

Sakai and Library Content and Services

Introduction

The original Sakai development group (Indiana, MIT, Michigan and Stanford) has grown to include more than 50 Sakai Educational Partners (SEPP). Within the Sakai institutions are 16 Digital Library Federation and 29 Association of Research Libraries members. With the rapid growth of Sakai and the need to meet development deadlines, Sakai has yet to explore the complex issues surrounding the effective use of library content and services in Sakai. There is an active SEPP Library Discussion Group and substantial technical leadership across Sakai that could and should be used to realize overall library goals, once they are set.

We believe, however, that it is appropriate for Sakai libraries, and specifically the four directors from the original Sakai development group, to lead the effort to develop a library framework and to recommend the priorities for developing library content and services in the Sakai environment. This proposed work would build upon the Mellon-funded Flecker-McLean report¹, which explored the complexities of finding and using a growing body of digital library resources in course management systems and recommended that “the community move from theoretical discussions of interoperation of content repositories and instructional systems to real-world demonstration projects”. We endorse the concept of demonstration projects, but we also propose to implement fully some features in Sakai.

Proposed outcomes for January 18th meeting

The Andrew W. Mellon Foundation has organized a meeting of the directors and information technology representatives from the libraries at each of the Sakai core institutions for January 18, 2005 in New York City. The outcomes of this meeting should be to define goals and scope of work for incorporating library content and services into Sakai, and to develop an outline for a project to carry out that work, with rough phases of work, costs, and division of responsibilities.

User Studies

Much or even most of what we know about user needs is based on the experience of librarians who work with faculty and students. As valid as this informal data is, we propose to conduct user studies in the form of observation sessions, interviews, and/or focus groups with faculty, graduate students and undergraduate students, in order to validate our assumptions and to build a larger knowledge base. We might also take advantage of work taking place at Stanford in their ATS [Academic Technology Specialists Program] and at Indiana University in the context of the Variations2 project.

¹ Flecker, Dale and Neil McLean. “Digital Library Content and Course Management Systems: Issues of Interoperation” *Report of a Study Group*. Digital Library Federation (July 2004), <http://www.diglib.org/pubs/cmsd10407/>.

To add to our own studies, we also propose to conduct an environmental scan to build on the work already completed by some of our colleagues (see Appendix A).

Content Types

The directors agree that the following are the suggested content types to be considered as priorities for Sakai library development.

Licensed Electronic Content

One of the largest investments libraries have made during the past 20 years has been in providing access to licensed electronic content: databases, electronic journals, etc. Indeed, public services librarians confirm that licensed electronic content is the library resource most used by instructors in their classes. Thus, it seems most appropriate for one of the major goals to be to develop methods for faculty to search these resources and to provide links to appropriate articles and recommended resources to their students from within Sakai.

Digitized Content

Libraries have made substantial investments in digitization of resources including books, manuscripts, art images, photographs, musical scores, audio, and video. To enable their greatest use, these rich sources of pedagogical material should also become easily available through Sakai. In addition, having these materials available online through Sakai enables new forms of teaching and learning, including direct annotation and markup of audio and video by students and instructors for classroom presentations and assignments. Work should be undertaken to support searching and use of digitized objects in digital repositories by teachers and students from within Sakai, including the integration of value-added tools for annotation and markup of digital media objects for completion of assignments or quizzes.

Faculty and students are interested in using resources digitized by their own institution as well as resources digitized by other institutions, so integration with systems that aggregate digitized content and metadata from many institutions into focused collections, such as the Aquifer initiative of the Digital Library Federation and various OAI-PMH metadata harvesters, is important.

Open Access Content

A growing amount of “born digital” content of all types, including articles, working papers, technical reports, and data sets, is being made openly available through institutional repositories such as those based on the MIT/HP DSpace software platform, discipline-focused repositories such as arXiv.org, and open-access journals such as those published by the Public Library of Science. In addition to making licensed content and digitized content findable and usable through Sakai, this “open access” content should be integrated.

Learning Content

In addition to traditional resources such as books, articles, photographs, videos, sound recordings, etc., libraries will be playing an increasing role in the archiving of teaching and learning content created by instructors, including various types of course notes, course modules, and learning objects, some of which may be created in course management environments such as Sakai. An area of possible work with Sakai is exploring means of storing learning objects created by instructors in library-managed repositories and enabling their discovery and reuse.

