A Novel Approach for Metametadata Views in Learning Domains

Amirah Ismail
Department of Computer Science
University of Warwick
Coventry, CV4 7AL
UK
Amirah.Ismail@warwick.ac.uk

Mike Joy
Department of Computer Science
University of Warwick
Coventry, CV4 7AL
UK
M.S.Joy@warwick.ac.uk

Abstract : In this paper we consider the use of metametadata in the context of a pedagogic architecture using learning objects. We propose a novel taxonomy for metametadata, supported by pedagogic ontologies, suitable for dynamic metametadata classification incorporated in such an architecture. This proposed taxonomy will act as the reference tool to link the decomposition of the semantics relationship for ontologies and metametadata. Examples of pedagogic metametadata are presented to illustrate the classification and the metametadata views will be introduced and discussed.

1. Introduction

A great challenge in knowledge based systems is to provide models that effectively represent domain-specific knowledge and integrate the pedagogical content for learning materials. Moreover, the applications which are interoperable within the semantic web environment require models that support the state-of-the-art information schema standards. These models will capture specific detailed information from each learning materials, and thus support allocation of work resources in the learning environment.

This study focuses on a novel metametadata schema which represents such a model. The schema is motivated by SCORM (Sharable Content Object Reference Model), for which the IEEE LOM element set provides for the descriptive metadata, and which contains guidelines on how to package XML metadata [14]. SCORM draws on a variety of standards to create reference model specifically for learning objects.

Metadata is data about data, *"structured data which describes the characteristics of a resource" [2]*. Metadata is used to describe, organize, locate and manipulate the data structure, content, context, quality and ownership. Metadata can be stored for different types of media format, such as texts, images, videos, *animation files, sounds, and so forth. A typical metadata file might (for example) consist of title, description, catalogue URL, resource type, intended end use, age range and keyword.*

Metadata can also be considered as data, and then metadata about such data is known as *metametadata*. The existing metadata schemas such as Dublin Core, IEEE LOM and IMS, contain metametadata that describe organization-specific metadata [15].

A simple example would be a table in a network database that consists of information about all the tables in the database management system, such as name, date of creation, use, creator, ownership, columns, rows, access of restrictions, etc.

DLESE [5] gives an interesting definition of metametadata: "Metametadata are features related to the creation of the metadata record including persons or organizations contributing to the content of the metadata (resource catalogers), copyright, terms of use, language and status of the metadata, date information and catalog record numbers".

Learning objects, which are used in pedagogic architectures, can be defined as "any entity, digital or non digital, which can be used and reused or referenced during technology supported learning" [16]. Wiley proposes a slightly narrower definition: "any digital resource that can be reused to support learning" [9].

Metadata which is used by learning objects can be categorized into *intrinsic* and *extrinsic* metadata [3]. Intrinsic data refers to properties of an object derived from its content and to basic information about it, such as title, author and subject. Extrinsic data describes the context of the work and may be used for management purposes, and might include information about the author, such as email, department or digital signature. These intrinsic and extrinsic characteristics can be used to help identify and locate learning objects.

It is desirable that metadata be exchanged between different applications. In order for metadata standards to interoperate, both the syntax and semantics of the metadata schemas need to be considered [7]. The semantic web refers to the web supplemented by precise codified meaning associated with the information contained in the web, and this supports humans in using the web [1]. This semantic content is facilitated by *ontological* metadata which represent the semantics of the data.

An ontology [3] is an abstract model of real-world phenomena and the relationships between the entities relating to the phenomena. Research into ontologies is cross-disciplinary and includes contributions from (amongst others) information science, computer science, artificial intelligence, e-commerce and knowledge management [4]. Since ontology research is a relatively recent area of interest, appropriate research methodologies are not yet established, although a well-known set of guidelines for working with ontologies has been offered by Gruber [7].

The theoretical use of ontologies in e-learning applications has been investigated, and architectures for a prototype system for e-learning using ontologies have been proposed [10][13]. The theoretical foundations of integrating competency ontologies with e-learning have also been studied and have been partially implemented [6].

2. Why Metametadata

A principal motivation for using metametadata in the context of a pedagogic architecture which uses learning objects is that if the designer or administrator wishes to integrate metadata from various repositories or sources, the format and content of the metadata may vary considerably. A "high-level" view of the metadata, in the form of metametadata, will assist the process.

