Appendix A

Laying the Groundwork: Foundations for Sabbatical Leave

Workshops attended

- American Library Association, Library Information Technology Association (LITA) pre-conference entitled "Practical Digital Libraries Overview" given by Dr. Edward A. Fox, Head of the Digital Library Research Laboratory at Virginia Technical Institute on October 12, 2001. The focus of this session was on the creation of a digital repository for electronic theses and dissertations.
Practical Digital Libraries Overview

- Coalition for Networked Information (CNI) project briefing entitled "The Fedora Project: an Open Source Repository for the Management of Content and Services" given by Thornton Staples, Director of Research and Development at the University of Virginia Library and by Sandy Payette, Researcher at Cornell University, April 28, 2003. The focus of this briefing was to describe the institutional repository model known as Fedora, its architecture and its functionality using open source software.
The Fedora Way: A Programmer's View

- Northeast Regional Computing (NERCOMP) Library/IT Special Interest Group (SIG) program on the Massachusetts Institute of Technology institutional repository, DSpace, April 30, 2003. The workshop was entitled "DSpace@MIT: an Open Source Institutional Digital Repository"
NERCOMP Workshop: DSpace@MIT: An Open Source Institutional Digital ...

- Northeast Regional Computing (NERCOMP) SIG workshop entitled "New England e-Portfolio Summit", Simmons College, Boston, MA, June 2, 2004. A NERCOMP workshop focused on new developments in the area of electronic portfolios, especially developing a New England pilot project using
The Open Source Portfolio Initiative (OSPI) software.
NERCOMP workshop: New England ePortfolio Summit

- American Library Association, Library Information Technology Association (LITA) conference "Putting Technology into Practice, Oct 7-10, 2004. I attended sessions relevant for institutional repository work, including sessions on XML in digital libraries, major trends in digital libraries, optimizing the value of a digital library, development of database-driven websites, and OAIster which was a one year Mellon funded project to use Open Archives Initiative (OAI) protocol to harvest data from multiple academic digital repositories and develop an end-user search service. In addition I found interesting sessions on Shibboleth, an open-source authentication project, and CUGIR, an open archives geospatial information repository at Cornell University.
ALA | LITA National Forum 2003, Norfolk, VA

- Northeast Regional Computing (NERCOMP) SIG workshop entitled "Learning Objects", Smith College, Northampton, MA, Oct 14, 2004. This session focused on the development of learning objects and open resources websites such as MERLOT.
NERCOMP workshop: Learning Objects

- SPARC & SPARC Europe meeting "Institutional Repositories: The Next Stage" Nov 18-19, 2004. This workshop was devoted entirely to IR: populating and marketing strategies; legal issues such as authors' agreements, copyright, and creative commons; IR policies; IR business models; how to choose an IR platform; and open access legislation update. A truly critical resource.
| SPARC | 2004 IR Workshop |

- American Library Association, MidWinter meeting, Jan 14 – 18, 2005. I attended several sessions on repositories, including the LITA Emerging Technologies Interest Group meeting, Web Of Science, Top Technology Trends, as well as coordinating a meeting between the NECOP Universities and ProQuest Digital Commons to explore that solution as a potential collaborative IR solution
ALA | home

Conference and event planning:

- UMass Amherst Libraries' sponsored Scholarly Communication Colloquium series:
  Freedom of Speech & Speech for Free with keynote speaker Peter Givler, Oct.15, 2004
  In the Public Interest: Expanding Access to Scholarly Research
  with keynote speakers John Willinsky and Jane Griffith, Mar. 03, 2005

- ProQuest's Digital Commons demonstration at UMass Amherst Libraries, Oct. 28, 2004

- Northeast Regional Computing (NERCOMP) Library/IT Special Interest Group (SIG) program on the Massachusetts Institute of Technology institutional repository, DSpace. The workshop was entitled "DSpace: An Open Source Institutional Digital Repository", June 09, 2004.

- Association of Research Libraries. Office of Leadership and Management Services webcast entitled "Institutional Repositories: Revealing our Strengths; cosponsored by SPARC and CARL, June 10, 2004
Webcast: Institutional Repositories

- American Libraries Association. ProQuest's Digital Commons meeting for the NECOP Electronic Repository Task Force and Directors of Libraries, Jan. 15, 2005.

Committee work:

- New England Council of Presidents (NECOP) Electronic Repository Task Force, February – December, 2004. I was honored to be asked to represent the University of Massachusetts on this Task Force and was elected chair of the group. The charge and complete report are in Appendix A.

Presentations:

- NECOP Electronic Repository Task Force presentation to NECOP Directors of Libraries, presenting the findings of our research, Dec 17, 2004

- NELINET seminar: Institutional Repositories: Models and Approaches. I was asked to give a presentation defining an institutional repository, its core features and functionality, key issues and staffing concerns. The presentation was entitled: "Repositories: New Beginnings for Libraries", January 27, 2005.

THE UMASS AMHERST LIBRARIES AND UNIVERSITY OF MASSACHUSETTS PRESS

are pleased to present

A COLLOQUIUM

for Faculty, Librarians, and Graduate Students

3:00 p.m. to 4:30 p.m., October 15, 2004
Cape Cod Lounge, Student Union Building

FREEDOM OF SPEECH & SPEECH FOR FREE:

University Libraries, University Presses, and the Law

What are the hot-button legal issues confronting academic publishers, librarians, and scholars in the digital age? How are questions of intellectual property rights and freedom of speech playing out on campuses and in the courts—from electronic reserves and "open access" to the Patriot Act and the Office of Foreign Asset Control's Information Regulations? How has the environment of scholarly publishing changed in recent years, and what are the prospects for the years ahead?

Addressing these questions will be

PETER J. GIVLER,

Executive Director of the Association of American University Presses, a national association that includes 125 scholarly presses. His talk will be followed by responses from the floor and a period of questions and answers.

After the colloquium, the University of Massachusetts Press will hold an Open House at its new offices in the recently renovated East Experiment Station. All are welcome.

