Learning Object Repository Technologies for TeleLearning: The Evolution of POOL and CanCore

Griff Richards
TeleLearning-NCE,
Vancouver, Canada
richards@techbc.ca

Rory McGreal
Athabasca University,
Athabasca, Canada
rory@athabasca.ca

Norm Friesen
University of Alberta,
Edmonton, Canada
norm.freisen@ualberta.ca

Abstract

Repositories provide mechanisms to encourage the discovery, exchange and re-use of learning objects. This paper describes Portals for On-line Objects in Learning (POOL), a consortium project of the TeleLearning NCE to build a learning object repository scalable to the national level. Funded in part by the Canarie Learning Program, POOL contributes to the development of two focal technologies: "POOL POND and SPLASH" a distributed architecture for a peer-to-peer network of learning object repositories, and CanCore, a practical metadata protocol for cataloguing learning objects.

Keywords: Learning object repositories, CanCore, POOL, metadata.

Introduction

During the seven years of the TeleLearning NCE, the use of web technologies for on-line learning has become nearly ubiquitous. The acceptance of this technology by both learners and instructors as a convenient medium for educational transactions has been followed with an unprecedented investment in time and resources into the creation of materials and network infrastructure for distance and augmented learning. Digital learning objects are the computer files that store graphics, lessons, animations and other computer-mediated activities that constitute the content and process of this world of on-line learning. As knowledge nuggets in an e-learning gold rush, they represent an ever-increasing store of intellectual potential. Although many learning objects could be re-used in different instructional contexts, many of these investments are used for its highly specific audience but remain unknown beyond the immediate creators and consumers.

Repositories provide mechanisms to encourage the discovery, exchange and re-use of learning objects. This paper describes Portals for On-line Objects in Learning (POOL), a consortium project of the TeleLearning Network of Centres of Excellence and its partner organizations to build a learning object repository scalable to the national level. Funded in part by the Canarie Learning Program, POOL contributes to the development of two focal technologies: a distributed architecture for a peer-to-peer network of learning object repositories called "POOL POND and SPLASH", and CanCore, a practical metadata protocol for cataloguing learning objects. While these are new technologies, they are designed to co-exist and be compatible with other repository technologies. For example, POND represents a direct linking of POOL with pre-existing or third party object repositories, and CanCore is conformant with both ADL/SCORM and the IMS Global metatagging schemas.

Learning Objects – the Building Blocks of e-Learning

Although their definition varies by author and organization, learning objects are essentially the digi-

tal files that are used to generate e-learning activities. They include such things as audio-visual media files, java applets, and interactive exercises which make up the learner's experiences. Whether lumped into a generic category or split into categories such as "information objects" or "instructional objects" or "reusable learning objects", there is no dispute that digital learning objects are the basic building blocks of the e-learning economy. More detailed introductions to learning objects are provided by Downes (2001), Barritt & Lewis (2000) and Wylie (2001).

The promise of digital learning objects is re-usability. If constructed appropriately, warehoused wisely and catalogued accurately, a learning object might find usage beyond its original audience, and instructional context. Given the relatively high cost of developing good learning objects, the promise of reusability receives considerable attention from administrators and publishers trying to amortize the cost of production and maximize the potential return for each of these digital investments.

For educators, the promise of re-usability goes beyond the economic argument to encompass notions of quality, and the re-use of exemplary teaching strategies in other contexts. CAREO (www.careo.org) and MERLOT (www.merlot.org) are web portals founded partly on the premise that academic peer review of learning objects can improve the quality of learning objects and enhance the quality of on-line education. Learning objects are posted not just to advertise their availability, but also so others can observe the way they are crafted to suit the needs of the learners, see how they can be adapted into new instructional settings, or how the instructional strategies might serve as models for other content areas.

