An Evolutionary Approach to Building a Learning Object Repository

Marek Hatala & John Nesbit

Technical University of British Columbia
2400 Surrey Place,
10153 King George Highway
Surrey, BC, Canada, V3T 2W1
email: {marek.hatala | john.nesbit}@techbc.ca

ABSTRACT

To support an increasing demand for customized access to learning opportunities, including on-demand learning, just-in-time learning, multiple modes of delivery and learner support, there is an emerging drive toward structured information. The growing requirement for high-quality, reusable components – learning objects – is being addressed by Canadian initiatives to build repository prototypes capable of demonstrating business models, public-private models, and the use of broadband-enabled eLearning. In this paper, we describe how the Technical University of British Columbia is addressing its own need for a robust repository solution that will support its core activities and maintain interoperability with national initiatives. The learning object repository development model we have followed involves three steps, based on the Seeding, Evolutionary growth, Reseeding (SER) model.

KEY WORDS: Internet-based educational systems, Advanced technology in education, Learning object repositories, Learning objects metadata

1. INTRODUCTION

With a growing number of organizations moving their training and education programs into the web environment, there is an increasing demand for high-quality, reusable components – learning objects (LOs). This demand comes from the realization that the development of high-quality learning objects is resource intensive and time consuming. There is a wealth of content available in public and private organizations but either it is not accessible to external users or it is difficult to find. In Canada, current publicly funded initiatives to build learning object repositories aim at establishing infrastructure for collections of high quality learning objects and related business models [1]. Another issue being addressed at the same time is the development of standards for descriptions of learning objects to enable national and worldwide interoperability.

The Technical University of British Columbia (TechBC) is a new, Canadian, high-tech university that has adopted innovative methods of course delivery. Each course has a substantial web component rich in media content. To explore how to manage a rapidly growing number of learning objects we have developed a prototype learning object repository. Features implemented in the prototype enable us to construct a requirements specification for a planned full-scale repository implementation while working with real educational content during a full development and delivery activities. Adopting metadata schemas supporting the Canadian Core Learning Resource Metadata Protocol (CanCore) [2] will guarantee interoperability with the national repositories currently under the development.

In the next section we provide an overview of the recent initiatives in the development of standards for metadata schemas and repositories for learning objects. Section 3 describes the context for the LO repository at TechBC and presents questions we are trying to understand and find answers for. Section 4 presents the implementation of the repository and paper concludes with plans for future research.

2. LEARNING OBJECT REPOSITORIES AND METADATA STANDARDS

There is an increasing demand for customized access to learning opportunities with on-demand learning, just-in-time learning, and multiple modes of delivery and learner support. We are seeing the emergence of a drive toward structured information [3] because it enables independence of content from the delivery system, dynamic rendering of content, and the development of more powerful content management systems. Structured information is essentially the application of a standards-

based mark-up language and metadata schema to tag learning objects.

The standards are important for interoperability between learning and business systems. Several standards for describing metadata have been developed through collaboration between the private and public sectors. The IEEE Learning Object Metadata [4] defines a set of metadata elements that can be used to describe learning resources. The IMS Global Learning Consortium has identified a minimum set of IEEE metadata elements called IMS Core [5]. Recently, new alliances have arisen to concentrate efforts on unifying different standards. For example, a Memorandum of Understanding (MoU) has been signed between the Dublin Core Metadata Initiative, IEEE LOM, ARIADNE (Alliance of Remote Instructional Authoring and Distribution Networks for Europe), EdNA (Education Network Australia), GEM (Gateway to Educational Materials), and the IMS Global Learning Consortium [6].

Standards allow developers of web-enabled learning technology to support an interoperable infrastructure for worldwide eLearning. These standards are crucial for building comprehensive learning object repositories. The goal of several current projects in Canada is to demonstrate LO repository business models, public-private models, and the use of broadband-enabled eLearning. These include POOL (Portal for Online Objects in Learning) and BELLE (Broadband Enabled Lifelong Learning Environment). The POOL and BELLE projects, and some of their participating institutions, have adopted CanCore as a metadata standard. CanCore is a national protocol that provides a standard for describing all multimedia educational objects. It builds on the Dublin Core and IMS standards to enable compatibility and allow seamless searches of educational object repositories located in Canada and around the world. It was developed by national and provincial educators and technology developers who were mainly POOL and BELLE project members.

