# Finding and Sharing eLearning Objects

***Standards Based Interoperability and Federated Repositories***


The Academic ADL Co-Lab
with support from
The William and Flora Hewlett Foundation

# Abstract

*For learning communities to effectively share their content with each other, interoperable repository search features and common addressing systems for locating and accessing content objects are necessary. The Academic Advanced Distributed Learning (AADL) Co-Laboratory gathers information about repository systems and provides opportunities for eLearning communities to share ideas and discuss the challenges faced by organizations involved with repository projects. This paper discusses federated repository concepts and development areas involving infrastructure, standards, and software tools that relate to the repository goals of eLearning communities.*


# Introduction

As more and more eLearning repositories are created to store and manage learning object content, end users will benefit by having effective ways to search against multiple repositories and conveniently access available content from them. While the material within different repositories may be varied in terms of data formats and instructional purposes, standardization of methods for networked searching is crucial for enabling interoperability among multiple silos of content objects. Competition and variation among content creators may be a good thing for driving innovation and improving content quality, but variation within high level metadata systems and search mechanisms will only make retrieval portals less effective for end users. A successful implementation for interacting with multiple repositories will:

- Support a wide variety of media types;
- Avoid restrictions on how a particular learning community can interact with their own repository content; and
- Collaborate with other repository systems for interoperable cross-repository searching.

Put another way, repository systems should have independence and versatility as well as interoperability and technological modularity.

# The Federated Repository Vision

Schools, organizations, and libraries use different information systems for managing their learning content, and feature a number of standalone delivery methods utilizing different technologies and tools. As the list of places vying to provide a home for learning content expands, variations in Web server file directories and content location tools lead to islands of content with little migration potential. To establish pathways among directories requires attention to both the functional and technical attributes of their systems. The federated repository concept takes a step back to expand the scope of collaborative systems, and promotes the use of modular components sharing services rather than idiosyncratically using only their own. [1] The vision of a federated repository system features a single portal for searching a number of different content repositories without having to navigate through different search engine interfaces. Only search results that are complete and also contextually useful are returned.

## *Web search engines — can't we just use Google?*

Several Web site search engines (Google, Yahoo, MSN, etc.) are available to help find information on the Internet by searching a database of Web pages for keywords that describe what is being looked for. Except for the imposition of advertising displays, these search engines are generally free to the Web-surfing public.

Search engines are created by a combination of sophisticated software tools (often called robots or spiders) that index pages from the Web, and humans who help by imposing order from the automated results. Search engines perform index building and results ranking in different ways, cataloging different

pieces of information about content and providing different user interfaces for searching. Despite their usefulness, Web search engines are generally inadequate for interacting with learning object repositories because the conceptual model of the Web that is built into these indexing/ranking systems is simpler than that of typical learning object systems. Some of the shortcomings are:

- Little, if any, quality ranking can be relied upon to differentiate search results.
- Web search engines can find only what is publicly exposed to the Web. A learning object within a repository will not necessarily be seen by the indexing process.
- A Web search engine might find references to a desired content object, but may not provide access to it.
- Web search engines do not currently differentiate learning objects or material designed for education from less focused material.
- Spiders are unable to utilize the descriptive information that a learning object repository is likely to have available.
- The Web currently has no digital rights management (DRM) sophistication.

### Improving Web search engine results

There are, however, data mechanisms that can be introduced to Web site documents to make them more accessible to search engines and increase the ranking for search results. Indexing engines usually look for <META> elements that define a comma-separated list of keywords or phrases that give a short description of the content. Search engines may present these keywords as the result of a search (see *figure 1* – Use of the <META> element to add meta-information for search engine robots). [2]

```
<HEAD>
  <TITLE>Finding and Sharing eLearning Objects</TITLE>
  <META  name="author" content="AADL Co-Lab">
  <META  name="keywords"
         content="repository,
                  federated content structures,
                  federated repositories,
                  learning objects">
  <META  name="desc"
         content="An overview of federated repository concepts
                   for eLearning collections">
</HEAD>
```

**Figure 1.**
Use of the HTML <META> element for providing search engine information.
*Source: W3C HTML 4.01 Specification.*

There are no limits or restrictions on how many or what kind of keywords and descriptions may be added to an HTML document, which is why Web searches often result in extremely diverse results. While innovative uses of the HTML <META> element certainly improve Web search results, interacting with a federated group of eLearning repositories requires a more robust approach.

# Metadata — the foundation for federated repositories

The descriptive information used to drive search and access services is typically called metadata. Learning object content must be adequately described by metadata for optimal search and access results. Metadata supports the two fundamental objectives for a successful federated repository system:

1. Discover and locate available learning objects.
2. Provide access to learning objects – with direct or indirect retrieval methods.