Early references to learning objects often over-simplified the notion of their being the "building blocks" of e-learning – interlocking elements to be combined in many creative ways to suit the needs of the learners. While this attractive analogy implies standardization is the key to interoperability, as with real building blocks, we can expect learning objects to come in many shapes and sizes, and commercial brands which for reasons of functionality, sophistication and competitive marketing will probably not all be compatible nor interlocking. Fortunately the learners, like children, will be oblivious to this fact and will integrate them into their learning experiences and use them in ways unimagined by the original designs and creators.

Learning Object Repositories – Digital Storehouses for Learning Objects

Repositories may be simply viewed as places to put digital objects. A central repository would be one that stores the objects for a defined community or organization. As objects can vary in number, size and file type, it is unlikely that a single central repository would be able to collect or even physically hold all of the available learning objects in any given field (Hamilton, 2001). As with libraries, organizations or communities may have something about everything or everything about something, but having everything about everything is unlikely. Thus, a decentralized or "distributed" model of a learning objects repository is a likely scenario. The key to a successful repository strategy will lie in the ability of repositories to share information and exchange records about learning objects, and their provision of access to the objects themselves.

Repositories might hold collections of learning objects as a warehouse might store books, or they could hold collections of information about learning objects as a library catalogue might hold descriptions about books. The catalogue descriptions are referred to as the "metadata". Some repositories may specialize in the type of information they carry; for example, the Australian AVIRE (Woodbury, 2001) contains only architectural objects, and the meta-data is specialized to the needs of the architectural community. MERLOT has a more open approach and welcomes information about learning objects in a wide range of content areas. MERLOT holds descriptions of learning objects, peer reviews of learning objects, lesson plans or assignments that use the learning objects, and in some cases marketing information about availability,

price and conditions of sale. A wide variety of services may be offered by a given repository based upon its mandate and the level of resources available to support the wants and needs of its user community.

Since not all repositories store the actual object files, a key function of repositories is to identify the storage location of the objects, and provide an indexing system that enables the efficient search and discovery of the objects. The way in which repositories accomplish the first is a function of their architecture; the latter is a function of their catalogue information or "metadata".

The Architecture of POOL, POND and SPLASH

"POOL, POND and SPLASH" evolved as a catch phrase to explain a distributed architecture that could flexibly meet the needs of many groups. Designed to support the individual instructor or learner SPLASH is conceived of as a small single-user repository that would be made freely available for download off the internet. SPLASH combines a database program and a peer-to-peer search engine with a CanCore meta-tagging interface. Built on Sun Microsystems JXTA platform (www.jxta.org), each SPLASH site holds those objects of immediate importance to the owner, and has the ability to search other SPLASH peers and, subject to permissions granted, to exchange learning objects or learning object metadata with other members of the network.

SPLASH development is partly driven by the notion that the most important place to hold a learning object is close to the developer and close to the user. SPLASH enables instructors, developers and learners to become consumers of, and contributors to the POOL of learning objects. While SPLASH allows individuals to collect and manage learning objects, perhaps creating portfolios of their personal learning experiences, and reducing the transience of the e-learning experience, the main success of SPLASH at the community level may simply be in its proliferation of desktop tools that encourage and assist the learning community to meta-tag their objects using CanCore and thus create a large virtual pool of otherwise undiscoverable learning objects.