This event is cosponsored by the Faculty Senate's Research Library Council, Research Council, and University Press Committee and by the Friends of the Library.

AAUP & ARL

The Association of American University Presses (AAUP) and the Association of Research Libraries (ARL) have designated 2004 as the "Year of the University Press." This year-long focus on university presses is intended to celebrate the important role presses play in the scholarly communications process. The campaign is an outgrowth of a recent formal recognition by the two organizations of their complementary roles in the scholarly communications system and the need to work together in this time of economic and technological turbulence to ensure a strong system for the future.

For further information, please contact Emily Silverman (UMass Amherst Libraries, 413-545-0995) or Alice Maldonado (UMass Press, 413-545-2217).

IN THE PUBLIC INTEREST:

Expanding Access to Scholarly Research

A Colloquium for
Faculty, Graduate Students and Librarians

Thursday March 3, 2005
3:00-4:30 p.m.

Cape Cod Lounge, Student Union Building

Wine and cheese reception following

What are the implications of "Open Access" to scholars, publishers and libraries? As interest in Open Access grows, its impact on scholarship is being investigated. Two experts will discuss the issues and implications.

John Willinsky
Professor, Department of Language and Literacy Education,
Faculty of Education, University of British Columbia
Principal Investigator of the Public Knowledge Project

Jane Griffith
NLM Assistant Director for Policy and Legislative Development
The National Library of Medicine
National Institutes of Health

Sponsored by:

Vice Provost for Research **Dean of the Graduate School**

Research Library Council **UMass Amherst Libraries**

For further information, please contact:
Leslie Schaler, UMass Amherst Libraries (545-0162) or lschaler@library.umass.edu

New England Land-Grant Universities
Institutional Repository Task Force
Charge to the Task Force

The deans and directors of libraries of the New England Land-Grant Universities have established a Task Force on Electronic Repositories to study the purpose and potential of this approach to preserving and providing information. Claudia Morner, University Librarian, University of New Hampshire, will serve as liaison between the Task Force and the Deans and Directors group.

All six New England Land-Grant Universities are in early stages of thinking about electronic repositories. The librarians feel that this is an excellent time to share the information gathered by our various experts and to work together in analyzing options. Specifically, they ask the Task Force to undertake the following:

1. Purpose
Develop a short statement explaining the purpose(s) and potential of an electronic repository. Is the repository primarily a storage solution? Is it a productivity tool? What long-range benefits to faculty and the institution can be anticipated?

2. Software
What are the major software systems being developed? What research universities are leading/joining the major systems? What are the key operational differences in how the software is used? Are there significant differences in purchase price or support cost?

3. Hardware and storag
What are the requirements? Could savings be achieved by sharing hardware costs?

4. Content issues
What types of content are our universities considering as they envision an institutional repository? This question relates to purpose, but also to the procedural questions of how information will be approved and submitted. What is the status of material in the repository? Is it all "open source"?

5. Access
What restrictions on access or use of materials?

6.Operations
What would be the process of entering information into the repository. What are the staffing needs for cataloging and entering information submitted by approved parties?

7. Models
Having discussed the above issues, can the task force develop a few alternative models of how an institutional repository might be structured and staffed? What would be the start-up and ongoing costs?

8. Cooperation in implementation
What cost savings, if any, could be realized through cooperation among some or all of the universities in establishing a repository? Is there a good basis of agreement about the technical and procedural issues that would be required in a collaborative effort?

February 20, 2004

Electronic Repositories

Report presented to the NECOP deans and directors of libraries
December 17, 2004

In February 2004, the deans and directors of libraries of the New England Land-Grant Universities established a Task Force on Electronic Repositories to study the purpose and potential of this approach to preserving and providing information and to work together in analyzing options. Specifically, they ask the Task Force to undertake the following eight areas in the scope of their study: the purpose and potential of a repository, major software solutions, hardware and storage requirements, content issues, access restrictions, operational processes, potential models, and cooperation incentives. These areas will be addressed in this white paper. Supplemental material, along with the current status of concept development and implementation of repositories at each institution, is included in the appendices.

Purpose and potential of a repository
Develop a short statement explaining the purpose(s) and potential of an electronic repository. Is the repository primarily a storage solution? Is it a productivity tool? What long-range benefits to faculty and the institution can be anticipated?

The nature of scholarly communication has been rapidly changing in the networked information revolution of the past few years. Institutional repositories have emerged as a new strategy to leverage and support these new areas of scholarship and scholarly communication. Clifford Lynch, executive director of the Coalition for Networked Information, stated in the ARL Bimonthly Report 226 that a 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." For the purposes of this report, the word institution can be replaced with consortium.

A repository is one way to showcase the entire range of research conducted at the university, providing a highly visible platform for personal scholarship and a centralized location for department, institute and research center materials that would otherwise be scattered throughout the institution. It provides a way for faculty to leave their legacy and be recognized for leadership in their chosen field. Using a repository for depositing research findings addresses the issue brought up by recent NIH documents for publicly funded research to be made publicly available. In addition, a repository can contain the ongoing teaching and learning materials that faculty and students create as well as provide a location to document the ongoing activities of the institution in the form of archives, events, performances, and other items of record.

Common features of a repository include the following elements
- content in digital form in a wide variety of types (text, audio, video, images, data sets)
- community focus, which determines what is included in the repository. The community members are the authors and copyright owners of the deposited content.
- institutional support, requiring collaboration across an organization. The repository requires long-term financial support to ensure that the content is preserved and maintained.
- durable, permanent content. Digital preservation techniques still need to be enhanced to ensure lasting retrieval.
- access to cotent by a broad audience, a community-shared alternative to local storage of content, fosters serendipitous discovery across disciplines.

Core functions that are essential for a repository

- material submission, some way for the author to deposit material, provide for editing to assure quality of content, conversion to archival format such as PDF
- metadata application, such as author, title and descriptive information and administrative data such as date and time of submission
- access control, or digital rights management, to provide for controlled access to the repository content. Even if the entire community has access to the content, there needs to be a way to restrict the ability to add, delete, edit, and approve content.
- discovery support, usually a search engine that supports browsing and full-text searching of the content
- distribution and dissemination of content to enable display and download capabilities
- preservation, some mechanism(s) for the content to be preserved and retrievable over time, including a persistent documentation identification system.

