Choosing Software for an Institutional Repository

Jody DeRidder, 1 April 2004

Institutional Repositories: what is the desired scope, and their future use?

Clifford Lynch, the Director for the Coalition for Networked Information (CNI), has a useful definition and vision for institutional repositories:

> …a university-based institutional repository is a set of services that a university offers to the members of its community for the management and dissemination of digital materials created by the institution and its community members. It is most essentially an organizational commitment to the stewardship of these digital materials, including long-term preservation where appropriate, as well as organization and access or distribution…. I believe that a mature and fully realized institutional repository will contain the intellectual works of faculty and students--both research and teaching materials--and also documentation of the activities of the institution itself in the form of records of events and performance and of the ongoing intellectual life of the institution.
>
> (Lynch 2003)

Including teaching materials implies learning objects, which can take a variety of forms. The educational community has perhaps already surpassed the digital library community in their development of digital object sharing (such as learning modules), for developing curriculum and courses. If we consider carefully what we may want out of our institutional repository in the future, and can choose a system which would integrate well with the hoped-for interoperability of digital libraries, it would save us headaches down the road and reduce our overall costs (in time, energy, and money).

As the Summary Report of the CNI Executive Roundtable on Institutional Repositories points out:

> It seems clear that there are a number of infrastructure components… that could serve a wide range of digital asset management applications, not just institutional repository systems, and we should be thinking about articulating and designing such services explicitly as infrastructure that will support this broader context…. A high priority should be …to try to ensure that disciplinary and institutional efforts are coordinated, follow common standards and architectures were possible, and evolve in a complementary fashion. We will need to understand how and when transitions of materials between institutional and disciplinary systems should take place.
>
> (Coalition for Networked Information 2003)

The consideration of the interoperability of an institutional repository, both with existing systems and potentially with future systems and other digital libraries, is a valid one. It will not serve us to waste our time and effort on a system with no future; and if we invite the academic community to commit their time and valuable resources to our repository, it would behoove us to be able to maintain the availability of their materials throughout future developments. We cannot afford to offend and disillusion our contributors.

With future interoperability in mind, I explored the proposed methods of sharing objects between digital libraries.

The May 1995 charter of the Digital Library Federation called for the development of a distributed, open digital library (DODL); in February 2003 this goal was reaffirmed, and an Initiative Committee appointed to establish functional specifications and appoint a technical subcommittee (*Federated Digital Library*, 2003). David Seaman (Director of the Digital Library Federation) has a vision for the deep sharing of digital objects in such a complete form that end-users could download the "master files" as malleable objects for local recombinations and a variety of uses; he states that we "need to exploit all available standards and technologies to infuse malleability, interoperability, and repurposing into the use of digital library objects (Seaman, 2003).

The Open Archival Information System (OAIS), which underlies the DODL initiative, spells out the business model for the sharing of digital objects: establishing a common framework of terms and concepts, identifying the basic functions (ingest, data management, archival storage, administration, access, preservation planning), and defining the information model and types of metadata needed (Day 2003, slide 5; and Consultative Committee for Space Data Systems 2002). OAIS was adopted in 2002 as an International Standards Organization Standard (NASA 2004).

Metadata: the three main contenders for sharing complex digital objects

The complexity of recommended metadata for OAIS implementation (OCLC/RLG 2002, 13) and the functionality described by David Seaman (Seaman, 2003) implies that METS (Metadata Encoding Transmission Standard) is the closest existing metadata format which will be functional for use in creating these distributed objects (and it will likely contain other complex metadata such as MIX: Metadata for Images in XML Schema, an evolving NISO (National Information Standards Organization) standard (*MIX* 2003), and MODS (Metadata Object Description Schema), which is based on MARC21 (*MODS*, 2004).

However, METS has competition from at least two other sources.

One recent contender against METS is MPEG-21 DIDL (multimedia Digital Item Declaration Language, developed for audio and video) (Burnett 2003, and Bekaert 2003).

> The vision for MPEG-21 is to define a multimedia framework to enable transparent and augmented use of multimedia resources across a wide range of networks and devices used by different communities.
>
> (ISO/IEC 2001, 5)

There is exploration already within the OAI (Open Archives Initiative) community to build on the existing OAI-PMH standards to support this metadata scheme in addition to simple Dublin Core for the sharing of objects (Van de Sompel 2003). MPEG-21 is the multimedia industry's answer to interoperability of digital objects, encompassing perhaps even more metadata than METS, but of a different variety (Bormans 2002). Thus far, the only repository in the states that appears to be working to implement this is the Los Alamos National Laboratory (Bekaert 2004).

A second competitor to the METS schema is the development of metadata specifically designed for learning objects: SCORM (Sharable Courseware Object Reference Model) (Advanced 2000) developed by the US Government's initiative in Advanced Distributed Learning (ADL). This is the educational industry's answer to interoperability of digital objects.

> The SCORM initiative is to develop common open-architecture specifications for 'net-centric' learning internationally. Its goal is to set specifications that focus on the ability to deliver and track content from multiple sources…. [SCORM] incorporates many emerging standards into one content model."
>
> (*Designing Courses* 2004).

An institutional repository can easily cross over into the realm of sharing learning objects, and many are developed for just this purpose. Learning objects are self-contained, reusable, small units of learning, that can be aggregated and are metadata tagged (Millar 2003); they may consist of one or several files. Some examples are: a time-lapse movie of a plant growing, an animated flow-chart describing an industrial process, a short-story told by a native speaker of a particular language, a java applet, an interactive computer simulation, or the description of a laboratory experiment (MacLaran, undated).The Learning Technology Standards Committee of the Institute of Electrical and Electronics Engineers (IEEE), who developed LOM, Learning Objects Metadata, on which SCORM is based, define a learning object as:

> …any entity, digital or non-digital, which can be used, re-used or referenced during technology supported learning. Examples of technology supported learning include computer-based training systems, interactive learning environments, intelligent computer-aided instruction systems, distance learning systems, and collaborative learning environments. Examples of Learning Objects include multimedia content, instructional content, learning objectives, instructional software and software tools, and persons, organizations, or events referenced during technology supported learning.
>
> (IEEE Learning Technology 2003).

Metadata standards for a LOM (IMSMD) have been developed by IMS (Instructional Management Systems) and the most recent version of SCORM (v1.2) references the use of the IMS Content Packaging specification v1.1.2 and the IMS Learning Resource Meta-data specification v1.2 (IMS 2004).

At least one person has tried to map the content of a learning object (IMS Content Package) into METS with mixed results: "If the goal is to support complete to-and-fro lossless translation, I think one will be disappointed. (It might be possible in that one might be able to wrap an object of one standard in the structure of another but one would not be able to meaningfully use the materials. If one wants to move the content structure from one system into another, the content-packaging ideas are roughly consonant [)]." (Yee 2003)

Rick Beaubien (also of UC Berkeley) expounds on the similarities and differences between IMS-CP and METS (and their interoperability) and states that the focus of IMS-CP is on structured presentation while METS focus is on structured content (Beaubien 2003: *BrainDump*).

Each community is developing the best metadata for the type of item with which it is concerned. While any digital item can be base-64 encoded and wrapped inside a METS record, the usefulness of that item on the receiving end will be determined by the existence and availability of software that can interpret it once unwrapped. Also, the resulting object would be "descriptive metadata poor from a library cataloging practice standpoint; and technical metadata poor from a preservation standpoint" (Beaubien 2003: *RicksBrain*). Still, Rick goes on to say, this would be more functional than trying to translate METS into IMS-CP (Ibid.).

By July 2003, the Dublin Core Metadata Initiative Education Working Group hopes to deliver a machine readable compound schema wedding both DCMI and IEEE LTSC LOM metadata (DC-Education 2003, Dublin Core 2003)

Another concern is whether our institutional repository will be interoperable with Blackboard and any current attempts at a UTK portal. If this functionality is desired, it would be helpful to involve ITC (Innovative Technology Center) personnel in the repository choice and development.