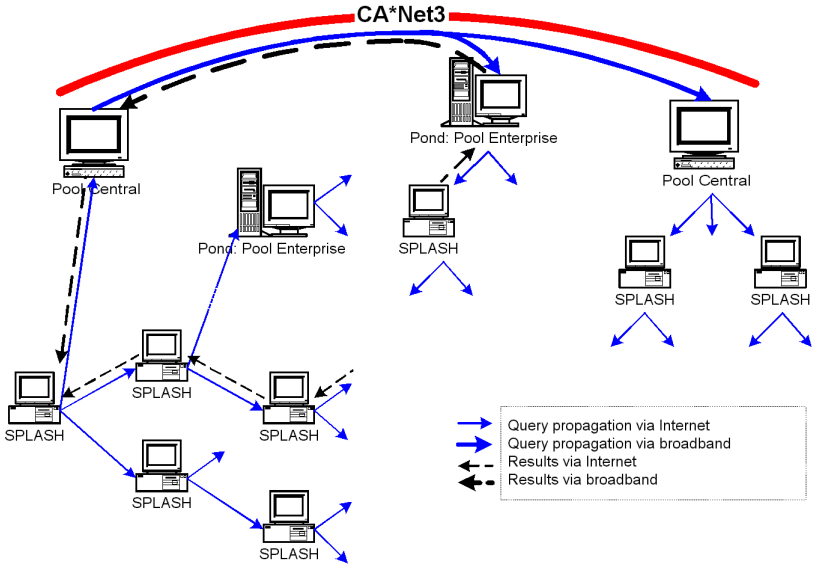

Fig. 1. POOL network architecture

Communities and organizations are a reality of the world of education and training. Ministries of education, universities, school boards, schools and employers are typical of organizations that will have an interest in providing access to their constituents with access to specific collections of learning objects. (They may also have interest in denying them access to other "unauthorized" learning objects.) These organizations will also have special needs to govern the access, workflow and life-cycle management of their learning objects. They may have access to financial and technical resources that will enable them to

build specialized and robust databases such as the POOL (www.newmic.com/pool) prototype developed by IBM Canada, or the CAREO (www.careo.org) repository developed in the Province of Alberta. While serving the defined needs of their communities, these sites also have the potential to become "PONDS" in our network architecture - community sites that primarily serve the interests of their clients, but through their interoperability with the POOL network provide their members with the opportunity to search beyond their immediate collection of learning objects.

Finally, a third level of aggregation, POOL Central, was devised to replicate search requests in topological regions of the repository network, and overcome the horizon effects that arise in decentralized peer-to-peer networks such as Gnutella. The designation of a number of "super nodes" could facilitate a faster and more exhaustive search of all of the member repositories via a high-speed and high bandwidth connection to the Ca*Net 3 optical highway that spans Canada.

The significance of POOL, POND and SPLASH is that it defines not so much a repository structure, as a method of linking repositories. Any repository of any size can be cross-searched simply by adding on the SPLASH search mechanism, adhering to the CanCore/SCORM metadata standard, and having a willingness to be included in the aggregate repository initiative.

CanCore – The Canadian Core Learning Resource Metadata Protocol

The Internet has become one of the best research tools for students and their teachers. Learners, however, are finding that there is just too much information online. This makes it very difficult to unearth and identify good, quality learning resources. As part of the POOL project, a national metadata protocol, known as CanCore (Canadian Core Learning Resource Metadata Protocol) has been created to help address this problem and facilitate the sharing and management of online educational materials or learning objects.

As has been noted, a learning object is a resource with an explicit educational application. It can be digital, for example: a simple Word, PDF or Word Perfect text, an e-book, or a Flash animation. Or, it can be physical like a text book or CD-ROM. For the purposes of interoperability, POOL is concerned only with those learning objects that can be accessed on the Internet. Metadata facilitates access to learning objects by providing a controlled and systematic way of describing each object. Metadata describes and links learning objects (Innes, McGreal, & Roberts, 2002).

Metadata, quite simply, is a label that is placed on any object. Common examples of metadata are the labels on cans of vegetables, the signs on shops, street addresses, and library cards. Metadata is the term used to describe a package of information about an electronic resource. It acts very much like a catalogue record in a library, describing the resource it refers to by providing information such as author, title, subject matter, copyright information, and location (McGreal & Roberts, 2001).

The need for metadata in educational repositories is becoming abundantly clear. The full text search approach currently widely used on the World Wide Web is proving inadequate for the location of high quality resources appropriate to specific learning contexts, levels and styles. Furthermore the number and range of non-text educational materials abounds as multimedia technologies and bandwidth for distributing them advance.

