Reusable Learning Objects: a Survey of LOM-Based Repositories

Filip Neven
K.U.Leuven, Computer Science Dept.
Celestijnenlaan 200 A
B-3001 Leuven, Belgium
+32-16-327659
Filip.Neven@cs.kuleuven.ac.be

Erik Duval
K.U.Leuven, Computer Science Dept.
Celestijnenlaan 200 A
B-3001 Leuven, Belgium
+32-16-327066
Erik.Duval@cs.kuleuven.ac.be

ABSTRACT
In this paper, we survey the field of learning object repositories. Learning objects are typically relatively small content components that are meant to be reusable in different contexts. Associated to these learning objects are metadata, so that they can be managed, searched, etc. As the international standardization in this area is making important progress, the number of these repositories is growing rapidly, and the whole field of learning objects is rapidly maturing as a research area in its own right.

Categories and Subject Descriptors
K.3.1 [**Computer uses in education**]: Computer-assisted instruction (CAI), Distance Learning.

General Terms
Reusable Learning Objects, Metadata, Digital Libraries

Keywords
Metadata, Learning Object Metadata (LOM), Learning Object repositories

1. INTRODUCTION
In this paper, we will compare the features and architecture of repositories that typically contain learning objects or references to them, as well as metadata based on the Learning Object Metadata standard (LOM). We will refer to such Learning Object Repositories as LOR's. After a short explanation on the nature and use of LOM repositories, we will summarize and explain the criteria we used for the comparative analysis. Next, we will summarize our findings and provide a non-exhaustive overview of pro's and con's of different features and architectural properties of a LOR.

1.1 LOM Repositories (LOR's)
In order to improve the availability of good quality multimedia learning material, 'share and reuse' of content components, called 'Learning Objects' (LO's), is crucial. Moreover, an adequate search and retrieval system is required to support flexible access to the LO's. LO Metadata capture characteristics of the LO's and their potential educational application, using a newly approved IEEE standard [1].

Multimedia'02, December 1-6, 2002, Juan-les-Pins, France.

A LOM repository or LOR stores both LO's and their metadata, either by storing them physically together or by presenting a combined repository to the outside world, while the metadata and LO's are actually stored separately. The metadata scheme a LOR uses is based on the IEEE LOM standard, through a process that is typically referred to as an 'application profile' [2].

A LOR allows registered or unregistered users to search and retrieve LO's from the repository. Searching for LO's is based on criteria that relate to LOM data elements. A LOR typically supports simple and advanced queries, as well as browsing through the material by subject or discipline. In a simple query, keywords given by the user are matched against the text in a number (or all) of the metadata elements. An advanced query allows a user to specify values for specific metadata elements (e.g. 'easy' or 'medium' for 'Difficulty level'), and sometimes also to rely on logical combinations of search criteria. Browsing typically allows the end user to descend in a tree of disciplines and sub-disciplines to get an impression of the objects available in different domains.

2. COMPARISON CRITERIA
In this section, we summarize the criteria we used in our comparative analysis. The goal of these criteria is not to select the 'best' LOR. Evaluating a LOR should be done with context, purpose and target user group in mind.

First, there are a number of criteria that apply to the LO's involved, like the subject(s) they cover, the metadata scheme used, and the number of LO's currently available.

Another non-architectural distinction between LOR's is whether they provide some kind of personal service to the user. A simple example of this kind of service is a personal workspace with bookmarks to LO's that the user finds interesting. A more sophisticated example is a system that keeps track of user interests based on which LO's he actually downloads. Such a system can then notify the user when new such LO's are inserted. Another example is a system with which the user can create personal templates for the metadata of the LO's he introduces.

More important in the context of this paper are criteria that relate to the architectural properties of a LOR. One fundamental difference between LOR's is whether they are client-server based or follow a peer-to-peer approach. In this overview, we will only discuss the client-server based systems. (More details on a peer-to-peer based system we developed can be found in [3].)