Blackboard (Release 6, the one in use at UTK) should support SCORM 1.2 (Blackboard, Inc. 2002). One document states that Blackboard "accepts SCORM-compliant Microsoft® LRN™ content" (Advanced 2000). Exploring further:

> Learning Resource iNterchange (LRN, pronounced 'Learn') is Microsoft's implementation of the IMS Global Learning Consortium (IMS) Content Packaging Specification, version 1.1, IMS Learning Resource Meta-data Specification, version 1.2 and the Advanced Distributed Learning (ADL) Initiative Sharable Content Object Reference Model (SCORM) Specification, version 1.2."
>
> (Microsoft 2001)

A note in a Blackboard document states that LRN Content files only run on Internet Explorer 5.0 or higher (Blackboard Inc.: *Microsoft LRN Content*). Therefore, it may be that implementation of interoperability with Blackboard will need to be custom configuration. Blackboard does allow developers to download and build on their API (with major legal caveats and limitations), creating "building blocks" to extend its functionality, but no changes or alterations may be made to its code as such (Blackboard 2004). Still, the option of building what is needed is available.

So what should determine our focus?

With all these considerations, and the evolving metadata standards, perhaps the best advice comes from Anne Kenney of Cornell University Library:

> There is a tendency to underestimate the importance of metadata costs, which can be half the total expenditure for a digitization project…The creation of appropriate metadata depends on anticipating needs of the user. Do we really understand very much about our users? Do users want what libraries are offering – large collections of mainly unmediated and uninterpreted digital objects? Or do they want what museums traditionally offer – interpreted objects in a context that tells a story? We need to pay more attention to identifying our audiences and to developing technological responses suited to their needs…The matter of selection – for whom and why – is crucial in developing a business plan, since every pixel caries a price… we must pay attention to the long-term implications of our present-day choices.
>
> (Kenney 2000,1-2)

One possible solution is to offer an online survey to faculty and staff, to determine their interests, focus, and future hopes for use of an institutional repository, and to identify collections of resources that may already be available. Some key questions to ask might be whether they want these items available via OPAC or Blackboard, and whether they will want, eventually, to access items from other institutions and to have our institution's items available from elsewhere. This approach would help to identify our client base, our user needs, promote the repository, and help to create institution-wide support.

While the university is intended to be our initial client, it is important to remember that integration with current systems and preparation for future sharing with other digital libraries, learning portals and web services, is critical for future use, metadata migration, and interoperability.

Institutional Repositories: the software out there now

The repositories in existence now range from enterprise content management systems to learning object sharing systems to simple ETD repositories. The primary concerns in this survey were the existence of the following parameters:

- simple clear web interface for users to enter metadata and upload resources
- preservation of (or at least on-site storage of) the user resources and metadata
- searching/browsing capabilities
- support for a variety of formats to at least be loaded up and offered for download during retrieval
- metadata sufficient for at least OAI records (whether a repository is built in or not), and hopefully for search capabilities beyond keyword
- functionality

The software examined has been divided somewhat arbitrarily according to the primary focus of its contents and use; the categories below are "software oriented toward gray literature and research", "software oriented toward learning objects", and "software oriented toward total content management". Since Fedora really doesn't fit any of these categories, it stands alone. And with the viewpoint that the more complex and complete the metadata, the more functional it will be for future interoperability as well as for preservation, it was the first one examined.

Fedora
(Flexible Extensible Digital Object and Repository Architecture)
http://www.fedora.info/

Fedora was created to implement the sharing and preservation of digital library objects using a profile of the METS metadata scheme. Of the software supporting METS, Fedora is currently the most functional, and definitely the best implemented on the library sharing level (*Mellon Fedora Technical Specification* 2003; and Payette and Lagoze 1998). However, it has not yet conquered the task of creating a user-friendly interface for upload and metadata creation by the document creator, and to a large extent, this is because the METS metadata scheme is so complex (*METS* 2003).

The technical docs for Fedora show this software to be rather remarkable and well thought out; items are even bound to separate disseminator objects and behavior methods to control display, use and access, with a software check in place to verify that the correct version is accessed; all export, preservation, and ingest methods are carefully documented, sensible, modular, and clear (Fedora 2003: *Mellon Fedora Technical Specification* 2003) . Should METS become the metadata standard of choice for digital libraries (quite likely at present), then Fedora is uniquely suited to underlie the Distributed Online Digital Library (DODL) of the future (*Digital Library Federation* 2003).

However, because of the complexity of METS records, the specifications for creating a METS-supported object in Fedora (or any other software) are incredibly difficult to implement for the casual untrained creator. If we want researchers to be able to describe and upload their resource via a web interface, Fedora would need serious augmentation. The software accepting the input (which has not yet been written, to my knowledge) would have to be transforming/augmenting that input metadata into a METS record, and transforming the resource into a base64-encoded bytestream, as well as storing it online

somewhere and inserting a corresponding URL to it into the METS record, for access/display as opposed to transmission (*METS: An Overview* 2003; *METS Version 1.2* 2003); it would also have to somehow extract the structure of the document and create links between the parts, then add appropriate dissemination and translation behaviors to the object, which would have to be predefined (for example, how to display it; how to serve it up on a web page; how to translate it into an OAI record; how to translate it into a MARC record) (*Mellon* 2003). Each different format (image, Word Doc, text file, html page, whatever) would have to have different handlers predefined and created that would work for each such object *without* variation or special treatment within object.

The Library of Congress attempted to create such a software for user web input of METS record metadata, but found it unworkable:

> Taken as a whole, data entry turned out [to] be somewhat cumbersome and labor intensive. The user has to visit many screens and click on many things to enter data for a complex object… There are other shortcomings beyond cumbersome data entry…. The people who enter data need to be familiar with METS, for example, they have to understand how a structMap ought to be set up.
>
> (Fleischauer 2003)

And this scenario entirely omits the need to create different levels of accessibility and authorization, and does not even address creating a layer for searchability, access and retrieval.

One exploration into the creation of a METS record from user description of an item is evolving in the EVIADA (Ethnomusicological Video for Instruction and Analysis Digital Archive) at Indiana University Digital Library.

> The depositor will enter explanatory information about the entire video and about individual scenes. Such annotation will include scene and action descriptions, organization of scenes into chapters, including transcriptions or translations, and creation of higher-level analytical and organizational categories. During this process, the depositor will work with a cataloger to ensure consistency in the use of descriptions, subject categories, and keywords.
>
> (EVIA 2002, 17)

According to one of the programmers on the project, this information only provides one portion of the data for the METS record:

> The interfaces we are currently working on, will allow the ethnomusicologist to segment their collection into a hierarchy of Events/Scenes/Actions and adding annotations at each level.
>
> From the viewpoint of the METS document, the Event/Scene/Action hierarchy is captured in the StructMap area of METS. The annotation text which is both, controlled and free-form, is converted to a MODS object and stored in the DmdSec section. It is linked with the appropriate Div in the StructMap section. So, the descriptive information and the logical structure are added to the collection METS document using the Annotation interfaces.
>
> Apart from this descriptive information, information about the files and their

> technical/provenance information will also be part of METS. We plan to gather this information during the digitization stages, but we don't have a framework in place right now on how this is exactly done."

(Vaidya 2004)

This is an example of the kind of work it would take to set up the back end of an interface for a user to input a particular kind of complex object (videos) into Fedora. We would have to have similar software and processes in place to gather the requisite metadata information of each type of item being uploaded into our repository.

Software Oriented Towards Gray Literature and Research

DSpace

http://www.dspace.org/

When asked about the way in which they hope to implement METS in the next upgrade, MacKenzie Smith answers that they (she and Rob Tansley) are working on a METS exporter and plan to implement an importer as well. They are transforming the DC to MODS for the descriptive portion of the METS record, and basically not using the structmap portion of the record at all (Smith 14 March 2004: *Re: METS*).