Managers of e-learning databases of all types quickly learn that metadata is essential to ensure that the information can be discovered and made accessible. With this knowledge, however, managers quickly realize that implementing a metadata model that is standards-based, interoperable, and widely used is no small feat. The CanCore (www.cancore.ca) protocol has been developed to facilitate such implementations, addressing the problems involved in applying metadata to e-learning content (Friesen, Roberts & McGreal, 2001).

CanCore has been developed as the metadata application profile for POOL and other CANARIE- sponsored projects. It is already becoming widely known in the Canadian and international e-learning communities (see http://www.cancore.ca). CanCore provides a streamlined version of the complex IMS (Instructional Management System) metadata standard (http://imsproject.org/), which forms the base for the SCORM (Sharable Courseware Object Reference Model) metadata standard (http://www.adlnet.org/Scorm/). CanCore is fully compliant with these international standards and with the emerging IEEE Learning Object Management protocol P1484.12 (http://ltsc.ieee.org/index.html).

The CanCore Protocol is a set of elements for the uniform description of modular, digital educational resources. These elements represent a subset of the data elements specified in the IMS Meta-data Information Model. The CanCore specification takes a middle-ground approach between the minimalism of the 15-element Dublin Core Metadata Initiative (DCMI) (http://purl.oclc.org/dc/) and the structural approach of IMS.

CanCore has taken only the active "core" elements from IMS, which are considered essential for learning object implementations. The researchers produced a set of formal guidelines explaining the purpose and use of each element and providing standard vocabularies where they are appropriate. This simplification provides developers of learning materials with a prescribed and recognized formula that can speed up the development of standards-based projects. By simplifying the IMS element set, and providing guidelines and assistance, the CanCore group hopes to save time for developers and projects wanting to adopt the metadata standard. The full IMS element set is not suited to direct implementation. Conformity with all 86 elements of the IMS metadata specification is often not taken up by vendors because of the huge job of classification it creates.

The CanCore Protocol has been developed to provide a common element-set for Canadian and international educational object repository projects. Besides POOL, these include other CANARIE-supported projects including the Broadband Enabled Lifelong Learning Environment (BELLE) (http://www.netera.ca/belle/), and LearnCanada (http://www.learncanada.ca/), and CAREO. New Brunswick's TeleCampus (http://telecampus.edu) is also moving to the use of CanCore. Funding and support for the development of the CanCore Protocol has been provided through these projects, and by the Netera Alliance, TeleEducation New Brunswick, and the Electronic Text Centre at the University of New Brunswick.

CanCore Rationale

CanCore has been considered to be an essential piece of the total POOL project from the beginning. The existing metadata solutions were considered either inadequate or too complex for a real world implementation. In implementing a distributed learning object repository project, interoperability among different content repositories is indispensable.

Neither the Dublin Core nor IMS specifications present ready-made metadata solutions for the collection and sharing of learning objects. The Dublin Core provides a "minimalist" set of 15 elements for the description of information resources in general, but does not provide elements for describing educational resources in particular. The IMS Metadata Information Model, meanwhile, takes a structuralist approach to metadata, and uses 86 elements to cover an extensive set of attributes specifically intended for learning objects. Even the IMS realizes that many implementers have no interest in developing products that require so many metadata elements.

Moreover, the IMS provides only the briefest descriptions of the purpose and character of each of its 86 metadata elements. (For example, element 1.3 general.catalogentry is described only as the "designation given to the resource".) Consequently, the actual implementation of the IMS element set is necessarily a complex, resource-intensive undertaking, requiring elements to be chosen, interpreted, used and tested by those sharing, collecting, or developing educational resources. Also, widely varying interpretations of the

utility, purpose and scope of individual elements threatens to cause considerable interoperability problems. CanCore provides a model or benchmark interpretation of the meaning, purpose and scope of 36 IMS fields that are considered important for promoting the interoperability of learning objects.