A robust system must have the appropriate metadata to support the processing tasks of identification, discovery, location, administration, selection, and retrieval.

## *Metadata quality*

Quality and accuracy of metadata is best determined by the owners of the content, but it is not unusual for metadata to be created by others. As explained in *Encouraging Reuse at the Source* (AADL) [3], metadata production is often a challenge for repository projects. Many repository projects depend upon the public rather than the content developers for metadata production. Making matters worse, metadata is often created within a particular repository (e.g., MERLOT [4]) using a local metadata system. The use cases presented in the AADL paper highlighted three key metadata recommendations for content creators:

- Metadata production should be possible at any point within the production process.
- It should be possible to generate metadata automatically from production processes.
- Versioning information should be generated automatically and independently during the production process, separate from the metadata creation.

While compiling metadata in the form of feedback from readers may be valuable for eLearning community interaction, the best source of metadata for searching and retrieving is the content creator.

Since interoperability is a goal of federated systems, the descriptive information must be bound into standards-conformant data schemes in order to be shared effectively by different system tools. This is most typically done with XML data markup, which is currently the most practical way to ensure that the shared metadata can be useful and coherent when distributed and aggregated. For successful system interoperability, a learning object metadata (LOM) scheme should be designed so that the data can be accurately interpreted and efficiently processed. Learning object systems already benefit from this kind of standardization within learning management systems by using metadata conventions specified by the SCORM reference model for sharable content objects.[5] Several metadata standards (IEEE-LOM, Dublin Core, OAI, METS) [6, 7, 8, 9] exist, and a federated repository system may support more than one metadata specification. At a Repositories Summit event hosted by the Academic ADL Co-Lab in 2003 [10] (reviewed in the *What We Mean When We Say "Repositories"* AADL paper [11]), two themes shaped the conference agenda: 1) maintaining quality content and 2) maintaining quality metadata with repository systems. While well-constructed metadata can assist with content retrieval and provide helpful contextual sorting and ranking feedback for search results, the inherit quality of content within a repository can not be controlled or determined by a federated system. That is, content quality issues remain the concern of eLearning communities.

# Building a Federated Infrastructure – Two Different Approaches

The goal of providing a single point of discovery and access for federated repositories is being pursued with different strategies. One, a registry approach, aims to collect metadata from contributing repositories into a central registry which in turn serves as a gateway to the content. Another approach focuses on a middleware strategy that relies upon an abstraction layer between application tools and the repositories being interacted with. Both methods rely upon quality metadata and interoperability, and both aim to bridge the worlds of content management technology and digital libraries.

## *CORDRA — a registry approach*

CORDRA stands for "Content Object Repository Discovery and Registration/Resolution Architecture" and is a registry-based approach to the challenge of dealing with the issues of discovery and access for a federated repository system. Reference models like SCORM [5] and metadata standards like IEEE-LOM [6] have done a great deal for improving upon ways to use sharable learning objects within a learning management system (LMS), and the CORDRA project intends to provide a reference model that enables effective search and access methods across federated repository systems.

The CORDRA concept began in 2003 as a U.S. Department of Defense (DoD) effort to establish a federated system for SCORM compliant learning content held by different branches of the DoD. This project, known as the ADL-Registry (ADL-R), is the first active CORDRA implementation project. The ADL-R development work is shared by the Advanced Distributed Learning Initiative (ADLI), the Corporation for National Research Initiatives (CNRI), and the Learning Systems Architecture Lab (LSAL). Future CORDRA activities include implementation plans for other communities. When multiple CORDRA registries are operational, the development of a Federated CORDRA, a federation of CORDRA federations, is anticipated.

### A repository of metadata: registering learning objects

The CORDRA infrastructure scheme is based on registering an organization's content in a way that will facilitate broad reuse of sharable learning content beyond the original organization. A CORDRA repository is essentially a collection of registry metadata. That is, the CORDRA implementation does not provide a home for the actual content, but rather catalogues the content in the form of a registry built from metadata records (see *figure 2* – The CORDRA model).

A CORDRA federated registry is built upon a prescribed metadata model. A metadata record for each learning object is deposited into the registry, which is in turn searched against by an application server to locate the registered content. The CORDRA reference model does not require a specific metadata binding model. Any metadata standard that satisfies the core discovery and access requirements can be used to build a CORDRA implementation. The ADL-R project uses the IEEE-LOM metadata model. Because the metadata records are organized as a registry-repository independently of the content-repository, options for filtering directory content or introducing levels of access privileges may be implemented.