According to Leslie Myrick, Digital Library Programmer/Analyst of NYU Digital Library Team (the creators of METS), the structmap section is the heart of the METS record and the only required section, so use of METS for documents for which no structure is provided is really overkill. Also, extraction of useful structure from the document itself is limited at present, both by what information is stored within the documents (for example, a Word Document will store the name input by the software installer name as the author of the document) and by the limitations of tools currently developed for extraction (Myrick, 15 March 2004).

So, for the moment, there is a gap between institutional repositories and the creation of effective METS records for the uploaded documents, so that they can effectively be shared across libraries as envisioned by David Seaman in his discussion of DODLs. However, there is movement in the Adobe community to develop XMP (eXtensible Metadata Platform), a mix of RDF (Resource Description Framework) and XML, used to create and store metadata within resource files such as PDFs (Adobe 2004). Further changes in this area may well support our ability to create functional and fairly complete METS records from uploaded PDFs with a minimum of additional user-provided metadata.

DSpace is based on a profiled version of Dublin Core metadata (MIT 2003: *Metadata*); it supports a friendly user interface for metadata entry and file upload; communities can be configured and restricted in access, and collections within communities can also be restricted. Reviewers and accepters can be assigned to moderate the incoming submissions. The software is written in non-modular Java and currently based on a

PostgreSQL-centric database. The upcoming 1.2 version of DSpace (beta version due out in April) is supposed to support the capability of creating arbitrary hierarchies of communities (Bass 2004) which may be useful in organizing some structure in the holdings; currently the structure of holdings is limited to communities and collections within those communities (Bass 2002, 1). User-configurable portal access is not supported. Hopefully this new version will support full-text indexing and better support for HTML display (referenced files, though included in the upload, are not viewable online); and hopefully, the ability to assign distributed administrators for each community. Still, the software will not yet be modular; version 1.2 will still contain MIT-specific code, and extraction of items is non-trivial. There are, however, at least 2 groups exploring how to combine DSpace with software that shares objects more readily:

> http://www.npaci.edu/DICE/SRB/ (Storage Research Broker, making objects available via of federated grid technologies, San Diego Supercomputer Center, Chris Frymann, presenter) and http://www.oclc.org/research/software/srw (Search and Retrieve Web service, based on SOAP and WSDL, OCLC, Ralph LeVan, presenter)
>
> (Simpson 2004; *Notes*)

This author also thinks that "indexing and searching directly inside DSpace will always be a secondary function at best", with archiving, browsing and retrieval the primary function (Simpson 2004; *Thoughts*).

A plugin has been written to enable DSpace to harvest from OAI repositories and offer the records via the standard interface (*OAI harvester Plugin for DSpace* Undated), but there is one issue with OAI not yet addressed. The Dublin Core creator field is not used in MIT's DSpace; all authors go under contributor (Branschofsky 2003) to give them equal billing. To alter this in order to meet the standards of OAI, additional programming is required by the implementer (Smith, 14 March 2004: *Re: Subversive*). More appropriate would be a revision of DSpace to allow for more than one author in the creator field, in accordance with the NDLTD (Networked Digital Library of Theses and Dissertations) interoperability metadata standard (Atkins et al. 2003). Theses Alive! dealt with this and other issues by building their own EUL-DSpace Add-On (Edinburgh undated), for use solely with their ETD submissions. Here's a metadata comparison they did between the two parts: http://www.thesesalive.ac.uk/archive/MetadataSchemas.pdf.

DSpace is currently implemented by MIT, and all changes are tightly controlled by a very small and close-mouthed development team. According to Jason Simms at UTK Sunsite, the documentation is terrible, and implementers are forced to depend on one another's help since the support from the developers is so poor (Simms 2004). Hopefully, DSpace development will soon become open-source (Smith 19 March 2004).

The remainder of the repositories I examined, with the exception of Digitool, do not currently support, nor *appear* to be working towards support of METS, MPEG-21 or

SCORM for future interoperability. A comparison of seven of these at a primary level is available online (Open Society 2004).

EPRINTS
http://software.eprints.org/handbook/

Eprints is perhaps the most widely used institutional repository of research documents, with 143 live repositories, including Indiana and Virginia Tech Universities (Eprints.org 2004); it was "designed more specifically for institutional or disciplinary repositories of papers, as opposed to arbitrary digital materials (Lynch)." The types supported include articles, book sections, monographs, conference or workshop papers, books, theses, and patents (Gutteridge 2003: *GNU*); the primary formats are HTML, PDF, Postscript, and ASCII, with associated files and URLs.

CDSware
http://cdsware.cern.ch/
Online demonstration: http://cdsware.cern.ch/DEMO/

This was developed by CERN, the European Organization for Nuclear Research, and is available for open source download. It uses MySQL and apache, PHP and Python, runs on Linux, and runs an OAI data and service provider. CDSware offers configurable portal-like interfaces, google-like searching syntax, electronic submission and upload of a variety of documents. The metadata is based on MARC21 metadata standard (it extracts DC xml from MARC xml); a metadata example can be viewed at http://ditdevpc1.epfl.ch//search.py?recid=304&ln=en&of=hm. Searching is combined metadata/reference/full-text searching. Types accepted include articles, preprints, theses, books, video, conferences, multimedia, journals, photos, and more. Text submissions can be PDF or postscript (which gets converted to PDF on the fly), and the support docs say they also accept Tex/LaTex and Word. Upload of item is not required for submission. Support files can also be Word docs or compressed postscript. Maximum and minimum sizes for files are enforced.

Image file choices are jpeg and PDF; additional files may be added. At upload, user is asked for category of image: Diagrams and charts, experiments and tracks, Life at Cern, or Personalities and History of Cern (so this is configurable). Metadata requested is title, author, date of creation, description, keywords, and a couple of optional fields: author's comments and other (local) reference numbers. Each submission is assigned a unique id. The uploader can choose to submit text or image with or without simple refereeing, and revisions are accepted, so versioning is supported.

This one probably falls between Hyperwave and Eprints or Dspace, as far as usability and extensibility; better than Dspace (DSpace does not yet offer full-text searching or MARC) and not as fancy as Hyperwave. Fairly complete user help pages are available (Atlantis 2002)

ARNO:
(Academic Research in the Netherlands Online)
Software available from http://arno.uvt.nl/~arno/arnodist/.

The ARNO project is funded by IWI (Innovation in Scientific Information Supply). Project participants are the University of Amsterdam, Tilburg University and the University of Twente.

ARNO is a perl-cgi script supported by a relational database. It offers the choice of archiving the resource or merely entering the metadata and URL; it also supports peer review and moderator evaluation of submitted documents and metadata prior to archiving. OAI is supported. There is no division of communities or groups among the contributors. Versioning of documents is supported, and it accepts uploads based on known mime types (Bentum, Brandsma, Place and Roes 2001). ARNO can be configured for LDAP authentication, and provides persistent URLs. The OAI repository is built from the Virginia Tech Perl OAI-PMH software (*ARNO – install* 2003).

Metadata formats supported are configurable and extensible. Requires additional open-source software: for search and browse, iPort is recommended, for which the only documentation seems to be in Dutch (iPort, Undated); and for peer review and workflow, they recommend Metis, a java implementation developed by the University of Colorado (University of Colorado 2003). Bulk uploads are available, as is command-line interface; the system base is Linux/Unix (SURF 2003 *Appendix C*). Most of the documentation is in Dutch.

bepress
(Berkeley Electronic Press)
http://www.bepress.com

University of California's E-Scholarship repository (http://escholarship.cdlib.org/) is powered by bepress, as are repositories at Boston College (http://dissertations.bc.edu/) and New England Law Library Consortium (http://lsr.nellco.org/).

Edikit® is the software at the heart of the bepress system, and a license must be purchased to be able to obtain it (Internet Journal 2003 *bepress Services*). It is apparently very extensible and configurable, and bepress assists in the implementation of customized repositories. Some of the features offered are:

- Automatic conversion of documents to PDF
- Ability to publish HTML
- Ability to publish non-static resources such as sound and video files, data sets, and executables
- Full-text searching

- Saved Searches
- Personalized email notification of newly published content
- Browsing by date or author
- GUI-based template editor
- HTML-based templates
- Customized controlled-vocabulary picklists for data entry Site customization at any level
- Auto-image generation
- Branded publication sites for participating research units
- Automated email interface between author and publication administrator
- Usage statistics at the publication and paper level
- Flexible document hierarchy
- "Push" email capabilities
- OAI compliant
- Data exporting as XML
- Data transfer to third party indexing services

Furthermore, bepress licenses include hosting, offsite backup, technical support, training, and upgrades to the software. Ownership of all content is maintained by the licensee."