The 36 elements of the CanCore Protocol already save users the task of interpreting, selecting and coordinating the use of metadata elements to achieve a basic level of interoperability. As a Canadian and international initiative, CanCore presents the possibility of supporting further economies of scale, by discouraging the emergence of duplicate, redundant or inconsistent implementation efforts. It ensures that educational resources can be shared seamlessly across the Canada and internationally.

CanCore is continuing to develop and as it matures the CanCore team is creating support documents and services for a general audience of developers, designers and educators. This includes the development of classification vocabularies as well as training and promotion programs. These activities are promoting the expansion of the acceptability of CanCore and thus supporting enhanced interoperability. CanCore, like Dublin Core and IMS/SCORM is being implemented using RDF (Resource Description Framework) and XML (eXtensible Markup Language). This implementation is currently being tested on the TeleCampus, which houses the most comprehensive collection of metadata for learning objects available on the Internet, using the customized technical support applications developed by POOL.

The real world trials of the POOL applications and CanCore are being conducted using the TeleCampus and other online repositories. The TeleCampus currently has more than 55 000 online programmes, courses, modules, or lessons (McGreal, 2002). It is being adapted as a specialized POND node of called "CanLOM", the Canadian Learning Object Metadata Repository to gather additional learning object metadata from SPLASH sites and as such is presently implementing CanCore into its IMS/SCORM-compatible database structure. These trials involve indexing, developing, assessing, testing and implementing the TeleCampus directory within the POOL interface. TeleCampus personnel, as the creators of the first IMS-compatible metadata directory on the World Wide Web, have considerable experience interpreting, manipulating and expressing metadata in a large application. This experience also includes technical expertise in indexing, assessing, storage, data table relationships, administration interface development, and management knowledge.

The Future of Learning Repositories

POOL is but one of many international efforts to create learning object repositories. Others, such as MERLOT, and CAREO have been created to meet specific community efforts, and there is a growing abundance of LCMS (Learning Content Management Systems) in the commercial e-Learning market (Washburn, 1999).

CANARIE, which has sponsored POOL and other repository initiatives through its Learning Program also recognizes the need for convergence of effort, and has initiated a loose series of informal strategy sessions through the Canadian Repository Action Group. CRAG would see the regrouping of the various repository projects into a single pan-Canadian effort to create a national strategy for the advancement of learning object repositories (McLeod, 2001).

Canada is not alone in these efforts to build repository tools. Australia, Sweden and Holland are also moving rapidly in repository research and development. Indeed, the POOL team is in ongoing correspondence with a Swedish group building Edutella – a peer to peer model which is being built using the same JXTA platform as SPLASH (http://edutella.jxta.org/). We would hope to see convergence of these international efforts so that a universal repository model can emerge.

Learning object repositories are the catalogues of the e-Learning era. They will be the fundamental first step in knowledge discovery and object exchange. They will provide the foundation for future learning

and commerce in the knowledge market. They will fuel e-Learning as the stock exchanges fueled the industrial era. This is why they are of priority interest.

Acknowledgements

The original idea for a Portal for Online Object in Learning must be credited back to Nicholas Galan, Ian Dowdeswell and Joanne Curry then of TeleLearning, and Chuck Hamilton of IBM Canada who wanted to create an E-Bay for learning objects. As momentum for the project grew so did a wider recognition of general needs to develop technical specifications for a learning object repository. Rory McGreal and Terry Anderson (both now at Athabasca University), recognized the need for a common metatagging protocol and they facilitated the CanCore project with team members: Alan Burk and Sue Fisher, University of New Brunswick; Norm Friesen, University of Alberta; and Anthony Roberts, TeleEducation NB. As the design team met to scope the second phase of POOL, it grappled with the need of POOL to address a variety of end users. Griff Richards conceptualized POOL, POND and SPLASH as a way of fostering co-operation between repositories projects of common goals but different scale. Also contributing were Randy Bruce of C2T2, David Porter, then of Open Learning Agency, and Marek Hatala, who went on to lead the technical team at TechBC (Gordon Yip, April Ng and Timmy Eap) to produce the early version of SPLASH. Collaboration between large projects requires special leadership, and we have been fortunate to have the support of Tom Calvert of TechBC, Joanne Curry of TL-NCE, Fred Lake of NewMIC and Jamie Rossiter of CANARIE.