Software:
What are the major software systems being developed? What research universities are leading/joining the major systems? What are the key operation ences in how the software is used? Are there significant differences in purchase price or support cost?

The major repository software systems marketplace being developed fall into two main areas: open source and commercially available solutions.

Open Source software solutions: By definition, the open source software movement is an effort to "rationalize the software business" within higher education and to enable higher education to unbundle software tools, not getting locked into one vendor. Open source software is freely available and OAI compliant. Further information comparing a variety of open source institutional repository software, not including Fedora, can be found in the report written in Oct 2003 by the Open Society Institute entitled "A Guide to Institutional Repository Software"

DSpace

Developed jointly by MIT Libraries and Hewlett-Packard (HP), DSpace is freely available to research institutions worldwide as an open source system that can be customized and extended. DSpace is a groundbreaking digital institutional repository that captures, stores, indexes and redistributes the intellectual output of a university's research faculty in digital formats with an emphasis on preservation of digital materials.

DSpace accepts all manner of digital formats. Some examples of items that DSpace can accommodate are: documents (e.g. articles, preprints, working papers, technical reports, conference papers); books; theses and dissertations; data sets; computer programs; visualizations, simulations, and other models; multimedia publications; and learning objects.

Sites using DSpace (DSpace Federation website) include: Cambridge University; Columbia University; Cornell University; Massachusetts Institute of Technology; Ohio State University; University of Oregon; University of Rochester; University of Toronto; and University of Washington.

EPrints

EPrints software has been created so that institutions can create OAI-compliant Archives quickly, easily and for free. EPrints servers are designed to help dissemination of research publications by sharing information (metadata) using the latest (OAI) standards.

Some examples of sites using EPrints software:

- arXiv, a large Open Archive in physics, mathematics and some related disciplines
- Behavioural and Brain Sciences, an archive of preprints of target articles from the peer-reviewed journal of the same name.
- California Digital Library (3 Archives using eprints.org)
- CogPrints, a multidisciplinary Open Archive in the Cognitive Sciences (psychology, neuroscience, computer science, biology, linguistics, philosophy)
- PubMed Central, a Life Sciences Research Reports site (not yet OAI-compliant)
- Psycoloquy, an archive of refereed reprints, commentaries and responses from the online journal of the same name (currently not OAI compliant but soon to be upgraded to EPrints 2)

Fedora

Fedora was jointly developed by the University of Virginia and Cornell University. The project was funded by the Andrew W. Mellon Foundation to build an open-source digital object repository management system based on the Flexible Extensible Digital Object and Repository Architecture (Fedora). The new system demonstrates how distributed digital library architecture can be deployed using web-based technologies, including XML and Web services. Fedora is a general-purpose digital object repository system that can be used in whole or part to support a variety of use cases including: institutional repositories, digital libraries, content management, digital asset management, scholarly publishing, and digital preservation. In the newest version, the Fedora Access interfaces now support date-time stamped requests, so that a client can "go back in time" and see a digital object as it looked in the past. Additionally, this release provides a migration utility for mass export and mass ingest of objects from either directories or other repositories.

Major features of Fedora software include: open source; web services; flexible digital objects with disseminators that associate a set of behaviors with a digital object; content versioning; XML ingest and export; digital object storage; access control and authentication; OAI metadata harvesting; migration and batch utilities. Upcoming versions of the software will add important functionality, such as Shibboleth-based authentication, fine-grained policy enforcement, workflow support, enhanced preservation features, and performance enhancements to support extremely large repositories.

Sites using Fedora include: Indiana University; New York University; Northwestern; Rutgers University; Tufts University; and University of Virginia.

Commercially available solutions: Many commercial vendors are now offering repository solutions. When evaluating these products, it is critical to ensure that the vendor's product adheres to industry standards and is easily portable to another platform. One area of the industry where pilot projects are taking place is the Integrated Library System vendor marketplace. One such example is Northeastern University who is working with Innovative Interfaces and DSpace to create an integrated product.

BioMed Central's Open Repository

Open Repository is a service from BioMed Central which builds and maintains repositories on behalf of institutions and organizations. This software allows organizations to quickly set up and run an industry-standard DSpace repository with minimal effort and investment in manpower and cost. At all times the customer retains ownership of the repository. Open Repository is built using Open Source technology platform: DSpace and allows institutions to migrate to their own in-house DSpace repository to take over its operation and maintenance.

Major features of Open Repository software include set up, customization, service maintenance and ongoing support, document upload and formatting, populating the repository through data feeds such as BioMed Central and PLoS, search functionality, customer service and support, and the ability to showcase and exposure of the organization's intellectual output.

There were no sites using OpenRepository listed in their documentation.

CONTENTdm

CONTENTdm provides tools to organize, manage, publish and search digital collections on the Web. This flexible, multifunction software package handles multiple format: documents, PDFs, images, video and audio files. CONTENTdm offers a scalable solution to grow with a repository's needs. This software is intended to provide a solution for digital library collections, not to provide a complete institutional repository solution.

Sites that use CONTENTdm include: Colorado State University Libraries; University of Washington Libraries; and Washington State University.

UMI/ProQuest's Digital Commons@

Digital Commons is a new Digital Institutional Repository service provided by UMI ProQuest. It is OAI-compliant and provided as a partnership with the Berkeley Electronic Press (Bepress). Digital Commons is designed so that UMI manages the repository software and hardware and the cost model is that institutions pay at different rates for different levels of access.

Features of this product include: immediate population of content – theses and dissertations uploaded by ProQuest; full text searching and new features such as access paths to this work via Google's scholar search capability; integrated management tool; remotely hosted service; incorporates a wide variety of publication types, formats, and dynamically generated resources (articles, preprints, reports, sound, video, data sets, executables); and provides usage statistics.