(Internet Journals 2003 *bepress Repository Technology*)

ETD-db
(Electronic Theses and Dissertations Database)
http://scholar.lib.vt.edu/ETD-db/

This is a fairly simple software, using web pages, perl, and MySQL; it was created solely for ETDs, and the metadata is ETD-specific. Browsing is currently available by author name and department; boolean searching is available among metadata fields. It is not currently configured to feed data into an OAI repository. ETD-db was developed at Virginia Tech and is free to members of the NDLTD (Networked Digital Library of Theses and Dissertations, which includes UT) (Digital Library and Archives 2003: *Welcome*).

An additional script is available to output records in MARC, and they use these records with Zgate (a free Z39.50 server) to allow Z39.50 searching of records (Digital Library and Archives 2003: *About the Non-CGI Scripts*).

The following five repositories are probably not useful for our purposes, but are included for informational purposes.

OPUS
(Online Publications)
http://elib.uni-stuttgart.de/opus/doku/english/index_english.php

This system is being tested by various German universities, and further development now in the hands of the University of Stuttgart. It is not open source at present. It was created for research papers which are displayed to users in PDF format, using DC (Dublin Core) with an mSQL database (Universität Stuttgart 2003: *About OPUS*). Authors may upload several files for archiving, but must also create and upload a PDF for use in the system. (Universität Stuttgart 2003: *Electronic Dissertations*). OPUS does not appear to be open source.

Greenstone
http://www.greenstone.org/cgi-bin/library

Greenstone is not set up for users to upload their own documents; an administrator would have to create a collection, define its parameters, and authorize users to add to those collections. Users could only add items to a collection that have the same format as the existing documents, and the same metadata. The only export method is to a "self-contained, self-installing Greenstone CD-ROM for Windows." (Witten and Boddie 2003, 31-32). Because of the limitations for casual users, this does not appear to be functional as institutional repository software.

i-Tor
(Innovative Technology for Open Repositories)
http://www.i-tor.org/en/toon

According to their online documentation, no metadata is required (*iTor… More Advantages* 2003), so all searching would be by keyword except for whatever subset of records were entered in with a structured metadata (*iTor… Features* 2003). Basically i-Tor offers a webcrawler service over specified webpages for which users have chosen the links; only if they enter dc metadata for the item will it be included in "advanced search" of chosen fields and in an OAI repository (*iTor… Applications* 2003). Since there is no preservation of materials, this does not appear to be useful for our purposes.

MERLOT
(Multimedia Educational Resource for Learning and Online Teaching)
http://www.merlot.org/Home.po

This software offers browse capabilities (sorted by communities or by broad subject area), no search capabilities, and no preservation of materials. Metadata is sparse, and items are accessed by linking to their location elsewhere online. Items are, however, peer reviewed (MERLOT 2002).

MyCoRe
(My Content Repository)
http://www.mycore.de/engl/index.html

This is an open source java serverside application designed (currently) to run on Windows, using IBM Content Manager and IBM DB2 database. It offers an OAI repository functionality and a Z39.50 interface for exchanging objects, as well as distributed search capabilities and workflow functions, with a hierarchical classification system (MyCoRe: *Project Description*). As all the documentation I could find is in German, I did not determine the metadata fields expected of contributors, nor the limitations on upload formats. Since this software runs only on Windows, it is probably not useful for our purposes.

Software Oriented Toward Learning Objects

Connexions

http://cnx.rice.edu/

This software is open source, developed primarily for open-licensed educational materials in fields such as music, psychology, and electrical engineering. Modules are browseable by author, title, keyword, courses. Each module may consist of several files. Provides authoring tools for creators to use to develop modules in XML; the modules are then stored in a common area, where they can be used, updated, and adapted to the user's needs. Collaboration is supported; browsing by author, title, keyword, and course. Offers links to similar content, other content by this author, and courses containing this content (Rice 2004).

This software stores user stores content in a custom xml format, CNXML, using MathML within certain tags, in order to display mathematical equations and symbols properly (Hendricks and Galvan, undated). This is the first real concern I have seen with managing display of diacritics or mathematic and scientific symbols, which will be problematic in any institutional repository. Requiring all textual content to be in xml promotes interoperability and ease of use and migration; creation online is supported, and both DTD and schema are available for download. Addition of images and other types of supporting files is optional. Metadata requested for new items is not extensive (for a new course: title, institution name, course code, instructor name, course home page, keywords and abstract), but the actual creation of the files requires extensive tutorial and online examples to instruct the creator in the use of xml (Hendricks, Galvan and Husband 2004), which can be daunting.

intraLibrary

http://www.intrallect.com/products/index.htm

This is a commercial product from a company based in Scotland. intraLibrary supports storage and retrieval, indexing and searching, harvesting and cross-searching across distributed collections, and personal collections of learning objects. It provides support for any digital format, and import and export of content packages conformant to IMS and SCORM specifications (Intrallect *intraLibrary*).

BELTS
(Basic E-Learning Tool Set)
http://belts.sourceforge.net/

Based on eXist (an open source XML database) as well as PostgreSQL and Jetty (a Java server), BELTS is written in Java. It appears to be oriented toward teachers sharing learning materials, and making use of them to create activities and curriculum. BELTS is not yet a fully-featured system, but does support basic activity creation, downloading and object replication between BELTS systems (The le@rning Federation 2004). A learning object can be any type of digital asset, including URLs, and the metadata supported is a particular profile built on the LOM (Australian 2003, 13-14).

The following two software repositories are probably not useful for our purposes, but are included for informational purposes.

Maricopa
http://www.mcli.dist.maricopa.edu/mlx/index.php

MLX is designed for sharing "learning objects", uploaded by faculty only. Searching and access is available to all, and searching is based on keywords within metadata fields, and by college; sorting allowed by newest and by title. The metadata fields include: Item (title), contact, credits, college(s), discipline(s) (for which this might be useful), summary, details, web links (to access the item), supplements, trackback (to note where on the web this has been used), stats (last modified, when created), and comments by users (A sample "packing slip" can be viewed at: http://www.mcli.dist.maricopa.edu/mlx/slip.php?item=121).

Objects are not actually preserved in the database. Once a year, contributors are emailed to verify that they still want to include their item in the database. Links are not checked or maintained, except by the contributor (McNally 2000). Because of the lack of preservation, this is probably not useful for our purposes.

ALOHA and CAREO
(Advanced Learning Object Hub Application)
http://aloha.netera.ca/about.php
(Campus Alberta Repository of Educational Objects)
http://careo.netera.ca/careo/userguide.pdf

ALOHA is a MySQL, JAVA-based XML indexing and searching service, supporting both IMS and Dublin Core metadata (and any others for which a schema is available; also, all upload formats are supported). It is designed to run on Windows 9x/2000/XP, Macintosh OS/Linux, with Java 1.41 or higher and Java Website installed.

> ALOHA is a piece of software designed for indexing, sharing, multi-purposing and repurposing learning objects. It is created to meet the needs of indexers, educators and learners, and includes versatile and powerful indexing tools and flexible searching of multiple educational object repositories.
>
> (*ALOHA* 2003, 3)

A new version is coming out (http://apollo.ucalgary.ca/) which has only recently been announced at this spring's New Media Centers (NMC) conference (Mattson 2004).

APOLLO is open source, and often uses CAREO as a front end (client to its server). CAREO is not open source, but source is available to partner institutions (Norman 2003); it was undertaken jointly by the Universities of Alberta and Calgary (National Library of Canada 2001). CAREO uses MySql, PHP, with IMS and MPEG-7 metadata (MaGee 2001).