The CanCore Schema provides 54 elements in total with 36 active elements, 15 placeholders and 3 reserved elements. The elements are organized into 9 groups describing different characteristics of the learning object. The table below provides a description of these groups, more detail information can be found on the CanCore website [1].

General	Groups information describing learning object as a whole. *Active elements:* Identifier, Title, Catalogentry.Catalog, Catalogentry.Entry, Language, Description, Coverage
Lifecycle	History and current state of resource. *Active elements:* Version, Contribute.Role, Contribute.Entity, Contribute.Date
Metametadata	Features of the description rather than the resource. *Active elements:* Identifier , Catalogentry.Catalog, Catalogentry.Entry, Contribute.Role, Contribute.Entity, Contribute.Date, Metadatascheme, Language
Technical	Technical features of the learning object. *Active elements:* Format, Size, Location, Otherplatformrequirements, Duration
Educational	Educational or pedagogic features of the learning object. *Active elements:* Learningresourcetype, Intendedenduserrole, Context, Typicalagerange, Language
Rights	Conditions of use of the resource. *Active elements:* Cost, Copyrightandotherrestrictions, Description
Relation	Features of the resource in relationship to other learning objects. *Active elements:* Kind, Resource.Identifier, Resource.Catalogentry
Classification	Description of a characteristic of the resource by entries in classifications. *Active elements:* Purpose, Taxonpath.Source, Taxonpath.Taxon.Entry, Keyword

3. LO REPOSITORY AT TECHBC

TechBC is building its undergraduate and graduate programs at a rapid pace. In its first two years of program delivery, the university offered 108 1-credit learning modules in three program areas: Information Technology, Interactive Arts, and Management and Technology. In the current development cycle, TechBC faculty supported by the Education Technology and Learning (ET&L) unit are developing 65 modules for the third year of the undergraduate program, and 15 modules for the graduate program starting in Fall 2001.

In the learning environment at TechBC each module has a substantial online component consisting of features such as web presentations with multimedia objects, quizzes, project assignments, sample exams, and asynchronous conferences. As the number of modules continues to grow, the need for a robust LO repository is becoming increasingly apparent. The repository will not only serve as a storage space for the learning objects and their delivery but will also support the reuse and sharing of learning objects between program areas and further encourage the interdisciplinary nature of the TechBC programs.

The fact, that the repository will be at the core of our business for several years to come, makes the choice of the repository and its structure a crucial one. In the process of specifying requirements, we realized that we needed to have a better understanding of the nature of the learning objects in use at TechBC and the workflow patterns that surround them. To address this need we have built a prototype repository that addresses both the immediate requirement to archive existing learning objects and the need to study LO workflow. The main benefit of this solution is that it enables us to make a final

decision about a full-scale repository and its structure once we understand what the real requirements are.

TechBC is a partner in the POOL project and is interested in sharing learning objects with other institutions. To support this goal it is essential that the LO repository be compatible with standards adopted by other repository projects. Consequently we chose to adopt CanCore as a basis for the internal metadata schema. Although CanCore provides the basic schema that supports sharing of learning objects across Canada, it is not sufficient for effective description of the learning objects in the TechBC context. Therefore, the prototype we have developed also serves the need for an evolving tool for the metadata schema that will be implemented in the future full-scale repository.

3.1 Relationship between production workflow and the LO repository

The development of learning modules at TechBC is an elaborate process involving several phases and roles. Modules move through planning, specification, design, production and post-production phases as they are being developed. Although the primary responsibility for carrying module development forward lies with faculty, module production staff in the ET&L unit provide necessary professional services such as instructional design, learning object search, licensing and copyright clearance, multimedia development, and editing.

Interactions among these roles has been fairly ad hoc, but the introduction of the prototype LO repository is bringing questions about workflow to the fore. Figure 1 shows a simplified model of workflow related to LO procurement and production. The module author, perhaps working with an instructional designer, identifies the need for a learning object. If the author develops the object herself, say in the case of a printable worksheet, she might be the one to enter metadata into the repository. Otherwise she might provide brief requirements to the course researcher, who searches international repositories and the entire web. If a suitable object is found, the licensing specialist is tasked with procuring usage rights. Only when rights have been obtained does the licensing specialist enter metadata to the repository and release the object for use. If no suitable LO is available, the author works with the instructional designer to create a detailed requirements specification to pass on to the media developer. In this case it is likely to be the media developer who records the metadata.