The most fundamental distinction between client-server based systems is whether the LOR consists of a single server holding the metadata and the LO's (or the references to these documents), or that the LOR is distributed. When the LOR is distributed, are the metadata and/or documents being replicated or do different servers cooperate in handling a search request, as in for instance a 'federated search'. Another issue is the type of storage used for the metadata and the LO's. Are LO's stored in the server or are they distributed over the Internet? And in what format are the metadata stored? Very valuable is the possibility of interconnecting different systems: another important factor then is whether a LOR allows to propagate a search request to other LOR's, or whether it gathers LO's and metadata from other LOR's.

3. COMPARATIVE ANALYSIS OF LOR'S

ARIADNE is a LOR that has evolved from over 6 years of collaborative work [4,7]. It consists of a hierarchical network of replicating nodes, where metadata of all objects are replicated, as well as the free LO's. LOM instances can be validated. Work is ongoing in the area of automatic metadata generation [5].

SMETE supports searching and browsing resources in the domains of Science, Mathematics, Engineering and Technology [8]. A personal profile gives access to a workspace with bookmarks. LO's are recommended, based on past user interaction with the LOR. Members with similar interests can be identified. Federated searches that support querying remote SMETE LOR's, are under development.

The Learning Matrix [9] and iLumina [10] are other examples of LOR in the same subject domain. iLumina groups related LO's in so-called "virtual collections" [6]. HEAL [11] is a LOR focused on health science education and will soon be integrated in SMETE. They are doing research on automatic content harvesting.

MERLOT [12] covers all disciplines. Besides the MERLOT-Central LOR, which provides access to the entire collection, a number of discipline specific websites offer modifications of the central functionality and interfaces to meet the requirements of specific disciplines. Reviewing by peers is central to MERLOT's strategy. Discipline teams rate documents on 3 criteria: content quality, potential effectiveness and ease of use.

The LearnAlberta Online Curriculum Repository [13] is in its early stages of development: at the moment, few LO's are available, but new material is being created. The goal is to create a collection of LOR's in the field of Kindergarten to grade 12 education, with access through a set of linked portals.

CAREO [14] is a LOR that holds links to LO's, as well as some LO's themselves. The entry page displays the newest and most popular LO's. A personal profile gives access to a workspace ('My objects') with bookmarks. Users can access a history of objects they have downloaded. A component is under development to collect LO's and their metadata from other LOR's.

The EdNA LOR [15] can harvest LO's from remote LOR's and make its own LO's available to other sites, relying on the Open Archive Initiative (OAI) protocols [16]. LO's and metadata are collected from Australian Museums Online (AMOL) and there is a two-way exchange of metadata between the EdNA LOR and the Education Channel in Victoria.

Lydia Inc. [1] is a commercial network of LOR's. It has a central site (Lydia Global Repository or LGR) where users can search for LO's after registering. The metadata for each LO describe ownership and price, and users can use Lydia's transaction basket to buy objects. Organizations can set up their own Virtual Private Repository (VPR). A VPR allows organizations to introduce private content and provides access to all free content available on the central server. A VPR is actually located on the central site. Another option is to install an Enterprise Repository, which is a local server that offers full replication of the features of the Lydia Global Repository.

4. FEATURE COMPARISON OF REPOSITORIES

In the table below, we summarize the different LOR's we studied, with a comparative analysis of their features and characteristics.

	ARIADNE	**SMETE**	**Learning Matrix**	**iLumina**	**MERLOT**	**HEAL**	**CAREO**	**Learn-Alberta**	**EdnA**	**Lydia**
Organ-ization	Foundation	Federation (Berkeley)	ENC.	Project	Cooperation	US Nat. Science Foundation	Universitie s	Alberta Learning	non-profit	private org.
Meta-data scheme	IEEE LOM profile	IEEE LOM profile	IEEE LOM profile	IEEE LOM profile	IEEE LOM profile	IEEE LOM profile (CanCore)	IEEE LOM profile (CanCore)	IEEE LOM profile (CanCore)	Dublin Core profile	IEEE LOM profile (SCORM)
Subject domain	All	Science, mathematic s, engineering and technology	Science, mathematics, engineering and technology	Science, mathemati cs, engineerin g and technology .	All	Health science	All	Kinder-garten to grade 12 (K-12) education	Education	All
# LO's	2498	1645	170	880	7408	N/A	1576	?	15782	48
IPR mgmt	Free and restricted	Free	Free	Free	Free	Free	Free	Free	Free	Free and restricted