CAREO appears to be of a level of difficulty of use that is good for catalogers and trained staff for uploading, and metadata entry, but not meant for the casual user. The metadata used by CAREO is the Canadian Core Learning Resource Metadata Application Profile (CanCore) of Instructional Management Systems (IMS) Learning Resource Metadata Information Model (CanCore 2003 and IMS 2001). Objects are archived as zip files containing the media, an HTML page referencing the media, and the metadata (CAREO, Undated, 9). While not set up to support an OAI repository, the metadata required is more extensive than simple Dublin Core, and provides sufficient information with which to create OAI records.

Because of the complexity of metadata entry, and the lack of open-source client software, this is probably not useful for our purposes.

Software Oriented Toward Total Content Management

All of the following software is commercial.

DIGITOOL, by Ex Libris
http://www.exlibris-usa.com/digitool.htm

The description is rather remarkable. It offers a range of options for element structuring and interlinking of objects and metadata, to support versioning, multiple formats, and hierarchical relationships. The web-based interface allows patrons to create personal collections or add to the library main collection; approval workflows are supported, and access can be regulated in accordance with user status and ID, course enrollment, or IP. Image thumbnails and full-text indexes for text-based documents are created

automatically upon upload; metrics are gathered automatically, and all assets are tracked. Searching includes metadata searches, full text, wildcard and Boolean search functions; results can be filtered, ranked and sorted, with automatic offers of related items. This tool supports OAI, Z39.50 and other data interchange protocols, and is designed to work seamlessly with ALEPH, our current OPAC system (Ex Libris 2004).
"Premier Partners" are listed as U of Maryland, Curtain U. and NYU; other implementations in the US include Brandeis, U of Iowa, Boston College, U of Notre Dame, Center for Jewish History (NYC) and Private Academic Library of Indiana (PALNI) (Ex Libris 2003).

An exploration of the implementation at John Curtain Prime Ministerial Library (the Electronic Research Archive, or ERA) yields some interesting information about how information is organized:

> Material is arranged in ERA in a hierarchical structure. By tracing the hierarchy, you can see where a record fits in and which creator is responsible for its creation.
>
> Creators can be individuals such as Frederick McLaughlin, private secretary to PM John Curtin, or agencies, such as Australia Post. A creator may have produced the records or simply been responsible for bringing them together as a collection.
>
> The hierarchical structure has either three or four levels, depending on the degree of detail needed to adequately describe the records. The levels are
>
> 1. Creator
> 2. Series
> 3. File
> 4. Item
>
> Example of a 3 level hierarchy for creator Frederick McLaughlin
> Creator - Frederick McLaughlin
> Series - Personal papers of F A McLaughlin
> File - Prime Minister´s visit to England via USA, Intinerary (sic) and engagements 1944
>
> Example of a 4 level hierarchy for creator Australia Post
> Creator - Australia Post
> Series - Philatelic items published by Australia Post
> File - First day covers commemorating the birth centenaries of Curtin & Chifley, 1985
> Item - FDC "Carried on Special Flight" with Curtin stamp, Creswick & Perth postmarks, 25 January 1985
>
> Each creator has one or more series, each series has one or more files, and if further subdivision is required, each file can have one or more items.
>
> (JCPML 2003, *ERA Help*)

A sample search turned up the following minimum metadata fields: system number, series ID, title, date, description, collection and creator; and, optionally, access (privileges) and biographical/historical note, extent, and available copies. The items themselves do not appear to be available online in this implementation.

In a private email, Leslie Myrick told me of her difficulties in making it work for a collection for an event two years ago; apparently the Digitool conception of METS

differed from theirs at the time (Myrick 10 March 2004). However, she later relayed a message from Jerry McDonough:

> Jerry says stay tuned on Digitool -- it's better than it was and they are actively trying to make it an IR. But it probably won't be ready for prime time til (sic) the end of the year in terms of structural metadata and exporting METS and SCORM.
>
> (Myrick 10 March 2004)

If Digitool will support both METS and SCORM, and tie into our Aleph database with MARC records, it could well be the best choice we could make for an institutional repository, to position us for future developments and interoperability, both with other libraries and with UT's Blackboard portal environment – as well as to integrate the holdings with our current OPAC system.

Hyperwave
http://www.hyperwave.com/e/

Hyperwave is a company offering a suite of content management tools. This is a broad, integrated enterprise solution intended to include e-learning applications and other documents, tying into LDAP and based on ORACLE; it offers search and retrieval, collaboration, user-defined portals and fine-grained access permissions for publishing to specific areas in the corporate knowledge base, with creators of documents using common Windows resources. This is easily the fanciest package out there that I have examined, in use by such large businesses as MetLife, BMW, Siemens, and the US Department of Defense (Kanda Software 2002).

However, from a cataloging standpoint, it has drawbacks. Metadata may be limited to author and creation date, with the optional ability to add keywords. (Kappe 2001: *Knowledge Management* and Kappe 2001: *Hyperwave*). Prices do not appear to be published online.

ContentDM
http://contentdm.com/products/overview.html

ContentDM is a commercial product, with pricing posted online (DiMeMa 2002: *Software Pricing*). Collections can be added to OCLC; it functions as OAI repository; and it is Z39.50 compatible through an open source product ZContent (DeMille 2003). Items can export in XML, SGML or tab-delimited text. Individual items are stored in web-accessible directories (there is no built-in security here).

Users can query multiple CONTENTdm servers (similar to ILS); access can be controlled by account, subnet, or IP. Uploads and metadata entry can be made from up to 50 remote locations using the "Acquisition Station" (DiMeMa 2002: *Overview*). Data can

be imported using tab-delimited text, which is useful for import of existing collection items and metadata from Microsoft Excel or Access.

The server runs on Linux, Solaris, NT 4.0, Windows 2000 or 2003, and the search client runs on Mac or PC with current browsers. Custom (in-house) applications and updates to this software may require licensing discussions (DiMeMa 2002: *Frequently Asked Questions*).

Currently the OAIster service provider unable to harvest over 1000 documents from these ContentDM OAI repositories without breaking them; but hopefully the new upgrade will repair this problem (Hagedorn 2004).

Conclusion

In choosing a software, or collection of software, with which to implement an institutional repository, we need to consider what we want it to do, what we will want to be able to do with its contents in the future, how interoperable we want it to be with our current library and UT portal systems, and last but not least, what our users want and need. The variety of choice is somewhat daunting, and the competing metadata standards and changing face of the digital library scene complicates the variety of futures that lay before us. The more metadata that can be gathered from ingested items, the more likely that we will be able to both preserve the objects appropriately, and migrate the associated metadata into forms which promote the sharing of digital objects. Balanced against this, we need to consider the limitations of what we can expect of the casual user, if creators are to enter their own metadata and upload their own objects. One software, Connexions, tries to bridge the gap by enforcing creation of objects in xml. Others, such as Digitool, seek ways to transform the metadata internally from one format to another to promote interoperability.

What we want from our repository will be metadata-driven. Many things are important here beyond user interface and ease of upload. Support for full-text and Boolean searching, creation of communities and sub-communities for organization and browsing, modularity and extensibility, ease of implementation and low cost are all important also. But beyond these possibly temporary and changeable aspects, is the content which we will be given to preserve. That content needs to be in a form, and with associated metadata, that will be amenable for both preservation and migration, into the foreseeable future.

References

Adobe Systems Incorporated. January 2004. *XMP Specification*. Available from www.adobe.com/products/xmp/pdfs/xmpspec.pdf; accessed 3 March 2004.

Advanced Learning Technology Resource Center. 2000. *SCORM Info and Tools* . Available from http://www.altrc.org/specification.asp; accessed 15 March 2004.

ALOHA: Advanced Learning Object Hub Application; Aloha User Guide Version 1.1b. Available from http://aloha.netera.ca/downloads/manual.pdf; accessed 26 March 2004.

ARNO – install. 2003. Available from http://arno.uvt.nl/~arno/arnodist/arno-1.0/INSTALL-1.0; accessed 23 March 2004.