Library Services and Library/Sakai Development

Libraries currently provide a variety of online services supporting access to content. Precisely how these services will best be used to integrate library content in Sakai will be determined in large part by the outcomes of additional user studies, but services of interest include the following:

- In academic libraries, librarians work with faculty to create course or subject pages with reading lists, web links, and links to appropriate databases geared to a specific course or department.
- Libraries provide a wide range of other services, such as online and chat reference, Interlibrary Loan, linking tools (such as SFX), bibliographic management tools, and federated searching tools to enhance access to library materials, either locally owned or available from other institutions or licensed resources.
- Other local search systems have been developed to search local image, audio, video, and electronic text collections.
- Library web sites are increasingly used to deliver both library services and content, e.g. course pages.

Appendix B provides information on additional standards and projects of potential relevance to the integration of library content and services within Sakai.

This paper was developed by:

Bill Gosling, University of Michigan
Michael Keller, Stanford University
Suzanne Thorin, Indiana University
Ann Wolpert, MIT

Jon Dunn, Indiana University
Susan Hollar, University of Michigan
Katherine Kott, Aquifer Project
Mark Norton, SEPP
Jerry Persons, Stanford University
MacKenzie Smith, MIT

APPENDIX A: Links to Information on Related User Studies

2004 Alt-i-Lab papers:

<http://www.imsglobal.org/altilab/index.cfm>;

<http://www.cetis.ac.uk/content2/20040729231332>

UC Berkeley Digital Resource Study:

<http://digitalresourcestudy.berkeley.edu/>

University of Hull Contextual Resource Evaluation Environment (CREE) project:

<http://www.hull.ac.uk/esig/cree/>

APPENDIX B: Relevant Technical Standards and Current Projects

1. Interfacing with digital repositories

- Repository Interfaces Layer – Johns Hopkins University Digital Knowledge Center (Sayeed Choudhury et. al.) <http://ldp.library.jhu.edu/repository.html>

From JHU’s Mellon grant proposal: “Given the variety of systems listed above, it is clear that some level of integration would be desirable. As the number of systems increases, so will the complexity of creating specific application-to-application interfaces. Creating an interfaces layer—a set of common services—offers the potential to reduce this complexity substantially. This layer would not need to specify the lowest common denominator of all the services supporting a given function, but would provide standard ways of interacting with the services that are available. These services could continue to expose their native interfaces independent of this layer.”

- CORDRA <http://cordra.lsal.cmu.edu/cordra/>

Content Object Repository Discovery and Registration/Resolution Architecture: An open, standards-based model for how to design and implement software systems for the purposes of discovery, sharing and reuse of learning content through the establishment of interoperable federations of learning content repositories.

CORDRA is designed to be an enabling model to bridge the worlds of learning content management and delivery, and content repositories and digital libraries. CORDRA aims to identify and specify (not develop) appropriate technologies and existing interoperability standards that can be combined into a reference model used to enable a learning content infrastructure.

CORDRA is:

- A formal model that can be used to design federations of repositories (the CORDRA reference model).
- A collection of operational systems built from the CORDRA model, including:
 - a prototype implementation of a repository federation; and
 - an operational *federation of federations* used to combined different CORDRA federations.
- The activities and projects surrounding the definition of the CORDRA model and creation of the operational systems.

The CORDRA activities are being coordinated by the [Advanced Distributed Learning Initiative](#), (ADL) the [Corporation for National Research Initiatives](#), (CNRI) and the [Learning Systems Architecture Lab](#)

(LSAL).

- D+ <http://devil.lib.ed.ac.uk:8080/dplus/>

The D+ project develops a software development toolkit (SDK) corresponding to 3 common services areas of the proposed services oriented [JISC e-Learning Framework](#), i.e. [federated search](#), [search](#) and [resolver](#) services. The SDK mediates the discovery of deep resources in distributed and heterogeneous repositories. Current list of [searchable repositories](#). The D+ SDK consists of an application framework that brokers the functionalities of the repositories, and enables the functions to be accessible via a unified set of web service interfaces. The functions are based on the reference models of the [IMS Digital Repositories Specification](#).