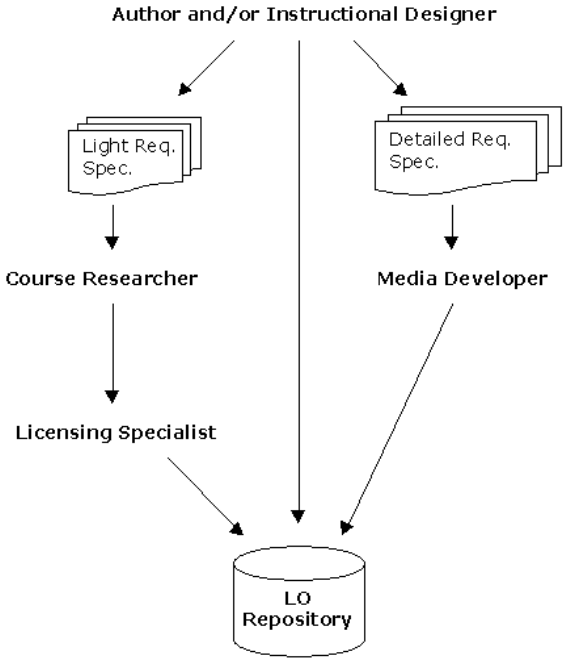

Figure 1. Simplified model of LO production workflow.

In considering how the simplified model might evolve, some questions that arise are: How will the course researcher role change when extensive global repositories become available? If extensive and time-consuming cataloging is required, should there be a dedicated LO cataloging role? Should the repository provide tools for managing licenses? Can the repository usefully evolve to mediate the interaction between the media developer, the instructional designer and the author? How far should the repository evolve to become a full content management system, tracking and manipulating an LO throughout its development and usage life cycle? Since the repository will classify LOs by academic discipline, and faculty within a discipline will, in effect, use it as a medium for the exchange of ideas, how far should the repository evolve to support communities of scholars?

3.2 LO repository as a portal

One of the main criteria for implementing the prototype repository was that it should not interfere with ongoing module development and delivery. At the same time, the repository has to contain real content that the faculty and other personnel are working with. Therefore we have build a LO repository prototype as a portal that collects LO metadata but does not provide content management capabilities. Instead, the repository stores links to the learning objects. This naturally fits into the web-based course management system (CMS) used at the TechBC. To submit the learning object into the repository, the user refers to the objects using its URL in the CMS system and fills in the metadata describing the object. The repository itself maintains the metadata records and provides a sophisticated search facility.

3.3 Support for the metadata expansion

One of the main features of the implemented prototype is a capability for end users to expand the core metadata

schema (i.e. CanCore). The expansion of the core metadata schema has been done in three different ways:

1. *Predefined vocabularies for freeform fields.* For some freeform fields (those that enable the user to enter free text, e.g. 'Contribute.Entity' in Figure 2) we have predefined a vocabulary of the most often used terms. This speeds-up a process of metadata creation.
2. *Expanded vocabularies.* For some fields that have vocabularies defined in the CanCore we have expanded these vocabularies by adding TechBC specific values. For example, for the 'Contribute.Role' field, in addition to the 'Author' value we have added the value 'Author.Techbc' indicating an author from the TechBC, etc. This expansion represents a subclassing of predefined values and enables us to perform more specific searches in the repository.
3. *Soliciting new values for the vocabularies.* For most of the fields with pre-set vocabularies we have enabled the user to define new values.

Table below summarizes expansions made to CanCore in TechBC context.

Element	Predefined vocabularies	Expanded vocabularies	Soliciting new values
MediaType		?	?
Lifecycle.Contribute.Role		?	
Lifecycle.Contribute.Entity	?	?	?
Metametadata.Contribute.Role		?	
Metametadata.Contribute.Entity	?	?	?
Educational.LearningResourceType		?	
Technical.OtherPlatformRequirements	?		?