Metametadata are structured descriptions about a set of metadata which intelligently describe and capture relevant identified characteristic properties and relationships between metadata types to aid locating and managing and retrieving data.

Metametadata are useful for the following purposes.

- Providing sufficient information about metadata to enable intelligent searching via the metadata.
- Implementing flexible dynamic semantic mappings between metadata vocabularies.
- Processing and displaying different explicit and implicit characteristics of the stored data sets.

- Associating sets of related data by identifying semantic relationships between the associated metadata.
- Providing consistent semantics and structures for metadata in the repositories or database schemas, browsing interfaces and presentation of content.

3. A Taxonomy of Metametadata

A classification scheme for pedagogic metametadata has been designed in order to provide a strong foundation for the future implementation of a pedagogic architecture supported by metametadata.

Implicit metadata for leaning objects is often used for administering purposes and can be captured through the context of the learning object. Explicit metadata is normally straight forward metadata that is coded in a simple format.

We can consider metametadata as also being either implicit or explicit. As an example of implicit metametadata, we might have a relationship that states that "Adam *wrote* 'LearningJava.org'" is similar to "Adam *created* 'LearningJava.org'".

In terms of metadata, we might have the following two related tags for the learning objects stored in LearningJava.org: <creator name="Adam"> and <writer name="Adam">. In other words, there is a semantic similarity between the tags and attributes stored in the metadata, and an element of metametadata might capture that similarity. Such implicit relationships might be queried by users through a database interface browser, so that 'LearningJava.org' would be selected by a query "web resources *authored* by Adam".

In another learning context, a programmer is developing and testing software for two projects, and the files are marked up with metadata. Using metametadata the similarities between classes of files with equivalent functionality may be represented.

For explicit metametadata to be viable, we need to understand how to identify (address) individual metadata elements external to a specific metadata instance. These can be linked with connector metametadata types that will identify metadata for specific locations, such as URIs included in structured metadata in other documents.

In the context of this paper, relationship types for metametadata are proposed to connect with these three metadata types: *Semantic* metadata that can be used to describe the subject matter of the resource or document; *Context* metadata which characterizes relationships with external entities or the meaning of the learning objects or documents (for example, author, publisher); and *Structural* metadata which indicates a description of the internal media type, structure and presentation layout, such as text, sound, image, simulation, video, etc.

The reason for proposing a taxonomy is to provide a common framework containing semantic definitions together with further contextual expression.

The IEEE LOM specification consists of a series of data elements that divided into the following nine categories as follows: General, Lifecycle, Metametadata, Technical, Educational, Rights, Relation, Annotation, and Classification [9]. IEEE LOM just provides guidelines for developing metadata, but does not yet cover the metametadata functionalities for capturing relationships between metadata, especially under the Educational category. Properties of instances of metadata need also to be taken into account to ensure the validity of information being retrieved via the metametadata which provide dynamic views on the learning objects.

The context of metadata may change over time. So, we need metametadata to link to the previous metadata instances and the current version of the metadata to identify modifications applied to them. The educational metadata category is designed to define the essential elements describing the pedagogical aspects of the learning objects or learning materials.