Atkins, Anthony, Edward Fox, Robert France and Hussein Suleman. *ETD-MS: an Interoperability Metadata Standard for Electronic Theses and Dissertations – version 1.00, revision 2*. Available from http://www.ndltd.org/standards/metadata/current.html; accessed 1 April 2004.

Atlantis Institute of Science. 2002. *WebSubmit Help*. Available from http://cdsware.cern.ch/DEMO//submit/docs/english/; accessed 29 March 2004.

Australian Flexible Learning Framework. 2003. *VET Learning Object Repository*. Available from http://www.flexiblelearning.net.au/projects/resources/VLOR_green_paper.pdf; accessed 24 March 2004.

Bass, Michael, et al. 2002. *DSpace -- Functionality: Internal Reference Specification*. Available from http://libraries.mit.edu/dspace-mit/technology/functionality.pdf; accessed 30 March 2004.

Bass, Mick. 12 March 2004. *Questions Regarding Structure of DSpace.* Available from http://mailman.mit.edu/pipermail/dspace-general/2004-March/000132.html; accessed 14 March 2004.

Beaubien, Rick. 2003. *BrainDumpRickBeaubien.* Updated 21 August 2003. Raymond Yee's Wiki. Available from http://raymondyee.net/wiki/BrainDumpRickBeaubien; accessed 18 March 2004.

Beaubien, Rick. 2003. *RicksBrainDumpDistilled* Updated 25 August 2003. Raymond Yee's Wiki. Available from http://raymondyee.net/wiki/RicksBrainDumpDistilled; accessed 18 March 2004.

Bekaert, Jeroen, Lyudmila Balakireva, Patrick Hochstenbach, and Herbert Van de Sompel. 2004. Using MPEG-21 DIP and NISO OpenURL for the Dynamic Dissemination of Complex Digital Objects in the Los Alamos National Laboratory Digital Library. *D-Lib Magazine* 10:2 (February 2004). Available at http://www.dlib.org/dlib/february04/bekaert/02bekaert.html; accessed 30 March 30, 2004.

Bekaert, Jeroen, Patrick Hochstenbach, and Herbert Van de Sompel. 2003. Using MPEG-21 DIDL to Represent Complex Digital Objects in the Los Alamos National Laboratory Digital Library . *D-Lib Magazine* 9:11 (November 2003). Available from http://www.dlib.org/dlib/november03/bekaert/11bekaert.html; accessed 18 March 2004.

Bentum, Maarten van, Renze Brandsma, Thomas Place, and Hans Roes. 2001. Reclaiming academic output through university archive servers. Submitted to the *New Review of Information Networking* (August 2001). Available from http://drcwww.kub.nl/~roes/articles/arno_art.htm; accessed 23 March 2004.

Blackboard, Inc. 2002. *Blackboard Learning System™ and SCORM 1.2.* Available from http://www.blackboard.com/docs/datasheets/SupportforScorm.pdf; accessed 24 March 2004.

Blackboard, Inc. Undated. *Microsoft LRN Content*. Available from http://company.blackboard.com/docs/cp/learning_system/release6/instructor/_Microsoft_LRN_Content.htm; accessed 25 March 2004.

Blackboard, Inc. 2004. *Blackboard Developers: SDK Agreement.* Available from http://www.blackboard.com/dev/SDKAgreement.htm; accessed 25 March 2004.

Bormans, Jan and Keith Hill, eds. 2002. *MPEG-21 Overview v.5*. ISO/IEC JTC1/SC29/WG11/N5231 Shanghai (October 2002). Available from http://www.chiariglione.org/mpeg/standards/mpeg-21/mpeg-21.htm#_Toc23297968; accessed 30 March 2004.

Branschofsky, Margret, Rebecca Lubas, MacKenzie Smith, and Sarah Williams. 2003. *Evolving Metadata Needs for an Institutional Repository: MIT's DSpace*. Available from http://www.siderean.com/dc2003/703_Poster40-DC2003posterSubmitted.pdf; accessed 30 March 2004.

Burnett, Ian, et al. 2003. *MPEG-21: Goals and Achievements* . IEEE Computer Society: IEEE MultiMedia , October-December, 2003 (107-986X), p. 60-70. Available from: http://www.chiariglione.org/mpeg/events&tutorials/MMpaper.pdf; accessed 18 March 2004.

CanCore (Canadian Core). 2003. *Canadian Core Learning Resource Metadata Application Profile*. Available from http://phenom.educ.ualberta.ca/n/cancore/; accessed 29 March 2004.

CAREO (Campus Alberta Repository of Educational Objects). Undated. *CAREO Repository User Guide: Draft.* Available from http://careo.netera.ca/careo/userguide.pdf; accessed 29 March 2004.

Coalition for Networked Information. 2003. *Summary Report of the CNI Executive Roundtable on Institutional Repositories: December 8, 2003* . Washington, DC: Coalition for Networked Information (c2004). Available from http://www.cni.org/projects/execroundtable/fall2003summary.html; accessed 2 March 2004.

Consultative Committee for Space Data Systems. 2002. *Reference Model for an Open Archival Information System (OAIS)* . CCSDS 650.0-B-1 Blue Book (January 2002) ; ISO 14721:2002. Available from http://ssdoo.gsfc.nasa.gov/nost/isoas/wwwclassic/documents/pdf/CCSDS-650.0-B-1.pdf; accessed 3 March 2004

Day, Michael. 2003. *Preservation Metadata Initiatives: Practicality, Sustainability, and Interoperability* . Marburg, Germany: ERPANET Training Seminar, Metadata in Digital Preservation (3-5 September 2003). Available from http://www.erpanet.org/www/products/marburg/slides/erpaTraining-Marburg_Day.pdf; accessed 18 March 2004.

DC-Education Working Group Breakout Session: 15 October 2002, Florence Italy. 2002. University of Washington Information School . Available from http://www.ischool.washington.edu/sasutton/dcmi/DC-Ed-Breakout-Report.html; accessed 17 March 2004.

DeMille, Aaron and Kenning Arlitsch. 2003. *ZContent: An Open Source Z39.50 Connection to CONTENTdmTM Digital Collection Management Software*. Available from http://www.lib.utah.edu/digital/ZContent.html; accessed 29 March 2004.

Demonstration EPrints Archive. 2003. EPrints.org. Online software application. Available from http://demoprints.eprints.org/; accessed 2 March 2004.

Designing Courses: Learning Objects, SCOs, IMS Standards, XML, SGML, etc. Updated 21 February 2004. Available from http://www.ibritt.com/resources/dc_objects.htm; accessed 18 March 2004.

Digital Library and Archives. 2003. *etds@vt : About the Non-CGI scripts.* Available from http://scholar.lib.vt.edu/ETD-db/developer/non-cgi.html; accessed 24 March 2004.

Digital Library and Archives. 2003. *etds@vt : Welcome to the VT ETD-db.* Available from http://scholar.lib.vt.edu/ETD-db/; accessed 24 March 2004.

Digital Library Federation Steering Committee Meeting 28 October 2003: Summary. 2003. Digital Library Federation. Available from http://gateway.library.uiuc.edu/lon/SCMtgOct03.pdf; accessed 2 March 2004.

DiMeMa. 2002. *CONTENTdm Digital Media Management Software: CONTENTdm Software Pricing*. Available from http://contentdm.com/products/pricing.html; accessed 29 March 2004.

DiMeMa. 2002. *CONTENTdm Digital Media Management Software: Frequently Asked Questions*. Available from http://contentdm.com/products/faq.html; accessed 29 March 2004.

DiMeMa. 2002. *CONTENTdm Digital Media Management Software: CONTENTdm Overview*. Available from http://contentdm.com/products/overview.html; accessed 29 March 2004.

Dublin Core Metadata Initiative. 2003. *DCMI Education Working Group.* Updated 1 November 2003. Available from http://www.dublincore.org/groups/education/; accessed 16 March 2004.