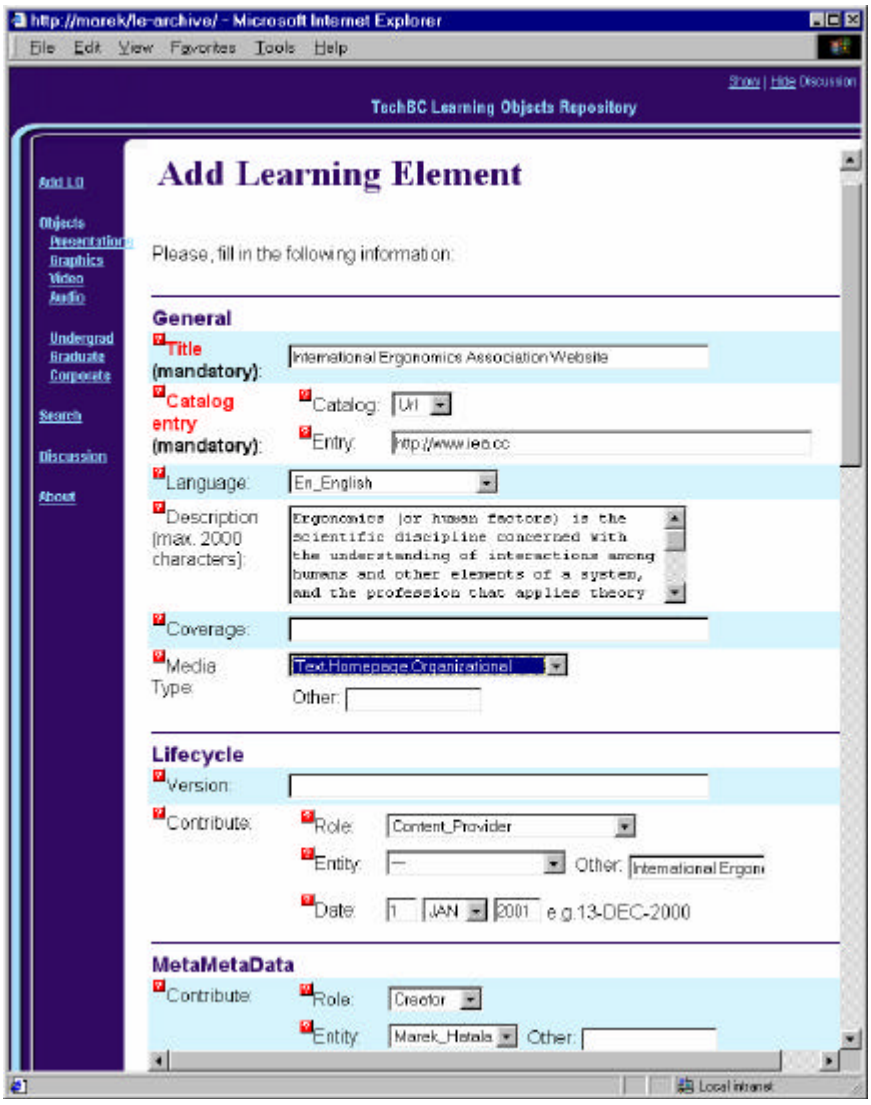

Figure 2. Screen shot from the prototype LO repository.

3.4 Adding elements to the LO Repository

The tagging of the objects is performed in two steps. In the first step, the object is tagged with metadata by the object creator (e.g., faculty member or media developer). The object creator fills in the form shown in Figure 2 and goes through the confirmation process where he or she can correct fields. The guidelines for filling fields are accessible through the small question mark icon in the field title. In certain fields, such as Media Type in Figure 2, the user either selects values from the predefined vocabularies, or specifies a value for describing the object in the 'Other' field. Newly defined values are collected by the system and may be either added into the vocabulary for the field automatically or stored in the system for later review and acceptance by an administrator. If added automatically, they appear in the vocabulary for the field next time the submit form is displayed. All new values are available in the search form.

In the next step, the content editor reviews newly added records for the completeness and correctness. The editor uses the same form as shown in Figure 2. A licensing specialist reviews and updates only records relating to external LOs for which copyright clearance must be obtained

When the system generates a metadata record for the object it also automatically associates a comment area with it. This area is available to all users to insert their comments and suggestions on the usage of the learning object.

