used XSLT to transform it into different formats and structures for publication. This approach allows course authors to focus more on the pedagogical aspects of the material rather than technical issues around delivery. The semantic structure enables computer programs to take care of the content processing. For example, explicit semantics in XML can facilitate automatic transformation of content for display online or in print and for automatic assembly of a course syllabus based on the selected learning objects. Moreover, there is no concern that such an approach will limit the expressive powers of the learning material, since additional design can always be added to enrich the automatic results.

Some development is still needed before we can realize the benefits of explicit semantic structures. Transforming existing educational content into learning objects with a semantic structure requires a significant effort, and there is a need for authoring tools that an average content creator can use comfortably. Once these obstacles have been overcome we can begin to realize the full benefit of semantically structured learning objects.

References

Anderson, T., & Downes, S. (2000). Models and strategies towards a Canadian on-line educational infrastructure. *Report for Industry Canada, Advisory Committee for Online Learning*. Retrieved September 22, 2004, from http://mlg-gam.ic.gc.ca/sites/acol-ccael/en/resources/R01_Anderson_Downes/

Bartz, J. (2002). Great idea, but how do I do it? A practical example of learning object creation using SGML/XML. *Canadian Journal of Learning and Technology, 28*(3), 73-89.

Bray, T., Paoli, J., Sperberg-McQueen, C. M., & Maler, E. (Eds). (2000). *Extensible markup language (XML) 1.0 (Second Edition)*. Retrieved November 11, 2003, fromhttp://www.w3.org/TR/2000/REC-xml-20001006

Dublin Core Metadata Initiative (2003). *Dublin core metadata element set, (Version 1.1: Reference Description)*. Retrieved January 21, 2004, fromhttp://dublincore.org/documents/dces/index.shtml

Friesen, N. (2004). Three objections to learning objects. In McGreal, R. (Ed.). *Online education using learning objects*. London: Routledge/Falmer.

IEEE Learning Technology Standards Committee, Learning Object Metadata Working Group (2000). WG12: Learning Object Metadata. *Working group information, announcements & news*. Retrieved October 13, 2003, from http://ltsc.ieee.org/wg12/index.html

IEEE Learning Technology Standards Committee, Learning Object Metadata Working Group (2002). *Draft standard for learning object metadata*. Retrieved December 16, 2003, fromhttp://ltsc.ieee.org/wg12/files/LOM_1484_12_1_v1_Final_Draft.pdf

IMS Global Learning Consortium. (2003). *IMS learning design specification*. Retrieved January 21, 2004, from http://www.imsglobal.org/learningdesign/index.cfm

Koper, R. (2001). *Modelling units of study from a pedagogical perspective: The pedagogical meta-model behind EML*. Retrieved January 21, 2004, from Open University of the Netherlands Web site: http://eml.ou.nl/introduction/docs/ped-metamodel.pdf

McGreal, R. (2004). Learning Objects: A Practical Definition. *The International Journal of Instructional Technology and Distance Learning, 1*(9). Retrieved September 22, 2004, fromhttp://www.itdl.org/Journal/Sep_04/article02.htm

McGreal, R., Anderson, T., Babin, G., Downes, S., Friesen, N., Harrigan, K., et al. (2004). EduSource: Canada's Learning Object Repository Network. *The International Journal of Instructional Technology and Distance Learning, 1*(3). Retrieved September 22, 2004, fromhttp://www.itdl.org/Journal/Mar_04/article01.htm

Wiley, D. A. (2000). Connecting learning objects to instructional design theory: A definition, a metaphor, and a taxonomy. In D. A. Wiley (Ed.), *The instructional use of learning objects.*[Electronic Version]. Retrieved January 21, 2004, fromhttp://reusability.org/read/chapters/wiley.doc

---