Institutions using Digital Commons include: Boston College; California Digital Library; Florida State; New England Law Library Repository (University of Connecticut, Cornell and Yale); University of New Brunswick; and University of Pennsylvania.

Hardware and storage
What are the requirements? Could savings be achieved by sharing hardware costs?

Hardware and storage require
savings could be achieved by bulk purchase of
commercially available solutions or, if going with open source, by locating hardware at one of the sites.

DSpace hardware specifications from the DSpace website:

There are no specific server requirements for DSpace except UNIX (and si
ational database system. Most
of the necessary drivers and libraries are packaged together with the DSpace system. This set of tools should run on any UNIX-type OS, such as Linux, HP/UX or Solaris, and other libraries can be substituted if run on another platform. General hardware architecture (server, memory, and disk storage) recommendations (not necessarily endorsements) are listed below:

- HP Server rx2600, powered by dual 64-bit Intel Itanium 2 processors (900MHz), 2GB RAM, 26 GB internal disk storage. HP StorageWorks Modular SAN Array 1000 (msa1000) with a single high-performance controller. Options include a second controller and, with the addition of two more drive enclosures, controls up to 42 Ultra2, Ultra 3, or Ultra320 SCSI drives. Total capacity can be six terabytes. **Cost starts around $40K and up to around $1.8M**

- SunFire 280R Server, two 900MHz UltraSPARC-III Cu processors, 8MB E-cache, 2GB memory, two 36GB 10,000rpm HH internal FCAL disk drives, DVD, 436-GB, or 12 x 26.4 Gbyte 10K RPM disks, Sun StorEdge A1000 rackmountable w/ 1 HW RAID controller, 24MB std cache. **Around $30K**.

- Dell PowerEdge 2650 with dual Xeon processors (2.4GHz), 2GB RAM, 2x73GB scsi disks. One 2.5TB Apple XServe. A DLT tape library to back up the DB/jsps etc. Around $10K. – not recommended for larger institutions

Fedora hardware requirements are similar to DSpace

BioMed Central's Open Repository and UMI/Proquest Digital Commons:

If the NECOP institutions decide to move forward with one of these solutions, there are no hardware or storage requirements since it is a remotely hosted option. There are a variety of licensing options depending on the amount of postings or uploads, quantity of storage and other factors. For example, Digital Commons costs include an initial pilot project startup fee of $12,500. An annual subscription license including unlimited postings/uploads, theses / dissertations = $34,500 depending on size of institution. ProQuest would grant the NECOP institutions a 20% discount if purchased as a consortium instead of individual purchases by each institution. Open Repository has three different pricing structures ranging in price from $9,050 one time setup fee plus $4525/yr maintenance fee to $18,100 one time setup fee plus $45250/yr maintenance fee. The functionality and customizability varies accordingly. Discounts for large quantities of storage are available.

Content and access issues

What types of content are our universities considering as they envision an institutional repository? This question relates to purpose, but also to the procedural questions of how information will be approved and submitted. What is the status of material in the repository? Is it all "open source"? What restrictions on access or use of materials?

The key definition for "institutional repository" is that it documents the digital intellectual products of an institution. At the heart of setting up and maintaining an institutional repository lies the content that is available through the repository.

Content in its purest form refers to the intellectual information stored on the site. For a collaborative project to be successful, content needs to be meaningful for users, easily accessible to and controlled by contributors and be governed by policies that address the mutual goals of participating institutions and contributors. Some examples of content have been listed above in the various software solutions.

Before undertaking a collaborative effort, the land-grant institutions need to determine the goals for a joint institutional repository. This includes how likely users will be to access the repository based on the content and how useful or attractive that content will be. The institutions will also need to set policies to govern what type of content would be submitted to the site, what areas to focus on, and what file formats to support. There may need to be some limitations for accessing content because of regulations or copyright issues for submitted research. Institutions would need to balance the need for access versus the need for preservation of content in regards to allowing restricted content on the site. A policy regarding submitted content would also address what types of content would be included. This policy would narrow the scope of the content by limiting what could be posted but yet ensure high quality content by allowing certain items to be posted. Another consideration would be what type of content should be included, such as geo-spatial datasets, databases of research data, conference proceedings, lecture or lab notes, and other research related materials. Geo-spatial datasets could take up considerably more storage space than lab notes; a policy decision that allows for this type of content would necessitate more digital storage.

Electronic portfolios

Electronic portfolios, although not specifically part of the Task Force's charge, have a complementary role to the repository content as described here and in more depth at http://www.libraryjournal.com/article/CA456768 An electronic portfolio, or e-portfolio, system is a web site through which students maintain developmental information about their university careers. The portfolio owner creates permission sets that define the portion of the portfolio visible to other students, faculty, staff, and guests. The portfolio serves as a digital representation of the student's development record in academic and co-curricular activities from first-year orientation to graduation, and may be maintained throughout the alumna's lifetime.

There are several emerging platforms that provide e-portfolio functionality in different ways, but key conceptual components include:

- website to deposit and display a student's work
- mechanism for faculty, staff, and peers to access, exchange, annotate and refine the student's work
- customizable interface that allows students to arrange and display their works in creative ways
- ability for portfolio owners to define the information others can access within the portfolio
- easy to use interface and a flexible system that allows for the complex model of the university and departmental development metrics

An effective electronic portfolio system will need to be tightly integrated with other systems within the campus infrastructure (portals, student information systems, etc.), and may, in fact, be a transit point between student works created in an instructor-centric Course Management System and a research community-centric Institutional Repository. For instance, the culmination of a student's undergraduate career may result in a body of work created in junior and senior level courses that is bundled into a senior thesis. The component parts of that senior thesis can be stored in the student's e-portfolio and visible to the world as a view set up by the student of the specific works in the portfolio. In turn, a research community may recognize individual works or the entire thesis as a contribution to the profession and want to include it in the community's institutional repository area. The act of students submitting works from their e-portfolios to the institutional repository becomes the key link between the systems.