References

ARIADNE Alliance of Remote Instructional Authoring and Distribution Networks for Europe http://www.ariadne-eu.org/

Barritt, C and D. Lewis (2000) Reusable Learning Object Strategy http://www.cisco.com/warp/public/10/wwtraining/elearning/implement/rlo_strategy_v3-1.pdf

CAREO Campus Alberta Repository of Educational Objects http://www.careo.org

Downes, Stephen (2001), "Learning Objects," IRRODL (Volume 2, Number 1, 2001) http://www.irrodl.org/current.html

Friesen, N., T. Roberts, et al. (2001). CanCore brochure, TeleEducation NB/CAREO. 2002. WWW. Available: http://www.cancore.ca/

Innes, J., R. McGreal, et al. (2002). Metadata specifications. Handbook on information technologies for education and training. H. H. Adelsberger, B. Collis and J. M. Pawlowski. Stuttgart, Springer-Verlag**:** 273 - 288.

Hamilton, Chuck (2001) "Software Combinations for Learning Object Repositories" presentation at 2001 Canarie E-Learning Workshop, Toronto, October 3, 2001 http://www.canarie.ca/funding/learning/workshop_2000/meeting/Hamilton.ppt

McLeod, D. (2001). Presentation to the CANARIE Learning Program Senior Steering Committee. Ottawa, Canadian Repository Action Group.

McGreal, R. (2001). “A metadata primer: Fostering an interoperable environment.” E-Learning **2**(10): 26 - 29.

McGreal, R. (2002). Téléapprentissage et mondialisation de l'enseignement. Éducation. [in press].

Washburn, C. (1999). Evaluating integrated course management products. NAWeb 99. Fredericton, NB, Centre for Instructional Development and Distance Education.

MERLOT (www.merlot.org)

Wylie, David A. (2001). The Instructional Use of Learning Objects http://reusability.org/read/

Appendix: CanCore Metadata Element Set

1 general

- 1.1 identifier
- 1.2 title

- 1.3 catalogentry
  - 1.3.1 catalog
  - 1.3.2 entry
- 1.4 language
- 1.5 description
- 1.7 coverage

2 lifecycle

- 2.1 version
- 2.3 contribute
  - 2.3.1 role
  - 2.3.2 entity
- 2.3.3 date

3 metametadata

- 3.1 identifier
- 3.2 catalogentry
  - 3.2.1 catalog
  - 3.2.2 entry
- 3.3 contribute
  - 3.3.1 role
  - 3.3.2 entity
  - 3.3.3 date
- 3.4 metadatascheme
- 3.5 language

4 technical

- 4.1 format
- 4.2 size
- 4.3 location
- 4.6 otherplatformrequirements
- 4.7 duration

5 educational

- 5.2 learningresourcetype
- 5.5 intendedenduserrole
- 5.6 context
- 5.7 typicalagerange

5.11 language

6 rights

6.1 cost

6.2 copyrightandotherrestrictions

6.3 description

7 relation

7.1 kind

7.2 resource

7.2.1 identifier

7.2.3 catalogentry

8 (Not Allocated in CanCore)

9 classification

9.1 purpose

9.2 taxonpath

9.2.1 source

9.2.2 taxon

9.2.2.2 entry

9.4 keyword

See: www.cancore.org/schema.html

Biographies

Griff Richards is Research Integration Officer for the TeleLearning Network of Centres of Excellence, 2400 Surrey Place Mall, 10155 King George Highway, Surrey, BC, Canada V3T 1M8.

Rory McGreal is Associate Vice-President of Research for Athabasca University, Athabasca, Alberta, Canada.

Norm Friesen is a doctoral candidate in the Faculty of Education, University of Alberta, Edmonton, Alberta, Canada.