2. OAI in its various flavors and foci

- OAI-PMH – resource harvesting
<http://www.dlib.org/dlib/december04/vandesompel/12vandesompel.html>
- OAI-rights working group
<http://www.openarchives.org/news/news2.html#oai-rights>
<http://www.lboro.ac.uk/departments/ls/disresearch/romeo/>
- mod_oai – “OAI-PMH semantics to the web crawling community”
<http://www.modoi.org/>

3. Various flavors of wrappers/containers, i.e. the containers for citations and content that Sakai can/will read

- IMS CP ... content packaging
<http://www.cetis.ac.uk/lib/media/CPbrief.pdf>
- IMS Simple Sequencing
<http://www.cetis.ac.uk/lib/media/CPbrief.pdf>
- ADL SCORM [Sharable Content Object Reference Model]
http://www.cetis.ac.uk/lib/media/WhatIsSCORM_web.pdf
- IEEE – LOM [Learning Object Metadata]
<http://www.cetis.ac.uk/content2/20041015015543>
- IMS RLI [Resource List Interoperability]
<http://www.cetis.ac.uk/content2/20040708173446>
- RSS [Really Simple Syndication]
- A related tools project: RELOAD <http://www.reload.ac.uk/>
A JISC funded project (X4L strand B) developing tools to facilitate the use of emerging Learning Technology Interoperability specifications such as those produced by ADL and IMS.

4. Internet2 middleware – we need to move well beyond authentication and authorization (AuthN and AuthZ) <http://middleware.internet2.edu/>

- OpenSAML [Security Assertion Markup Language]
-- formation/exchange of authentication, attribute, and authorization data

- I2IM [Integrated Infrastructure for Instant Messaging]
- Signet [Privilege Mgmt]
- VidMid [Video Middleware]
- URN registry
- OID registry
- edu.Person, edu.Group, edu.CouseID
- Shibboleth
- HEPKI-TAG [Higher Education Public Key Infrastructure – Technical Activities Group]

5. Standards and tools for search and discovery

- Z39.50 International: Next Generation (ZING) SRW/SRU
<http://www.loc.gov/z3950/agency/zing/srw/>
- NISO Metasearch Initiative
http://www.niso.org/committees/MS_initiative.html

6. Libraries as Repositories of Learning Materials

- JORUM <http://www.jorum.ac.uk/>

The JISC Online Repository for Learning and Teaching Materials (JORUM) will be a repository service for all Further and Higher Education Institutions in the United Kingdom (UK), providing access to materials and encouraging the sharing, re-use and re-purposing of them between teaching staff. The JORUM project has been funded by the Joint Information Systems Committee (JISC), as part of the Exchange for Learning (X4L) Programme, to prepare for the JORUM service, which is scheduled to start in August 2005. The project has two main areas of work:

- Area 1: Preparation for the JORUM service
- Area 2: Support for the X4L Programme

- CWSpace <http://cwspace.mit.edu/> and DSpace@Cambridge

CWSpace is an [MIT iCampus project conducted by the MIT Libraries' Digital Library Research Group](#) and [MIT OpenCourseWare staff](#). The research is based on [MIT's DSpace](#) digital archive system, and takes as its principal focus the archiving of educational materials found in [MIT OpenCourseWare](#).

The project includes a [CSAIL](#) faculty research project investigating a variety of automated metadata generation techniques and tools.

In conjunction with the related MIT Libraries' project DSpace@Cambridge, sponsored by [CMI](#), the University of Cambridge's [CARET](#) (Centre for Applied Research in Educational Technologies) is a consulting partner in the selection, adaptation, and development of

standard protocols, data formats, and supporting software such as Web Services for interoperability with learning management systems.

7. Technical Architectures for Digital Object and Metadata Interoperability

- SIMILE <http://simile.mit.edu/>

SIMILE is a joint project conducted by the [W3C](#), [HP](#), [MIT Libraries](#), and [MIT CSAIL](#). SIMILE seeks to enhance inter-operability among digital assets, schemata/vocabularies/ontologies, metadata, and services. A key challenge is that the collections which must inter-operate are often distributed across individual, community, and institutional stores. We seek to be able to provide end-user services by drawing upon the assets, schemata/vocabularies/ontologies, and metadata held in such stores.

SIMILE will leverage and extend [DSpace](#), enhancing its support for arbitrary schemata and metadata, primarily through the application of RDF and [semantic web](#) techniques. The project also aims to implement a digital asset dissemination architecture based upon web standards. The dissemination architecture will provide a mechanism to add useful "views" to a particular digital artifact (i.e. asset, schema, or metadata instance), and bind those views to consuming services.

SIMILE has been working with metadata from both the library and the educational technology domains, and has prototype tools to demonstrate interoperability across such schemas as IMS LOM (as used by OpenCourseWare), VRA (as used by ARTStor), and variants such as the schema used by the Harvard VIA image repository.