Operations:
What would be the process of entering information into the repository? What are the staffing needs for cataloging and entering information submitted by approved parties?

I. Tools: Metadata

"If digital information objects that are currently being created are to have a chance of surviving migrations through successive generations of computer hardware and software, or removal to entirely new delivery systems, they will need to have metadata that enables them to exist independently of the system that is currently being used to store and retrive them." (from *Introduction to Metadata: Pathways to Digital Information* / A.J. Gilliland Swetland).

Though a relatively newly coined term, descriptive metadata was the backbone of library catalog records long before the advent of the online catalog. MARC (Machine Readable Cataloging) is a descriptive metadata schema that facilitated the relatively quick rise of the Online Public Access Cataloging. It also served as a springboard for creative implementations such as the Library of Congress Community Information Format. Library staff at the University of Maine, working with Innovative Interfaces partnered with LC and the public library community to stretch the capabilities of MARC tagging.

The archive community, moving toward an online environment that favored access while preserving original content found the MARC schema to be lacking in certain key elements crucial to effective finding aids for large bodies of material. Other forms of metadata have emerged in light of online archivist efforts and the current rise in electronic publishing. Distinct metadata categories can be delineated: administrative (rights management), descriptive (cataloging), preservation, use (tracking) and technical (authentication).

Consistent, standardized metadata is critical to institutional repositories achieving one of their primary goals of long-term preservation. The Dublin Core metadata schema has gained a great deal of momentum for building repositories as it readily maps to MARC. At 15 elements it is simple enough to accommodate a broad range of collection types. Qualifiers, often referred to as attributes, further refine the markup to meet different user community goals. DSpace, the MIT repository model, is currently deploying a qualified version of the Dublin Core schema. (See http://dspace.org/technology/metadata.html)

II. Tools: The Open Archive Initiative (OAI)

The Open Archive Initiative is another critical development in support of institutional repositories, providing a protocol that facilitates the interoperability of systems and services. The OAI Protocol for Metadata Harvesting (OAI-PMH) allows metadata, rather than the content to be moved to an aggregated source to facilitate broadcast (federated) searching. Thus diverse collections can become available to a single search tool. Other mechanisms must be deployed to achieve the latter. The Open Archives Forum provides a very readable overview and tutorial for the Initiative. For a more technical treatment, see the Open Archives Initiative site. See Appendix A for an example of a record from the Maine Music Box collection that has been marked up for inclusion in the UIUC Digital Gateway to Cultural Heritage Materials (http://oai.grainger.uiuc.edu/AboutCollections.htm).

III.Staffing

An institutional repository project clearly provides opportunities for cross-departmental and cross-institutional participation. Many libraries already have significant digital projects originating within Special Collections, Archives and Systems departments. Staff within these units will have experience with various encoding standards and metadata schemas (eg. XML and Dublin Core). Public service staff will play a key role in marketing and encouraging participation from various departments through liaisons. Pooling this wealth of complementary experience to form an advisory committee is essential. Partnering opportunities with faculty willing to serve as early adopters is another critical piece. This can be an offshoot of a standing University Library Committee.

Many institutions have elected to proceed cautiously with a repository pilot where select materials have been submitted and processed. A budget for finite staffing hires is then plausible through grants. The problems with this approach are visibility and buy in for the broader university community. Cultural differences between disciplines and their comfort level with information technologies tend to result in an unbalanced representation of materials in the repository. As echoed by the early returns of the faculty survey at the University of Maine, faculty in general with prior experience participating in subject based repositories are primarily in the scientific disciplines.

Long term, the most practical option for an open repository is integration into the existing library staffing structure with some modifications. Budget reductions leading to significant tangible items to be catalogued coupled with the rise of Electronic Resource Management Systems may leave a window for a shift in duties toward the IR realm in cataloging departments. It is generally recommended that 1-2 FTE be devoted to the project. (See Appendix B for examples of DSpace job descriptions containing responsibilities that need to be addressed when developing a repository.) Since cataloging staff have extensive metadata experience both at the professional and paraprofessional level, they would be positioned to take the lead in maintenance and conversion. It is extremely helpful to have a librarian on staff familiar with programming and development experience with open access tools. User-friendly interfaces or submission forms can be programmed to facilitate metadata entry at the cataloging level or even more directly with the author. Some combination of these is most likely with some descriptive metadata elements and administrative metadata relevant to access restriction coming directly from the content author. Ultimately for a given digital object, a combined set of metadata will be applied with elements defined at the author, cataloging and technical support levels. Hiring student interns from various information system programs to be trained centrally but deployed in a distributed fashion to colleges and departments across campus is another staffing option.

Models

Having discussed the above issues, can the task force develop a few alternative models of how an institutional repository might be structured and staffed? What would be the start-up and ongoing costs?

The Task Force has looked at a variety of models that would provide best practices for consortial development of repositories. These include disciplinary models such as ArXiv, JSTOR and consortial models such as California Digital Library, Daedalus (http://www.gla.ac.uk/daedalus), Sherpa (http://www.sherpa.ac.uk) as well as the well-known DSpace Federation.

Please see the Task Force recommendations for next steps section below. Given that there are so many possible scenarios and combinations of scenarios, it is difficult to project specific start-up and ongoing costs at this time.

Cooperation in implementation
What cost savings, if any, could be realized through cooperation among some or all of the universities in establishing a repository? Is there a good basis of agreement about the technical and procedural issues that would be required in a collaborative effort?

As mentioned previously in this report, if NECOP chooses to work together on a repository, savings could be achieved by bulk purchase of commercially available solutions or, if going with open source, by locating hardware at one of the sites. Please refer to previous sections of the report for some specific hardware and licensing costs figures. We could also take advantage of the variety of staff expertise throughout the NECOP institutions.

The Task Force recommends that we examine best practices for development of repositories to develop agreement about the technical and procedural issues. The DSpace Federation has developed a model business plan that includes these topics. Library Technology Reports, vol 40, no 4 Establishing an Institutional Repository by Susan Gibbons has additional resource material.