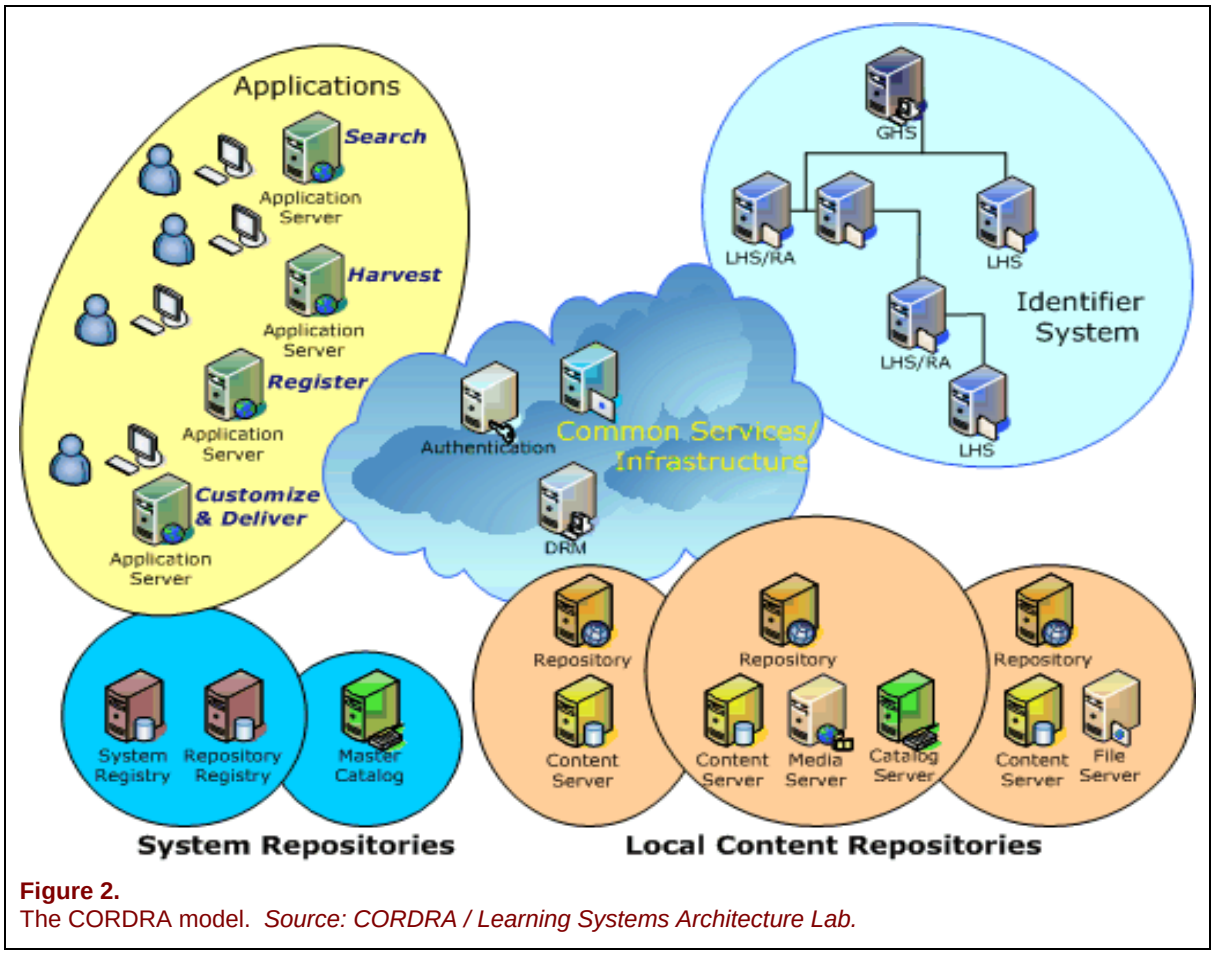


**Figure 2.**
The CORDRA model. *Source: CORDRA / Learning Systems Architecture Lab.*

## Locating and accessing content with CORDRA

CORDRA is about discovering, locating, and retrieving learning objects. Discovering that a learning object exists is one thing, but knowing where it can be found is another matter. While the CORDRA model does not require any one network addressing technology, the ADL-R project uses the Handle system. [12]

Handle identifiers (one for the content and one for the metadata record) are included in the registry metadata record and serve to uniquely label the content objects and the metadata records. The Handle system also provides the technology for locating intellectual property in the digital environment, providing a framework for managing content in any form at any level of granularity. As a network addressing agent, a Handle identifier serves as an unambiguous network address that can be re-directed to any Web location desired by the content owner. Possible re-direction results might be a direct path to the content, a password request, or perhaps a subscription fee notification.