Edinburgh University Library. Undated. *Theses Alive! At Edinburgh University Library.* Available from http://www.thesesalive.ac.uk/dsp_documentation.shtml; accessed 1 April 2004.

Eprints.org. 2004. Institutional Archives Registry. Available from http://archives.eprints.org/eprints.php; accessed 25 March 2004.

EVIA Digital Archive: Ethnomusicological Video for Instruction and Analysis. 2002. Indiana State University. Available from http://www.indiana.edu/~eviada/images/EVIADADevelopmentPhaseProposal-Web.pdf; accessed 18 March 2004.

Ex Libris, Ltd. 2003. *Digitool Overview*. Available from http://www.exlibris-usa.com/resources/digitool/DigiTool_overview_Dec2003.ppt; accessed 29 March 2004.

Ex Libris, Ltd. 2004. *Digitool Digital Asset Management: Overview.* Available from http://www.exlibris-usa.com/digitool.htm; accessed 29 March 2004.

Federated Digital Library: DODL: The Federated Digital Library; A strategic DLF initiative to develop a shared collection of academic information resources. 2003. Johns Hopkins University Scholarly Communications Group News & Events, 03-20-03. Available from http://openaccess.jhu.edu/news/details.cfm?news_id=25; accessed 2 March 2004.

Fedora™ Development Team. 2003. *Fedora™ Client Command-Line Utilities: Software Release 1.2* . The Rector and Visitors of the University of Virginia and Cornell University, c2003. Available from http://www.fedora.info/release/1.2/userdocs/client/cmd-line/index.html; accessed 4 February 2004.

Fleischhauer, Carl. 2003. *Report on METSmaker Software*. Washington, DC: Library of Congress, November 2003. Available from http://lcweb.loc.gov/rr/mopic/avprot/METSmaker2003_files/frame.htm; accessed 3 March 2004.

Gutteridge, Christopher. 2002. *GNU EPrints 2 Overview* . Greece: Proceedings of 11th Panhellenic Academic Libraries Conference. Available from http://eprints.ecs.soton.ac.uk/archive/00006840/02/eprintsoverview.pdf; accessed 2 March 2004.

Gutteridge, Christopher. 2003. *EPrints 2.3 Documentation: December 8, 2003* . Available from http://software.eprints.org/docs/eprints-docs.pdf; accessed 2 March 2004.

GNU EPrints Archive Software. Open Citation Project. Available from http://software.eprints.org/; accessed 2 March 2004.

Hagedorn, Kat. 15 March 2004. *Re: DSpace OAI compliance*. Personal email to Jody DeRidder, 15 March 2004. <khage@umich.edu>.

Hendricks, Brent and Adan Galvan. Undated. CNXML 0.5 Language Specification. Available from http://cnx.rice.edu/technology/cnxml/0.5/spec/; accessed 31 March 2004.

Hendricks, Brent, Adan Galvan, and Mark Husband. 2004. Connexions Tutorial and Reference. Available from http://cnx.rice.edu/content/col10151/latest/; accessed 31 March 2004.

IEEE Learning Technology Standards Committee. 2003. *WG12: Learning Object Metadata*. Available from http://ltsc.ieee.org/wg12/; accessed 18 March 2004)

IMS Global Learning Consortium. 2001. *IMS Learning Resource Meta-Data Information Model: Version 1.2.1 Final Specification*. Available from http://imsproject.org/metadata/imsmdv1p2p1/imsmd_infov1p2p1.html; accessed 29 March 2004.

IMS Global Learning Consortium. 2004. *IMS Learning Resource Meta-data Specification.* Updated 4 February 2004. Available from http://www.imsglobal.org/metadata/index.cfm; accessed 18 March 2004.

Internet Journals, Inc. 2003. *bepress Services*. c2003. Available from http://www.bepress.com/services.html; accessed 23 March 2004.

Internet Journals, Inc. 2003. *bepress Repository Technology*. c2003. Available from http://www.bepress.com/repositories.html; accessed 23 March 2004.

Intrallect. Undated. *intraLibrary: Digital repository and more…* Available from http://www.intrallect.com/products/intralibrary/features.htm; accessed 24 March 2004.

iPort. Undated. *iPort: Bestanden*. Available from http://dbiref.uvt.nl:4242/iPort?previous=2&sessionid=110800541220&request=databases ; accessed 29 March 2004.

iTor: Tools and technology for Open Repositories: Applications. 2003. Available from http://www.i-tor.org/en/system_info/about/applications_copy1/toon; accessed 31 March 2004.

iTor: Tools and technology for Open Repositories: Features. 2003. Available from http://www.i-tor.org/en/system_info/about/features_0_copy1/toon; accessed 31 March 2004.

iTor: Tools and technology for Open Repositories: More advantages. 2003. Available from http://www.i-tor.org/en/system_info/about/more_advantages/toon; accessed 31 March 2004.

ISO/IEC. 2001.*Technical Report: Information technology – Multimedia framework (MPEG-21)*. ISO/IEC TR 21000-1. Available from http://xml.coverpages.org/MPEG21-TR-21000-1.pdf; accessed 25 March 2004.

JCPML (John Curtain Prime Ministerial Library). 2003. *ERA Help: Understanding ERA*. Available from http://dtl.lis.curtin.edu.au/F/HI3GV5MDB9PSKD6BTDE1CQLGQPSALE26XBE6829DA3GBV9DFHH-01719?func=file&file_name=help-1; accessed 29 March 2004.

Jones, Richard. 2004. *DSpace vs. ETD-db: Choosing software to manage electronic theses and dissertations* . Ariadne 38 (Jan. 2004). Available from http://www.ariadne.ac.uk/issue38/jones; accessed 17 February 2004.

Kanda Software. 2002. *Kanda Integrates Star Tree With Hyperwave's Portal to Make Navigation Faster and More Intuitive for Army Users*. Available from http://www.kandasoft.com/press_Army.html; accessed 19 March 2004.

Kappe, Frank. 2001. *Knowledge Management with the Hyperwave eKnowledge Infrastructure*. Available from http://www.hyperwave.com/e/downloads/documents/whitepaper_km_eki.pdf; accessed 29 March 2004.

Kappe, Frank. 2001. *Hyperwave eKnowledge Infrastructure.* Available from http://www.hyperwave.com/e/downloads/documents/whitepaper_eki.pdf; accessed 29 March 2004.

Kenney, Anne. 2000. Summary of paper: Mainstreaming Digitization into the Mission of Cultural Repositories. CLIR (Council on Library and Information Resources) : *Collections, Content and the Web* (January 2000), published online as Reports pub88,

February 2000. Available from http://www.clir.org/pubs/reports/pub88/tech-response.html; with original article available from http://www.clir.org/pubs/reports/pub88/technology.html, both accessed 18 March 2004.

Lynch, Clifford. 2003. *Institutional Repositories: Essential Infrastructure for Scholarship in the Digital Age.* ARL Bimonthly Report 226 (February 2003). Available from http://www.arl.org/newsltr/226/ir.html; accessed 4 March 2004.

MacLaren, Iain. Undated. *Objects: What is a "learning object"?* Available from http://www.sesdl.scotcit.ac.uk:8082/celeste/obj3.htm; accessed 29 March 2004.

Magee, Michael. 2001. *CAREO Repository Background: Educational Object Repository Comparisons.* Available from http://www.kin.ucalgary.ca/cluster/metadata/Repository/repository_comparison.htm; accessed 29 March 2004.

Mattson, Mike and King Chung Huang. 2004. *Apollo*. Available from http://apollo.ucalgary.ca/project/events/NMCSpringConference04/presentation/; accessed 29 March 2004.

McNally, Karen, *Maricopa Learning eXchange (MLX).* The Labyrinth, (Spring 2000) Vol 8, Iss. 2. Available from http://www.mcli.dist.maricopa.edu/labyforum/Spr00/laby2.html. Accessed 11 March 2004.

Mellon Fedora Technical Specification: (December 2002) . 2002. University of Virginia Library; Cornell [University]. Available from http://www.fedora.info/documents/master-spec-12.20.02.pdf; accessed 20 June 2003.