Simple/ advanced search/ Browsing	Simple/ Advanced	Simple/ Advanced/ Browse by discipline	Simple/ Advanced/ Browse by discipline or resource type	Simple/ Advanced/ Browse by metadata fields	Simple/ Advanced/ Browse by discipline	Search and browse	Simple/ Advanced/ Browse by discipline	N/A	Simple/ Advanced/ Browse by discipline	Simple/ Advanced
Peer review	Metadata validation	No	Yes	Plan to create peer review, rating and recommen dation	Review of content quality, effectiveness, ease of use	No	No	Yes	No	No
Personal Features	Metadata templates	Workspace, recommen-dations, communitie s	No	Plan to create workspace s	No	No	Workspace , download history	No	No	Transac-tion basket, purchase history
Distribution	Hierarch. knowledge pool syst.	Central server	Central server	Central server	Central server	Central server	Central server	Portal to reposito-ries	Central server	Global svr/ enterprise systems
Replication/ federated search	Replication of metadata, and free LO's	Federation	Federation	N/A	N/A	N/A	N/A	N/A	N/A	Replication
Metadata store	Oracle database for metadata	?	?	?	RDBMS, with export to XML	SQL Server 2000	?	?	?	?
LO store	Document Repository	Links	Links	Links	Links	Links	Document Repository + links	Document Repository	Links	Document Repository
Connection with other LOR's	Planned	API for federated search under developmen t	Federated search	No	Import of LOM metadata records planned	Under investigation	Collects objects from other LOR's	Portal which links to different databases	Open Archive Initiative under develop-ment	API
Other	Soon: Automatic metadata generation	N/A	N/A	Collections of related LO's	Discipline specific websites	N/A	Newest and most popular LO's	N/A	N/A	VPR's can share content

5. EVALUATION

As a conclusion to this overview, we will briefly discuss some benefits and drawbacks of different design decisions.

When an organization decides to start with the implementation of a LOR, the most obvious thing to decide on is the field of interest. The organization can adopt LOM as such, or it may prefer to develop its own application profile [2]. Such a custom metadata scheme is especially useful when the objective is to gather a more specialized collection of learning items.

Also fundamental is deciding on whether to manage Intellectual Property Rights (IPR). This means some objects are freely available and others require payment or are only accessible to a certain group of users.

Apart from these fundamental issues, one has to decide about the functionality provided to the users. It is recommended to implement both simple and advanced search, where the advanced search function allows users to impose criteria on all metadata fields. Very useful for new users of the system is a browse function, with which new users can get a quick impression of the repository's contents, or which allows to quickly find related resources.

A very useful feature is peer-reviewing. Peer-reviewing can facilitate the task of evaluating the quality of a resource when it appears in the result page of a query. Peer reviewing is a time-consuming activity, and to do it properly it requires some kind of management from the organization serving the repository.

Other useful functions are a personal workspace a user can fill with objects from the repository. A more advanced system could generate user profiles based on the user's download behavior and point the user's attention to material he would be interested in. Other interesting features are chat and discussion forums.

Apart from these functional aspects, care has to be taken to make the right decisions concerning the architectural design. The most fundamental architectural issue is the distinction between a client-server system and a peer-to-peer network. Client-server systems are easier to implement and maintain, and backing up objects and metadata is straightforward. P2P is less efficient than client-server, since queries have to be propagated over a number of peers. The largest advantage of a P2P network is that it is perfectly scalable, where in a client-server based solution new servers have to be installed in case of overload, which could take quite some implementation effort. A drawback of P2P is that updates of the system require all peers to be updated, and the system should be able to cope with peers of different versions.

When a client-server based approach is chosen, one can choose between a single server or a distributed approach. A single server approach (or an approach with one machine holding the metadata and another serving the resources) is the easiest to implement, but performance problems could arise when the number of users grows. A distributed setup can cope with this growing number of users if it is well-implemented, and a distributed setup could allow institutions to have their own repository. Such a private repository is also a good solution for companies that do not allow their employees to have free access to the Internet, or if they don't want their repository to be accessible from outside their private network. In that case the firewall would have to be configured so that their server could only connect to the central server in order to be able to replicate or to enable federated search.