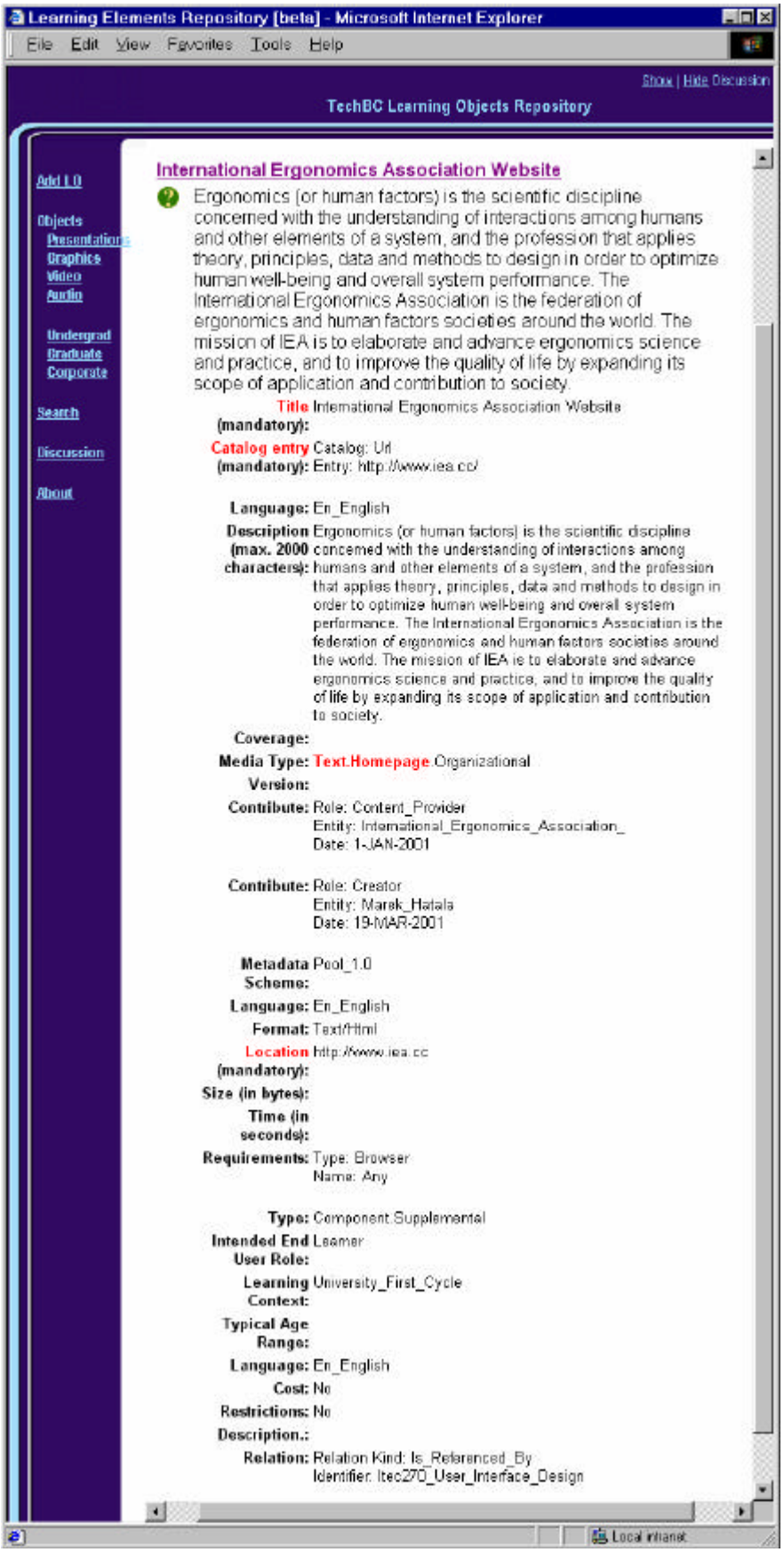

Figure 3 Metadata record for learning object

The resulting metadata record is shown in Figure 3. The title of the object is made into the hypertext link making it possible to open the object in the browser. Below the title a description of the object is shown together with the question mark icon. The icon is made into the hyperlink that opens the discussion space associated with the object. Below the description a full metadata record is shown (highlighted values are those that were used in the query when retrieving the object).

