



Guest Editorial on Metadata

Erik Duval¹ and Robby Robson²

¹Departement Computerwetenschappen, K. U. Leuven, Leuven, Belgium

²Saba Software, Leuven, Belgium

BACKGROUND

This is a special issue on the topic of *metadata*. An often-cited definition of metadata is ‘data about data’. In most cases, this means data that describe documents, for example, the author of a document, the date that a photograph was taken, or the Universal Resource Locator (URL) of a Web site. The World-Wide Web Consortium defines metadata as ‘machine understandable information for the Web’ <<http://www.w3.org/Metadata/>>. This definition emphasizes that, in order to be useful in a technological context, metadata ought to be provided in a form that can be processed by machines. Examples of such formats are HTML META-tags, XML records, etc.

In the context of this special issue, metadata is understood as a description of a learning object in a very general sense. A learning object is defined as “any entity, digital or non-digital, that may be used for learning, education or training. [. . .] A metadata instance for a learning object describes relevant characteristics of the learning object to which it applies. Such characteristics can be regrouped in general, educational, technical and classification categories.” (LOM, 2000).

As illustrated by the papers in this special issue, the main use of metadata in the context of learning environments is for identifying and accessing educational content. This is extremely important, as it enables the reuse of such content, a condition sine qua non if widespread deployment of educational technologies is to become a reality.

Some of the most relevant research and development in this area is taking place in the fields of information retrieval and digital libraries (CACM, 2001).

More specific educational applications that rely on metadata include learning management systems and adaptive learning (Greenberg, 2000). More specifically, a number of metadata-centric infrastructures are currently being developed and deployed by organizations such as ARIADNE <<http://www.ariadne-eu.org/>>, CSTC <<http://www.cstc.org/>>, EDNA <<http://www.edna.edu.au/>>, GEM <<http://www.geminfo.org/>> and SMETE <<http://www.smete.org/>>. These infrastructures typically include a central or distributed repository of learning objects and their associated metadata.

Looking at the broader picture, the field of educational metadata is undergoing consolidation, mainly because of the rapidly maturing standardization efforts. The most relevant of these efforts is the IEEE Learning Technologies Standardization (LTSC), Learning Object Metadata (LOM) working group, which has been working on an educational metadata specification for over 3 years with participation from industrial and academic partners. The result is currently in ballot and is expected to become a standard sometime in 2001. This specification is based on earlier work by the ARIADNE Foundation, as well as the IMS Global Learning Consortium. The LOM specification has also been adopted by the Advanced Distributed Learning (ADL, <<http://www.adlnet.org/>>) initiative, in its sharable content object reference model (SCORM).

In LOM, data elements are grouped in categories. The LOM scheme consists of nine such categories.

- General information that describes the resource as a whole.
- Features related to the history and current state of the resource and those who have affected this resource during its evolution.
- Information about the metadata record, rather than the resource that the record describes.
- Technical requirements and characteristics.
- Educational and pedagogic characteristics.
- Intellectual property rights and conditions of use.
- The relationship between this resource and other resources.
- Comments on the educational use of the resource and information on when and by whom the comments were created.
- Where this resource falls within a particular classification system.

The last category enables an end user to classify a learning object according to arbitrary classification structures. As any classification can be referenced, this category provides for a simple extension mechanism.

The work of the IEEE LTSC LOM working group is being further developed and adapted to the needs of a multilingual and multicultural context by the European CEN/CENELEC ISSS Learning Technologies WorkShop (LTWS, <<http://www.cenorm.be/iss/workshop/lt/>>). In that group, emphasis is put on, among others, translations and internationalization of the LOM specification, as well as on the use of taxonomies, vocabularies and ontologies, in order to achieve semantic interoperability. It is expected that the LOM specification will eventually be adopted by the ISO/IEC JTC1 SC36 on Information Technology for Learning, Education, and Training <<http://www.jtc1.org>>.