Recommended next steps, or, how to get the train out of the station

Opportunities exist for the NECOP institutions to develop a repository, either jointly or separately, while taking advantage of shared expertise. Faculty have a growing interest in archiving their digital intellectual content (notably pre-prints and electronic publishing) and students at a minimum want to access electronic theses and dissertations. For it to succeed, there needs to be university-wide participation in the endeavor, building partnerships with computer centers, departments and colleges, institutes, and other areas of the university community that contribute to its intellectual output.

Librarians are uniquely positioned to lead this effort. We are known for collaborative work and have already developed liaisons with the potential partners. Librarians have the expertise and experience needed to collect, manage, organize and preserve intellectual content and can leverage this to manage the Universities' digital assets. As the edifice of scholarly communication changes and shifts this is also an opportunity for libraries to take a high profile role in the interests of the scholarly community we serve.

Challenges still abound, most notably:
• addressing long term preservation (eg. addressing "bit rot"),
• finding the perfect search tool to avoid creating an "institutional dump" (Google Scholar?)
• gaining broad based participation by the Universities' authoring community (eg. marketing this service to faculty and other contributors to ensure their participation)
• funding support at the university, not the library, level to ensure longevity and build the trust of the community

Key Decisions

A NECOP decision is needed on whether to proceed with a collaborative project, separate projects, or a combination of the two. If there is agreement to move forward on a joint repository effort, the Task Force recommends that we continue the momentum built by this report in one or more of the following ways.

Other key decisions that need to be made prior to such project(s) include:
• Identify inter-institutional lead team for the project if different from the Task Force

• Determine vendor hosted, locally implemented, or combination of solutions if there are constraints here
• Determine areas of expertise, including staff, at each institution
• Determine timeline for moving forward with identified project(s)
• Examine other consortial models identified in this report for best practices

Next steps

The Task Force recommends a "kick off" meeting during the spring 2005 semester, preferably March, so that a project could be started as soon as possible.

Purpose of "kickoff" meeting: to provide a forum to move forward the conversation among the library staff from all participating institutions. The intent would be to have an all-day meeting to address both concept and practice in the developing IR environment. It would be an opportunity to promote integration within the existing library staffing structure, the most practical option even at a pilot project level given budget constraints.

Program:
The morning session would be devoted to general, broad concept of NECOP repository and current state of the art summary. We could pull in an outside "expert" in the repository field to discuss scholarly communication, specifically how repositories fit into that conversation and why libraries are key to this effort. Possible speakers include Ann Wolpert, MacKenzie Smith, Susan Gibbons. The afternoon would be concurrent sessions of working groups to take the conceptual framework and start putting together action steps, determine expertise we have, what's still needed, and the like. Background materials would be disseminated by the Task Force prior to the meeting so that workgroups would have some shared materials with which to begin their discussions. These materials would be delivered by mid-February.

Working groups:
Members of these groups would be taken from the staffing areas of expertise identified above
• Cataloging for metadata and access point issues
• Reference and Collection Development for marketing and content issues
• Special Collections/Archives for long term preservation issues
• Systems for technical issues
• General Oversight working group with a library director liaison.

Location (suggested):
New England Center at UNH

Repository "communities":
Because of their common missions, the land-grant universities of New England have potential research areas that could contribute to a joint institutional repository. Suggestions for initial repository communities:

• Co-operative Education publications and research. Although the focus of Co-operative Education may differ from institution to institution, the interest in publications and research will overlap.
• Science research materials, especially research done across the region or departments in the different institutions that do common or related research. One example would be the oceanic research done by the Universities of Maine, New Hampshire and Rhode Island.
• Electronic theses and dissertations submitted by the institutions.

• Digitized collections of common interest, such as government documents holdings of specific interest to the New England region and of particular interest to the institutions' research communities.

Another next step for the NECOP institutions is to look to the recent COIC meeting at the University of Maine as a model for similar events at the other institutions if this has not already been done. This type of event can build institutional support, help define local requirements through surveys, and determine where pilot projects would be the most effective. The primary audience for these meetings would be faculty and other content providers.

The members of the New England Land-Grant Universities Institutional Repository Task Force have appreciated the opportunity to work on this report and look forward to discussing it with you. Please let us know if we can provide additional information.

Submitted Dec 15, 2004 by the New England Land-Grant Universities Institutional Repository Task Force:

Peter Murray, University of Connecticut
Sharon Fitzgerald, University of Maine
Marilyn Billings, University of Massachusetts
Elizabeth Slomba, University of New Hampshire
Christopher Burns, Pongracz Sennyey, University of Vermont

Appendix A: Example of OAI metadata record from the Maine Music Box collection


```
<?xml version="1.0" encoding="UTF-8" ?>
- <OAI-PMH xmlns="http://www.openarchives.org/OAI/2.0/"
  xmlns:xsi="http://www.w3.org/2001/XMLSchema-instance" xsi:
  schemaLocation="http://www.openarchives.org/OAI/2.0/ http://www.
  openarchives.org/OAI/2.0/OAI-PMH.xsd">
    <responseDate>2004-12-03T16:20:39Z</responseDate>

    <request verb="GetRecord" identifier="oai:musicbox.library.umaine.edu:
      KAS_001809" metadataPrefix="oai_dc">http://mainemusicbox.library.
      umaine.edu/musicbox/ASPOAIDP-DB/oai.asp</request>
  - <GetRecord>
    - <record>
      - <header> <identifier>oai:musicbox.library.umaine.edu:KAS_
          001809</identifier>
          <datestamp>2004-07-31</datestamp>
          <setSpec>KAS</setSpec>
        </header>
      - <metadata>
        - <oai_dc:dc xsi:schemaLocation="http://www.openarchives.
            org/OAI/2.0/oai_dc/ http://www.openarchives.
            org/OAI/2.0/oai_dc.xsd" xmlns:oai_dc="http://www.
            openarchives.org/OAI/2.0/oai_dc/" xmlns:xsi="http://
            www.w3.org/2001/XMLSchema-instance" xmlns:
            dc="http://purl.org/dc/elements/1.1/">
            <dc:title>In Native Worth : Creation</dc:title>
            <dc:creator>Haydn, Joseph 1732-1809</dc:creator>
            <dc:creator>Pearce, S. Austen (Stephen Austen) 1836-
              1900 Arranger</dc:creator>
            <dc:type>Sheet music</dc:type>
            <dc:publisher>London : Addison, Hollier & Lucas</dc:
              publisher>
            <dc:date />
            <dc:subject>Piano music</dc:subject>
            <dc:identifier>http://musicbox1.library.umaine.edu/
              musicbox/pages/full_record.asp?id=KAS_001809</
              dc:identifier>
          </oai_dc:dc>
        </metadata>
      </record>
    </GetRecord>
  </OAI-PMH>
```