MERLOT (Multimedia Educational Resource for Learning and Online Teaching). 2002. *MERLOT Peer Review*. Available from http://www.merlot.org/home/PeerReview.po; accessed 29 March 2004.

METS: An Overview & Tutorial . 2003. Metadata Encoding & Transmission Standard (METS) Official Web Site, Library of Congress. Updated 18 July 2003. Available from http://www.loc.gov/standards/mets/METSOverview.v2.html; accessed 9 October 2003.

METS Version 1.2 Documentation . 2003. Washington, DC: Library of Congress. Available from http://www.loc.gov/standards/mets/mets_xsdv12/mets.html; accessed 2 March 2004.

Microsoft TechNet. 2001. *Learning Resource iNterchange (LRN 3.0) Toolkit Solution Papers and Tutorial.* Available from http://www.microsoft.com/technet/itsolutions/education/toolkits/lrntoolkit/toolkitl.mspx; accessed 25 March 2004.

Millar, Glenn. 2003. Inside Out: Learning Objects 101: A Primer for Neophytes. *Sidebars* (November 2002), British Columbia Institute of Technology. Available from http://online.bcit.ca/sidebars/02november/inside-out-1.htm; accessed 29 March 2004.

MIT & Hewlett Packard Co. 2003. *DSpace Federation: Metadata*. Available from http://www.dspace.org/technology/metadata.html; accessed 1 April 2004.

MIT & Hewlett Packard Co. 2003. *DSpace Federation: The DSpace Federation Project*. Available from http://www.dspace.org/federation/project.html; accessed 30 March 2004.

MIX: NISO Metadata for Images in XML Schema. Updated 26 June 2003. Library of Congress. Available from http://www.loc.gov/standards/mix/; accessed 3 March 2004.

MODS: Metadata Object Description Schema. Updated 8 March 2004. Library of Congress. Available from http://www.loc.gov/standards/mods/; accessed 3 March 2004.

MyCoRe *Project Description.* Undated. Available from http://www.mycore.de/engl/index.html; accessed 23 March 2004.

Myrick, Leslie. 10 March 2004. *Re: Digitool*. Personal email to Jody DeRidder, 10 March 2004. leslie.myrick@nyu.edu.

Myrick, Leslie. 15 March 2004. *Re: DSpace METS*. Personal email to Jody DeRidder, 15 March 2004. leslie.myrick@nyu.edu.

NASA. 2004. *ISO Archiving Standards – Reference Model Papers*. Last revised 1 March 2004. Available from http://ssdoo.gsfc.nasa.gov/nost/isoas/ref_model.html; accessed 3 March 2004.

National Library of Canada. 2001. *Projects and Programs - Projects, Programs and Partners - Canadian Learning Object Repositories.* Available from http://www.nlc-bnc.ca/9/1/p1-264.1-e.html; accessed 29 March 2004.

Nixon, William. 2003. DAEDALUS: Initial experiences with EPrints and DSpace at the University of Glasgow . *Adriane* 37 (October 2003). Available from http://www.ariadne.ac.uk/issue37/nixon/; accessed 27 February 2004.

Norman, D'Arcy. 27 March 2003. *D'Arcy Norman's Learning Commons Weblog: Learning Object Repository Interest Level Rising*. Available from http://commons.ucalgary.ca/weblogs/dnorman/000063.html; accessed 12 March 2004.

OAI Harvester Plug-In for Dspace (and software). Undated. Old Dominion University. Available from http://kepler.cs.odu.edu:8080/kepler/dspace/downloads.html; accessed 4 March 2004.

OCLC/RLG Working Group on Preservation Metadata. 2002. *Preservation Metadata and the OAIS Information Model: A Metadata Framework to Support the Preservation of Digital Objects* (June 2002). Available from http://www.oclc.org/research/projects/pmwg/pm_framework.pdf; accessed 17 March 2004.

Open Society Institute. 2004. A *Guide to Institutional Repository Software,* 2nd ed. (January 2004). Available from http://www.soros.org/openaccess//software/; accessed 25 March 2004.

Payette, Sandra and Lagoze, Carl. 1998. *Flexible and Extensible Digital Object and Repository Architecture (FEDORA): ECDL98 Paper* . NY: Cornell University. Available from http://www.cs.cornell.edu/payette/papers/ECDL98/FEDORA.html; accessed 3 February 2004.

Rice University. *Connexions – Sharing Knowledge and Building Communities*. Available from http://cnx.rice.edu/; accessed 29 March 2004.

Seaman, David. 2003. *E-Content: Deep Sharing: A Case for the Federated Digital Library* . Educause Review (July/August 2003). Available from http://www.educause.edu/ir/library/pdf/erm0348.pdf; accessed 2 March 2004.

Simms, Jason. 2004. *Re: downside of DSpace?* Personal email to Jody DeRidder, 25 March 2004. jsimms@utk.edu.

Simpson, Mike. 2004. *Mike Simpson Notes During Conference*. Available from http://wiki.dspace.org/index.php/MikeSimpsonNotesDuringConference; accessed 16 March 2004.

Simpson, Mike. 2004. *Mike Simpson Thoughts During Conference*. Available from http://wiki.dspace.org/index.php/MikeSimpsonThoughtsDuringConference; accessed 30 March 2004.

Smith, MacKenzie. 5 March 2004. *DSpace Customization*. Available from http://mailman.mit.edu/pipermail/dspace-general/2003-September/000020.html; accessed 30 March 2004.

Smith, MacKenzie. 14 March 2004. *Re: METS application in DSpace?* Available from http://sourceforge.net/mailarchive/forum.php?thread_id=4051552&forum_id=13580; accessed 14 March 2004.

Smith, MacKenzie. 14 March 2004. *Re: Subversive use of dc.creator…* Posted to DSpace technical listserv; possibly available from http://sourceforge.net/mailarchive/message.php?msg_id=7514926; accessed 30 March 2004.

Smith, MacKenzie. 19 March 2004. *Summary of March 2004 DSpace user group outcomes*. Available from http://mailman.mit.edu/pipermail/dspace-general/2004-March/000140.html; accessed 23 March 2004.

SURF. 2003. *Appendix C: Comparitive review of DSpace, ARNO and NCP* Available from http://www.surf.nl/infodare/AppendixCv3_ComparativeReview.pdf; accessed 23 March 2004.

The le@rning Federation. 2004. *BELTS Project Overview*. Available from http://belts.sourceforge.net/; accessed 24 March 2004.

Universität Stuttgart. 2003. *About OPUS*. Available from http://elib.uni-stuttgart.de/opus/doku/english/about_english.php; accessed 24 March 2004.

Universität Stuttgart. 2003. Electronic *Dissertations / Habilitations.* Available from http://elib.uni-stuttgart.de/opus/doku/english/e-diss_english.php; accessed 24 March 2004.

University of Colorado. 2003. *Metis Workflow System*. Available from): http://www-serl.cs.colorado.edu/metis/index.html; accessed 29 March 2004.

Vaidya, Prajakta P. 2004. *RE:* . Personal email to Jody DeRidder, 5 March 2004. pvaidya@indiana.edu.

Van de Sompel, Herbert. 2003. *Digital Library Architecture: Based on MPEG-21 DIDL, the OAI-PMH and the OpenURL*, *Project Briefing: Fall 2003* . Coalition for Networked Information Task Force Meeting. Available from http://www.cni.org/tfms/2003b.fall/abstracts/PB-digital-sompel.html; accessed 3 March 2004.

Witten, Ian H. and Stefan Boddie. 2003. *Greenstone Digital Library User's guide (March 2003).* Available from http://aleron.dl.sourceforge.net/sourceforge/greenstone/User-2.39-en.pdf; accessed 10 March 2004.

Yee, Raymond. 2003. *MetsImsInteropPaper/MetsImsComparison* Updated 20 August 2003. University of California Berkeley. Available from http://raymondyee.net/wiki/MetsImsInteropPaper_2fMetsImsComparison; accessed 18 March 2004.