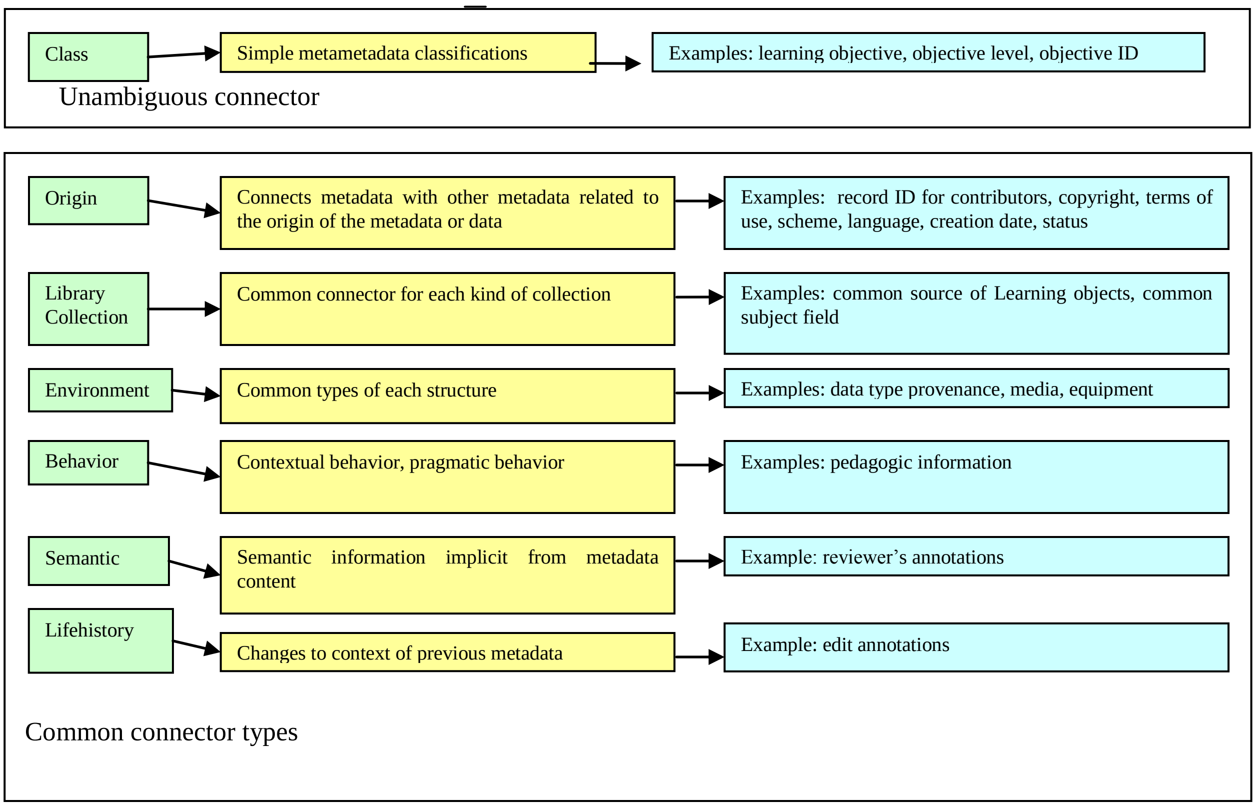

Figure 1: A proposed Pedagogical Metametadata Taxonomy

However, the educational category has not described the significant connections or relationships between each of the following metadata elements: Interactivity type, Learning resource type, Interactivity level, Intended end user role, Context, Difficulty, Typical learning time, Description and Language of the typical intended user [10].

The proposed metametadata relationship defines the semantic relationship between pedagogical metadata elements. Educational metadata from one category in the IEEE LOM specification covers the pedagogical aspects or elements for the learning objects.

Other elements listed – the interactivity type or level, semantic density and difficulty – have not been elaborated further here. There is a need to improve the semantic relationships between metadata under the educational metadata category in LOM in order to improve learning object reusability. Therefore, it is vital to find a semantic definition by describing each metametadata type that would link pedagogical aspects of chosen learning objects.

We propose a taxonomy as shown in Figure 1 for pedagogic metametadata which uses the IEEE LOM metadata specification elements, together with key pedagogic characteristics, and metametadata elements for relational and classification purposes.

The distinction between data and metadata is well understood, and metadata models may be described by classes, relationships and properties, known collectively as *types*. Our proposed taxonomy consists of a collection of types of metametadata, analogous to types of metadata, which we refer to as *connectors*.

These subdivide into two distinct categories, *unambiguous connectors* and *common connectors*.

1. *Unambiguous connectors*. These are classification metametadata, such as identifications for types of metadata which might be used for cataloguing purposes. There is only one type of unambiguous connector, which we refer to as the *Class* type of metametadata.

2. *Common connectors*. These represent any instances of relationships between selected metadata and other metadata. We can subdivide these into six generic abstract classes which we refer to as *types*.

- *Origin Type*: an attribute of the origin of the records. For example, two documents sharing a common author might use origin metametadata to store that relationship.
- *Library Collection Type*: information about commonality of a group of metadata. For example, the fact that a set of learning objects is sourced from a common repository might be represented by library collection metametadata.
- *Environment Type*: information about commonalities in the administrative or technical metadata. For example, a set of learning objects which share a common type of interface, which could be identified by the authoring tools (as specified in their metadata), would be linked by environment metametadata.
- *Behavior Type*: information about metadata behavior, such as contextual or pragmatic. For example, a set of learning objects which contains metadata indicating the cognitive abilities of the target students might be identified through behavior metametadata.
- *Semantic Type*: information about semantic content of metadata. For example, if a set of learning objects contains metadata which are reviews of each object, then a subset of those objects with positive reviews might be identified through semantic metametadata.
- *Lifehistory Type*: information about changes in metadata. For example, two Learning objects whose metadata had been edited at a similar time might be linked using lifehistory metametadata.