If the preferred system design is client-server based and distributed, new options arise. Metadata and/or documents could be partially or fully replicated to the other sites. One could choose to replicate metadata and documents of all public objects and replicate only the metadata of the private objects, or not to replicate anything of the private material. Anther option is not to replicate anything, but to implement 'federated search'. This means that a query invoked on one of the servers will be propagated to other servers. When considering a federated search approach, care has to be taken that not all queries are federated to the same server(s), since this would introduce a severe bottleneck. Another issue is the filtering of result items which occur in more than one site within the search results. This problem especially arises when a hybrid approach is chosen where partial replication and federated search are combined in some way. Federated search could also become impossible for a number of sites if another one would crash.

For the type of storage one thinks immediately of a relational database. A drawback here is that metadata records, which are usually presented in an XML or RDF format, have to be parsed into SQL insert statements. Conversely, records stored in relational tables have to be converted into an XML document. Recently there exist some implementations of XML databases, but using one of these is a less efficient solution.

Two approaches can be distinguished regarding document storage. Most repositories link to documents spread over the Internet. This is an easy approach, but obliges users to make their resources available on some website. When a Document Repository is used, users don't need to worry about this. But what to do with learning objects that are already hosted on a web server? If both kinds of storage should be supported, one could opt for a hybrid approach where metadata instances can point to a web page or to the Document Repository.

For a repository to become useful it is crucial to obtain a 'critical mass' of learning material. In order to obtain this critical mass the ability for different systems to interconnect is becoming more and more important. Systems should allow other repositories to collect their free content, and tools should be developed that gather objects from other systems. Federated search across systems could be made possible. Of course this introduces interoperability problems, especially because all systems use their own application profile. But some of the implementations show that automated object collection is already possible. A specification for the interoperability of repositories is being developed by the Open Archive Initiative (OAI) [1].

Other fields of interest are the automatic generation of a number of metadata fields to facilitate the task of those inserting objects and generating metadata. Some file formats (e.g. Microsoft Office) store some metadata with the file (read more about this subject in [5]).

6. REFERENCES

[1] IEEE Learning Technology Standardization Committee, Draft Standard for Learning Object Metadata, 18 April 2001.

[2] E. Duval, *Metadata Standards: What, Who & Why*, The Journal of Universal Computer Science 7 (2001), no. 7, 591-601, Special Issue: I-Know 01 - International Conference on Knowledge Management.

[3] S. Ternier, E. Duval & P. Vandepitte. *LOMster: Peer-to-peer Learning Object Metadata*. Proceedings of EdMedia-2002, pp. 1942-1943, June 2002.

[4] E. Duval, E. Forte, K. Cardinaels, B. Verhoeven, R. Van Durm, K. Hendrikx, M. Wentland-Forte, N. Ebel, M. Macowicz, K. Warkentyne & F. Haenni, *The ARIADNE Knowledge Pool System: a Distributed Digital Library for Education*, CACM, Vol.44, No.5, pp.73-78, May 2001.

[5] K. Cardinaels, E. Duval & H. Olivié. *Issues in Automatic Learning Object Indexation*. Proceedings of EdMedia-2002, pp. 239-240, June 2002.

[6] Gary Geisler, *Creating Virtual Collections in Digital Libraries: Benefits and Implementation Issues*, Proceedings of the 2nd ACM/IEEE-CS joint conference on Digital Libraries, pp. 210-218, January 2002

[7] ARIADNE http://www.ariadne-eu.org

[8] SMETE http://www.smete.org/

[9] The Learning Matrix http://thelearningmatrix.enc.org

[10] iLumina http://turing.csc.uncwil.edu/ilumina/homePage.xml

[11] HEAL http://www.healcentral.org/index.htm

[12] MERLOT http://www.merlot.org/

[13] LearnAlberta http://www.learnalberta.ca

[14] CAREO http://careo.netera.ca/cgi-bin/WebObjects/Repository

[15] EdNA http://www.edna.edu.au

[16] The Open Archives Initiative http://www.openarchives.org/

[18] Lydia Inc. http://www.lydialearn.com/