## Combining interoperability standards with CORDRA

CORDRA is a reference model that defines how to build federated repository systems to support the discovery and access to learning content. The CORDRA model is designed to work with existing technologies, and is essentially a profile of interoperability standards. That is, CORDRA attempts to provide the blueprint for joining those interoperability standards together as a cohesive whole. [13, 14]

## *Open Knowledge Initiative — a middleware approach*

The Open Knowledge Initiative (O.K.I.) started at MIT in 2001. It supports innovators in online education by focusing on modularity and integration in a learning environment, thereby freeing them from the constraints of one-size-fits-all products. O.K.I. develops specifications that describe how the components of an educational software environment communicate with each other and with other enterprise systems.

By focusing on modular components and other dimensions of interoperability, O.K.I. seeks to allow independent development and the updating of components in a complex educational environment, and the integration of that educational environment with an enterprise infrastructure. The tools created by O.K.I. are the Open Source Interface Definition (OSID) specifications. The OSID concept is used to differentiate its service based specifications from the broader class of application programming interfaces. The O.K.I. architecture exposes a carefully selected collection of services, and the definition of these services enforces a programming model that maintains a sharply delineated boundary between O.K.I. compliant applications and O.K.I service implementations. OSIDs are interfaces, not implementations, which serve to enable integration among enterprise systems to exchange and synchronize information (see *figure 3* – The O.K.I. repository model).

### Federating repositories with OSIDs

There are several different OSID specifications for services such as messaging, course management, authentication, scheduling, workflow, and repository. The repository OSID provides an abstraction layer between application programmers and the enterprise infrastructure of a content repository. As a tightly defined set of methods, the OSID approach offers some important benefits to repository developers:

- Simple integration with existing infrastructure;
- Local innovations can be shared among different organizations; and
- Adaptation to new technology without destabilizing the overall environment.

The OSID focuses on services, not data formats. In that spirit, the OSID repository service specification is neutral with regard to metadata – Dublin Core, IEEE-LOM, or even proprietary metadata schemes may be used.