The Dublin Core Metadata Initiative (DC, <<http://purl.org/DC/>>) has developed a specification consisting of 15 metadata elements for a broad range of purposes. It has set up a working group on educational applications (DCEd) that proposes two additional elements to the DC element set (audience and standards), a domain-specific qualifier to the DC relation element (conformsTo) and the endorsement of three elements from LTSC LOM specification (interactivity type, interactivity level and typical learning time).

In late 2000, the IEEE LTSC LOM working group and DCMI established a Memorandum of Understanding expressing “their joint commitment to develop interoperable metadata for learning, education and training”. It is expected that this will lead to technical solutions for increased ‘compatibility’ between the two specifications, and potentially to a full-scale harmonization in the longer term.

CONTENTS OF THIS SPECIAL ISSUE

A general, non-technical introduction to the field of pedagogic metadata is the subject of the paper by Robson. It also provides additional detail on the aforementioned LOM work.

The paper by Friesen focuses on the concept of ‘educational objects’, their use and reuse. Characteristics such as discoverability, modularity and interoperability are analyzed in detail. The author argues that the novelty of educational objects is much more linked to a teaching practice based on collaboration and reuse, rather than to the technological characteristics.

The report on the results of the GESTALT project by Konstantopoulos et al. focuses on the semantic and structural relationships between metadata for

content and for its users, the latter in the form of so-called 'learner profiles'. The paper situates both kinds of metadata in the context of a system architecture. The authors argue that, as characteristics of users need to be mapped against those of content, the schemas for content metadata and learner profiles should be closely integrated, if not the same. They also report on their practical experiences in the LOM based GESTALT demonstrator, implemented on top of CORBA for the broker service and LDAP for the learner profiles.

In their paper, Recker and Wiley focus on the use of multiple, 'non-authoritative' metadata that relate to different instructional contexts of use for learning objects. In contrast with the more objective 'authoritative' metadata that are typically provided at authoring time, the more subjective 'non-authoritative' metadata relate to the actual use of the learning objects. They can be exploited by automatic filtering systems that can act as recommender agents. Their Instructional Agent is a prototype implementation of their ideas, relying on the LOM specification.

Suthers describes the application of LOM metadata in the development of a database of diverse resources for Hawaiian schools. The paper details the development approach, as well as specific results on the use of vocabularies and structured descriptions. Through the LOM extension mechanism, additional detail is provided, to describe the audience of a learning object, community involvement, the discipline covered, the educational level and objectives, as well as the pedagogical approach.

The paper by Greenberg deals with metadata applications for a plant information center. The focus of this project is on the use by students of primary resources that are normally used by scientists, so as to make the students more acquainted with the aims and methods of botanical science. The resources include digitized herbarium specimens, reference books, research notes, lesson plans, etc. One of the major aims is to share collection holdings. The paper focuses on the role of different metadata schemes. Dublin Core is used for resource discovery and GEM for lesson plans. RLG Preservation metadata elements are used to document the archival copy of specimens. Specific PIC schemas have been defined for specimen images, FAQ lists an 'ask the expert' application. Metadata instances that comply to these schemes are managed in an Access database with an ASP-based Web interface.

CONCLUSION

We believe that educational metadata is on the brink of large-scale adoption, and that this will have a significant impact on research and development of

interactive learning environments in general and learning management systems in particular. We hope that this special issue will contribute to a better understanding of the wide-ranging issues relevant to educational metadata and its diverse applications.

Enjoy your reading!

REFERENCES

- CACM. (2001). Special issue on Digital Libraries. *Communications of the ACM*, 44(5).
- Greenberg. (2000). In J. Greenberg (ed.), *Metadata and organizing educational resources on the Internet*. New York: Haworth Information Press (Also published in the *Journal of Internet Cataloging*, 3(2/3)).
- LOM. (2000). IEEE LTSC, *Draft Standard for Learning Object Metadata*, working draft 5, November 2000. (<<http://ltsc.ieee.org/wg12>>)