Appendix B: Examples of DSpace job descriptions

According to the DSpace Business Plan Final Report, the following two positions are key to ensure the long term sustainability of the institutional repository. The people in these positions provide the oversight of the project as a whole coupled with moving the project into the daily operations of the library.

DSpace System Manager

Reporting to the MIT Libraries Associate Director for Technology as part of the Libraries' Systems Office, the DSpace system manager has primary responsibility for as aspects of the technical management of DSpace. The position will require a knowledgeable, enthusiastic, and self-motivated individual.

Responsibilities include:

- Primary responsibility for all aspects of the DSpace system's technical management
- Write and maintain technical documentation for the system
- Coordinate with DSpace-related research and development projects, and work with developers on integrating new functionality into the production system
- Coordinate related work by the Libraries' Systems Office and IS staff, including training
- Manage the DSpace open source code process and work with federation partners to successfully deploy and manage the system
- Perform system monitoring, testing and debugging
- Develop portions of DSpace related to system administration
- Monitor and upgrade DSpace utility programs and middleware
- Develop approved system enhancements
- Manage hardware contracts and system administration tasks for DSpace servers, documenting operational issues
- Assist with other library systems development projects as appropriate
- With the Product Manager, represent DSpace to the general library community
- Serve as backup for the User Support Manager during absences
- Participate in projects and teams working on activities related to DSpace

Qualifications:

- Master's degree in computer or library science, or equivalent experience
- Experience programming (especially java), working as part of a larger development team, and manage code written by others
- Understanding of network (especially web) development issues
- Experience with Unix systems and basic system administration skills
- Understanding of library mission and ability to communicate operational requirements clearly to key library staff at MIT and other institutions
- Excellent communication and interpersonal skills

Desired:

- Experience building and maintaining complex library systems
- Experience with Open Source development projects and procedures by which source of a project is shared with the community

DSpace User Support Manager

Reporting to the MIT Libraries Associate Director for Technology as part of the Libraries' Systems Office, the DSpace User Support Manager has primary responsibility for managing the communication with and support of DSpace users (both submitters and researchers). The position will require a knowledgeable, enthusiastic, and self-motivated individual.

Responsibilities include:

- Primary responsibility for all aspects of the DSpace system's user management, including coordination and communication with Libraries Public Services staff and Subject Specialists on support for MIT DSpace communities
- Write and maintain user documentation for the system, including system help pages
- Perform user training for library staff and other users (submitters, researchers)
- Coordinate and manage the definition and setup of new DSpace communities
- Perform outreach to MIT community, including faculty site visits and open training
- Plan and implement usability tests of the user interface and recommend user interface changes based upon test results. Coordinate with user interface developer to implement changes
- Make recommendations on new functionality for DSpace based on feedback from faculty, submitters and the MIT Libraries and serve as liaison with the DSpace Steering Committee to develop policies and prioritize enhancements
- Work with public relations organizations (e.g. Libraries' Communication Coordinator, Provost's Office Communications Officer, News Office, etc.) to publicize DSpace
- Coordinate importing of historical collections into DSpace with the System Manager, including collection assessment, metadata consulting, conversion referral, and developing metadata crosswalks if necessary
- Provide consultation on Institute policies and legal and regulatory issues related to intellectual property and sponsored research as they relate to DSpace
- Work on projects and teams with library and other groups at MIT who are closely aligned with services provided by DSpace (e.g. Metadata Advisory Group, OCW initiative)
- Chair the DSpace Advisory Group and participate in the DSpace Policy Committee

Qualifications:

- Master's degree in library science, or equivalent experience
- Experience with using, and helping others in an academic setting use web-based software
- Extensive knowledge of library practices and goals, especially with regard to technology
- Working knowledge web-based publishing tools and practices (e.g. HTML)
- Excellent written and oral communication skills and interpersonal skills
- Understanding of library mission and ability to communicate system mission and functionality clearly to key library staff and users at MIT and other institutions

Desired:

- Experience supporting complex library systems
- Knowledge of the MIT community and research interests

Appendix C: Status report at NECOP institutions

December 2004

University of Connecticut:
UConn is going to implement Digital Commons from ProQuest for at least a trial period of a year. We're going to use that time to gain ground on building communities, objects, marketing and policies. At this point the universities' and libraries' IT staff are tied up with other projects and don't have the concentrated blocks to time to implement an open source solution. At the end of the trial period, our IT staff will have caught up some and be in a better position to consider an in-house versus out-source question.

University of Maine:
See next page

University of Massachusetts
The UMass Amherst Libraries have begun the conversation of the changes in scholarly communication with faculty by hosting two colloquia. The first colloquium, held in the fall of 2002, addressed the dramatic rise in the cost of scholarly publishing and introduced the open source concept using the Bio One model. This past fall the colloquium addressed the issues confronting university presses with copyright, open access, and scholarly communication. There is a colloquium planned for the spring to address open access initiatives followed by a colloquium next fall on institutional repositories. The UMass Amherst Libraries have done a number of digitization projects in the Special Collections and Archives Dept. including a finding aids project in association with the Five Colleges. http://www.library.umass.edu/spcoll/spec.html

University of New Hampshire:
The University of New Hampshire is undergoing a campus-wide evaluation of information management systems. Institutional Repository software is being considered during this evaluation. Any implementation or planning for an IR is on hold during this process.