Figure 3 shows a long format of the record. Typically, users display just a short format without metadata values and display full metadata only for those objects that are of interest to them.

3.5 Search mechanism

The search mechanism in the implemented prototype is a combination of three types of searches:

1. *Full text search.* In this search the system searches for the text specified by the user in fields holding text values. The user can select text fields s/he is interested in.
2. *Metadata fields search.* This search operates over the metadata fields with limited vocabularies. The user defined values are compared with the values in the metadata records.
3. *Taxonomy based search.* Some metadata fields have predefined vocabularies with values creating taxonomy. For example, the vocabulary for the 'Media Type' field contains values 'Text', 'Text.Homepage' and 'Text.Homepage.Organizational'. The taxonomy based search uses this information to find relevant objects. Figure 3 shows the metadata of the record retrieved by the search

for ‘Text.Homepage’ (highlighted in Figure 2), although the record itself represents a more specific value, i.e. ‘Text.Homepage.Organizational’.

The results of three searches are combined together and sorted according a multi-tier ranking mechanism.

4. LO REPOSITORY DEVELOPMENT MODEL

The LORE development model involves three steps, based on the Seeding, Evolutionary growth, Reseeding (SER) model [7]. The SER model describes the necessary evolution of systems that are embedded in organizations. The designer can never fully envisage how the system will come to be used. The users themselves must have a stake in the evolution of the system, and the overall design must be flexible and amenable to evolution.

Seeding creates, through a participatory design, an initial state encompassing an initial prototype of the repository with the initial metadata schema. Evolutionary growth occurs through use. Emerging patterns in the use of the prototype and metadata lead to modifications of the prototype and the development of metadata schema as further knowledge is accumulated. The reseeding process occurs after some period of use and can be thought of as a process of revolution rather than evolution. For example, key elements in the metadata can be reorganized, or extensive changes made to tools in response to emergent use patterns. The previous section described how the seed of the repository was created in terms of the metadata schema and initial content. Evolution of the repository is sustained by the faculty and ET&L production staff with only limited software development support. The repository evolves through the way the learning objects are tagged, the metadata schema is expanded, and metadata descriptions are accumulated.

Reseeding involves more radical changes to the metadata schema and repository structure, and may require the re-tagging of learning objects in the repository. Reseeding would generally require the involvement of the original repository designers and the team preparing the implementation of the full-scale learning object repository. The nature of the reseeding process would be motivated by the results of the evolution process and formal evaluations of the prototype repository.

5. CONCLUSION AND FUTURE WORK

The LO repository at the TechBC is currently in an evolutionary stage in which faculty and media developers are adding learning objects and tagging with metadata as they proceed with current module development activities. In the next development cycle, when revisions of previous modules will be carried out, we plan to incorporate learning objects from the older modules. At the same time, we will continue to participate in the POOL project and will incorporate any features that will be required to maintain compatibility with emerging standards. This will give us enough information to proceed with the final requirements specifications and metadata schema for the full-scale learning objects repository at TechBC.

REFERENCES

[1] Learning Program, Canarie Inc., http://www.canarie.ca/funding/learning/backgrounders(99).html

[2] Canadian Core Learning Resource Metadata Protocol, http://www.cancore.org

[3] D. Porter, Learning That Fits: Objects Lessons from the Web, Proc. SURA/ViDe Annual Digital Video Workshop, GCATT, Atlanta, GA, March 20-21, 2001

[4] IEEE Information Technology - Learning Technology - Learning Objects Metadata (P1484.12), Learning Object Metadata Working Draft (WD5), http://ltsc.ieee.org/doc/wg12/LOM_WD5.pdf

[5] IMS Learning Resource Meta-data Information Model Version 1.1 - Final Specification, http://www.imsproject.org/metadata/mdinfov1p1.html

[6] Dublin Core Metadata Initiative, http://purl.org/dc

[7] G. Fischer, R. McCall, J. Ostwald, B. Reeves and F. Shipman, Seeding, evolutionary growth and reseeding: Supporting the incremental development of design environments. *Human Factors in Computing Systems (CHI ‘94)*, April 24 - 28, Boston, MA, 1994.