4. Metametadata Scenario

In the following learning scenario, a student has been asked to do a Java programming assignment relating to class inheritance as part of an intermediate level software course. The student, who is a “visual learner”, is not fully conversant with the material required to complete the assignment, but believes that he/she is competent with basic Java material. The student is provided with access to a pedagogic tool to assist them, but the student's institution does not have further financial resources to apply to the activity.

The versioning history will be a tree with known and unknown branches which can be,

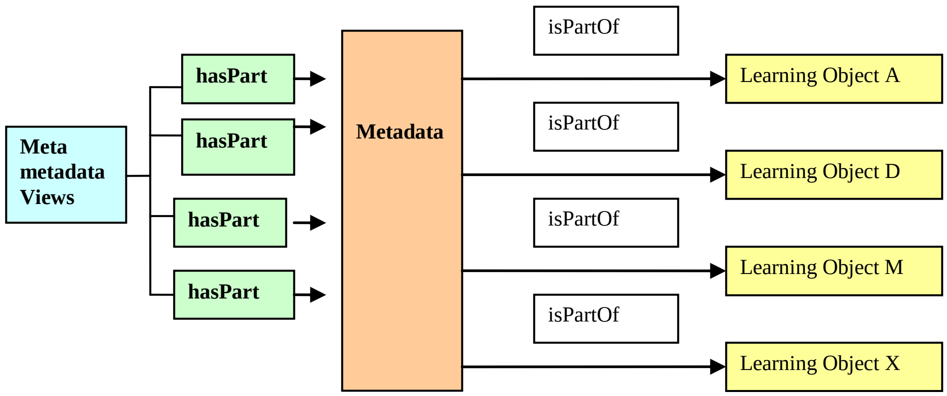

Figure 2: Metametadata views in a learning scenario

These metametadata views for learning acitivities is shown by figure 2. The data type relation for commonalities for each metametadata types are represented by data type, *hasPart* and *isPartOf* that is based on LOM data set. This facilitate the workflow for retrieving metametadata that matched and mapped the equivalent need for the student.

The tool provides a variety of learning materials, packaged as learning objects, and uses adaptive mechanisms to select appropriate materials for the student's use. The tool's architecture uses the metametadata approach to structuring and annotating its data store that we have described in the previous sections.

In order to support the adaptivity, information stored as metametadata is used to identify which suitable learning objects should be retrieved from the data store, and how they should be presented to the student. Initially, the tool identifies, through the *library collection* metametadata, which sources of learning object are from repositories which are either free, or for which the institution has already purchased a subscription.

The metadata of each learning object which relates to class inheritance in Java may contain information related to the specific topic, learning level and focus of that object. This is basic pedagogic information which *class* metametadata would relate to.

The student is sensitive to the quality of educational materials he/she engages with, and has specified that they should only be presented with learning objects of good quality. Such materials may have received high reviewer ratings, or be authored or edited by internationally renowned educators, and this information is available using *semantic* metametadata, since it relies on the semantics of the metametadata.

Java technologies are constantly changing, and although the basics of the language have not changed, there have been detailed enhancements. For example, the recent inclusion of Generics relates to the student's study topic, and it is important that he/she be presented with up-to-date material.

Origin metametadata can be used to filter out-of-date learning materials, and *lifehistory* metametadata to further identify learning objects which have either been brought up to date or have been corrected, and hence inform the tool of the integrity and reliability of the materials selected. Due to the student's visual learning style, those materials which display appropriately in context can be identified using *environment*metametadata.

Not only does the student require materials which address the topic, the difficulty level, the student's learning preferences (etc.), but he/she also needs an heterogeneous mix of materials as chosen by the tool's adaptivity mechanism. For this, *behavior* metametadata provides contextual information supplementary to the basic pedagogic information already available. In this scenario,

We have not prescribed the tool architecture or how it provides adaptivity, nor how that metametadata should be generated from the metadata bundled with the learning objects – these are details outside the scope of this paper. We have simply identified how such a tool might take advantage of metametadata, as classified by our taxonomy, in a natural and straightforward manner.