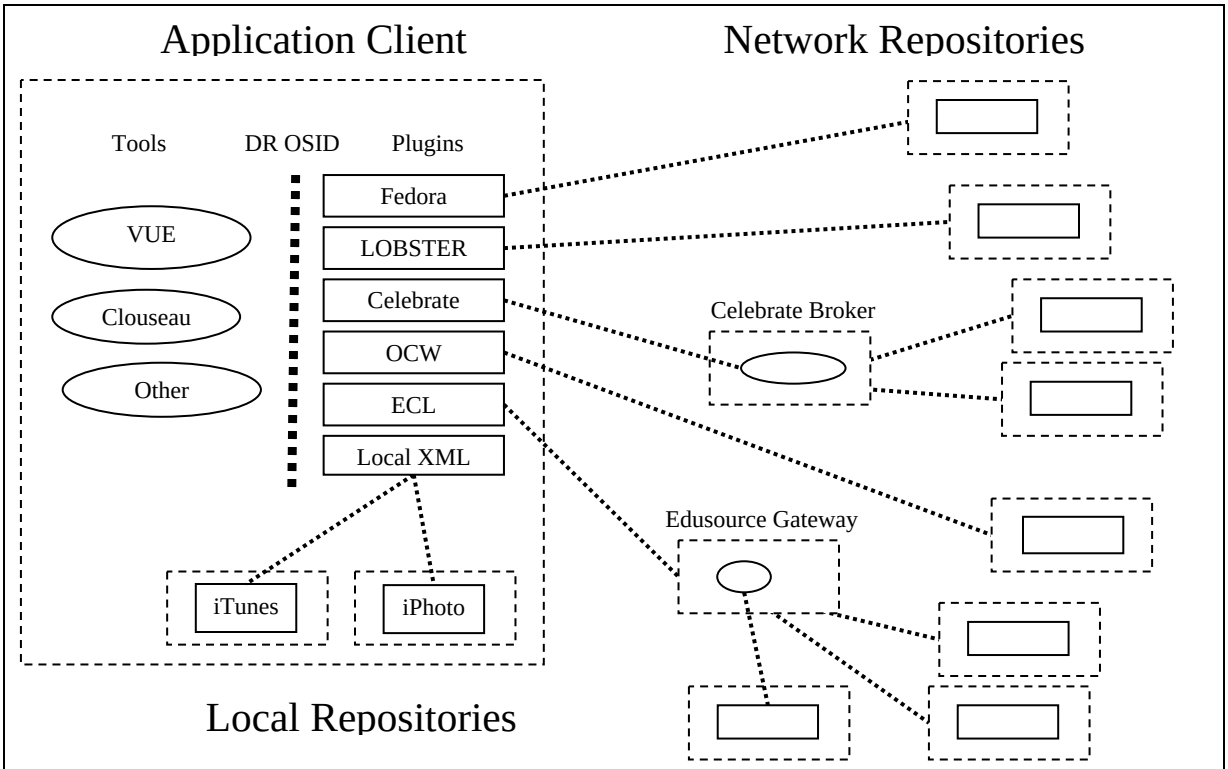


**Figure 3.**
The O.K.I. repository model. *Source: Open Knowledge Initiative: O.K.I. Overview*

While the OSID middleware concept is not an off-the-shelf solution, the implementation of O.K.I. federated repository systems will become increasingly attainable as industry affiliates recognize the value of a service-based approach that defines how components of an educational environment can work together. With those guidelines in place, software developers can build in OSID support thereby creating a plug-in menu of network solutions. The anticipated result will allow institutions to share innovative applications, take advantage of new technologies as they become available, and update components individually without destabilizing the overall environment.

O.K.I. has embraced the simple reality that things change in terms of services and data interchange functions and the OSID approach provides for a predictable evolution of expansion and change in the information exchange arena. Providing a middleware solution like the OSID between different software tools and different network repositories anticipates the reality that all systems do not have the same technologies, creating an environment where the need for sharing will continue to grow. [15]

# Summary: objectives of a federated repository system

The AADL paper *Content Repositories as eLearning Tools* [16] reviewed feedback from eLearning communities about requirements and priorities for learning object repositories and related network services. While some of the services important for eLearning communities are beyond the scope of the search and access goals of a federated system, the eLearning community needs for repository functionality includes the key motivations driving the technical work of projects like O.K.I. and CORDRA:

- Users want to easily discover learning content.
- Content creators want their content to be found.
- Researchers want to find the right content in context.

- Content queries should produce precision results, returning only what is needed.
- Flexibility and an approach that will scale are important for integration and interoperability with existing systems and applications.
- Participating eLearning communities should be able to enable local policies and business rules with no restrictions or interruptions to repository services.
- Ease of use is essential.

The building of successful federated repository systems depends upon interoperable implementation components so that a predictable evolution can allow for inevitable technology changes and also support expanding market possibilities. Standardized bridges between data formats and application services will allow for changing functionality and minimize negative impacts of inevitable technology change.

## References


1. Phillip D. Long, "Learning Object Repositories, Digital Repositories, and the Reusable Life of Course Content." http://www.campus-technology.com/print.asp?ID=9258
2. W3C World Wide Web Consortium, "HTML 4.01 Specification" http://www.w3.org/TR/REC-html40/
3. Academic ADL Co-Lab, *"Encouraging Reuse at the Source: Use Cases for Creating Metadata and Facilitating Content Management within a Development Environment."*
4. Merlot (Multimedia Educational Resource for Learning and Online Teaching). http://www.merlot.org/
5. SCORM-2004 (Sharable Content Object Reference Model). http://www.adlnet.org/index.cfm?fuseaction=scormabt
6. IEEE LOM 2002, IEEE Standard for Learning Object Metadata, IEEE Standards Department, Institute of Electrical and Electronic Engineers, Inc. http://ieeeltsc.org/
7. Dublin Core Metadata Initiative. http://www.dublincore.org/
8. Open Archives Initiative (OAI), http://www.openarchives.org/
9. Metadata Encoding and Transmission Standard (METS) http://www.loc.gov/standards/mets/
10. "From Local Challenges to a Global Community: Learning Repositories and the Global Learning Repositories Summit." http://www.academiccolab.org/resources/FinalSummitReport.pdf
11. Academic ADL Co-Lab, "What We Mean When We Say 'Repositories': User Expectations of Repository Systems."
12. The Handle System, http://handle.net/
13. CORDRA (Content Object Repository Discovery and Registration/Resolution Architecture). http://cordra.net/
14. Daniel Rehak, Philip Dodds, and Laurence Lannom, "A Model and Infrastructure for Federated Learning Content Repositories," http://cordra.lsal.cmu.edu/cordra/information/publications/2005/www2005/cordrawww2005.pdf
15. The Open Knowledge Initiative (O.K.I.), http:/www.okiproject.org/about/
16. Academic ADL Co-Lab, "Content Repositories as eLearning Tools: Community Building with Repository Services."