5. Conclusions

We have presented a novel taxonomy of metametadata for use in a pedagogic architecture employing learning objects. Our results are founded in pedagogic ontologies, and it is our intention that they should be used to develop and implement such an architecture.

Our claim is that the use of dynamic metametadata using our classification will promote the emergence of learning repositories to share not only knowledge resources, tools and services, but also to optimize the practices of teachers and learning designers. Furthermore, this approach will support rich searching mechanisms that will contribute to a high-quality, flexible educational experience in a learning domain.

6. References

[1] Berners-Lee, T., Hendler, J., and Lassila, O., "The Semantic Web", *Scientific American*, **284**(5), 2001, pp. 345-355.

[2] Bechhofer, S., and Goble, C., "Thesaurus Construction through Knowledge Representation", *Data & Knowledge Engineering*, 37, 2001, pp. 25-45.

[3] Burnett, K., Ng, K. B., and Park, S., "A Comparison of the two Traditions of Metadata Developments". *Journal of the American Society for Information Science*, **50**(13), 1999, pp. 1209-1217.

[4] Ding, Y., and Foo, S., "Ontology Research and Development: Part 1- a Review of Ontology Generation". *Journal of Information Science*, 28(2), 2002, pp. 123-136

[5]DLESE, *Best Practice*, 2004. Available from: http://dcs.dlese.org/preview/adn/adn.do?command=bestpractices&pathArg=/itemRecord/metaMetadata [Accessed 5th January 2007]

[6] Fernandez-Lopez, M., "Overview of Methodologies for Building Ontologies", *Proceedings of the IJCAI-99 Workshop on Ontologies and Problem-solving Methods*, Stockholm, Sweden, 1999. Available from: http://www.lsi.upc.es/~bejar/aia/aia-web/4-fernandez.pdf [Accessed 5th January 2007]

[7] Gruber, T. R. "Toward Principles for the Design of Ontologies used for Knowledge Sharing", *Proceedings of the International Workshop on Formal Ontology,* Padova, Italy, 1993. Available from: http://ra.crema.unimi.it/lenya/turing/live/know/knowpubs/gruber93toward.pdf [Accessed 6th January 2007].

[8] Hirata, K., Ikeda, M., and Mizoguchi R., "Total Resolution for Human Resource Development Based on Competency Ontology", *Proceedings of ICCE2001*, 2001.

[9] IEEE Learning Technology Standards Committee (LTSC), *Draft Standard for Learning Object Metadata Version 6.1* .Available from: http://ltsc.ieee.org [accessed 11 January 2007].

[10] IEEE LOM. IEEE Std ™-2005 1484.12.3. IEEE Standard for Learning Technology. Extensible Markup Language (XML) Schema Definition Language Binding for Learning Object Metadata. 2005. Available from: *http://stdsbbs.ieee.org/descr/1484.12.3-2005/* {Accessed 10th February 2007]

[11] Modelling Multiple Media Data Collections. Available from http://nordbotten.com/ADM/ADM_book/Ch4_MM-Modelling.htm [Accessed 9th February2007]

[12] Miller, E., "An Introduction to the Resource Description Framework", *D-Lib Magazine*, 1998. Available from http://www.dlib.org/dlib/may98/miller/05miller.html [Accessed 11 January 2007].

[13]Qin, J., and Paling, S., *Converting a Controlled Vocabulary into an Ontology: the Case of GEM*, 2001, Available from: http://informationr.net/ir/6-2/paper94.html [Accessed 8 January 2007].

[14] SCORM. Shareable Content Object Reference Model Initiative. 2001. Available from: http://xml.coverpages.org/scorm.html . [Accessed 7 January 2007]

[15] Gardner, T**.** A Developer Perspective on IMS.1999. Available from :
http://www.ariadne.ac.uk/issue20/ims/[Accessed12th February 2007]

[16] Wiley, D. A. "Connecting learning objects to instructional design theory: A definition, a metaphor, and a taxonomy", in D. A. Wiley, ed., (2000). *The Instructional Use of Learning Objects: Online Version*. Available from: http://reusability.org/read/chapters/wiley.doc: [Accessed 2nd February 2007]