University of Vermont:
At the University of Vermont a Digital Center (UVM-DC) has been established comprising multiple interested entities in campus (such as the Library, assorted departments and the Computer Center). UVM-DC has investigated some software packages and has created a DSpace prototype repository for a diversity of content. Use of the prototype has varied widely and there is no concerted effort to establish an official institutional repository at this time.

INSTITUTIONAL REPOSITORY:
CONCEPT DEVELOPMENT AT THE UNIVERSITY OF MAINE

I. Self Archiving

In the fall of 2003, the institutional repository concept was not widely recognized at the University of Maine. Library staff began by contacting liaisons in academic and administrative departments to see what efforts might be taking shape. We found that the majority of potential stakeholders were not familiar with an institution-based model. We did find a number of faculty, particularly in the fields of Physics and Chemistry, who were contributing to subject based repositories.

More frequently we found faculty and research institutes on campus were actively involved in establishing local digital archives for their current materials. Again, many of the ongoing efforts are in the sciences. Maine Sea Grant has an archive portal established for news items, research and technical reports. The Maine Agricultural and Forest Experiment Station has a comprehensive catalog of materials available online in .pdf format . The Department of Spatial Information Science and Engineering has made .pdf versions of their theses and dissertations available. These examples have not deployed OAI tools and thus lack the interoperability that is desirable in an institutional repository model where different audiences will be drawing on content with a variety of goals in mind.

On the horizon is a collaborative project proposal involving several departments, including the library, to convert autonomous data collection into a shareable repository that supports synthesis of data through new metadata structures based on events and processes. The institutional setting is the data and data-gathering activities of agencies, and academic and research institutions operating within the Gulf of Maine watershed. Library and spatial information scientists working jointly with domain scientists will coordinate the metadata development.

II. Education and Awareness Efforts

a. Faculty survey - To further the objective of campus wide awareness of the potential of an institutional repository, the library elected to conduct a survey. We contacted MIT, and spoke with library staff actively involved with the DSpace project who graciously shared their survey tools. We elected to use a survey that would include a brief definition of the concept and several select links to IR models. We encouraged faculty not familiar with the concept to follow links to selected background papers. We asked the Provost's office to endorse the survey and distribute under his cover letter both electronically and through traditional print channels (http://library.umaine.edu/IRsurvey.htm) .

The survey was introduced November 15, 2004 just a few days prior to the Conference on the Intellectual Commons. As has been widely cited in the literature alternative publishing models and many institutional repository projects work hand in hand.

Our survey has been publicly available for only two weeks and as of this writing we have had 45 responses. Early next week we will send a second wave of reminders for participation, trying to catch the best window between end of classes and finals. We have faculty from a wide range of disciplines reporting thus far including: Art, Chemistry, Marine Biology, Engineering, Languages, Physics, Communications, Social Work, Environmental Science, Computer Science, Political Science, English, Education, Business and New Media.

Initial analysis indicates that 37% of respondents were aware of the IR concept and 92% are willing to participate. 95% would like to remain informed of future IR developments and 46% and 48% offered to serve as a resource person and participant in future discussions respectively.

Ease of location (navigation) was considered the most essential element in perceiving an IR as a valuable tool. This was followed by central access, preservation and the availability of non-traditional material types such as large datasets and rich media formats. Of least importance surprisingly was low effort in making personal research available.

There were several opportunities for respondents to submit general comments. Many felt the institutional repository was a good idea and enthusiastically support the library's central role of maintaining and preserving the materials. The concern that has come forward most often is the cost and what library services might be neglected to meet this new responsibility.

b. Conference on the Intellectual Commons (COIC) – November 20, 2004 the Library co-sponsored the Conference on Intellectual Commons with the University of Maine Information Science Collaborative and the Technology Law Center of the University of Maine School of Law.

The Conference on the Intellectual Commons presented emerging options for scholars, teachers, authors, artists, entrepreneurs, non-profit agencies, and others to access or re-use work produced by others in new ways, and to make their own work more widely available through the information commons. Presentations from Leaders in the field of scholarly publishing and open source software implementation included: Hal Abelson (MIT), Peter Suber (SPARC) and Jean Claude Guedon (University of Montreal). Attendees included faculty and other academic professionals engaged in scholarly publication projects.

University of Maine professor Harlan Onsrud led a summary session on next steps at the end of the day that engendered a lot of discussion and brainstorming for local action. Chief among the potential action items on a collective level were creation of an institutional repository for the creative scholarly output of faculty and students that supports increased accessibility and long-term archiving. This would move forward parallel to development of an institutional policy that encourages creators to participate in open licensing environments while ensuring that copyright clearly resides with creators. Those who attended the summary session agreed to further participation in discussions and pilots to forward both of these goals. See Conference site at: http://library.umaine.edu/COIC/default.htm

c. Collaborative meetings – The Office of Research and faculty from the Business School recently held a meeting to discuss a knowledge portal for the University. The Office of Research is most keen on the idea as a recruitment tool to heighten awareness of current research activities across the campus. Current efforts from independent "silos" are uneven at best and the Office is willing to lend support to a coordinated approach. This goal is but one of many potential facets of an open source repository and library staff took the opportunity to gain support for the IR at the meeting. Subsequent to this administratively convened session and the Intellectual Commons conference, a meeting is being scheduled to get the major stakeholders involved with university knowledge portals/ repositories together to make sure that everyone's needs are met, that we are not duplicating efforts, and to facilitate a seamless collaboration of knowledge management.

III. Library Initiatives

Fogler Library has been involved in several discrete projects deploying extensible markup language (XML) and other OAI tool sets to develop an infrastructure for content management and discovery with an indefinite life cycle. Notable examples follow.

a. **The Maine Music Box**, a project to explore how the teaching of music can be changed using digital technology. The collection consists of approximately 22,000 pieces of historical and popular sheet music published and widely played from the mid-19th century until approximately 1990. The strength of the collection is in music published between 1920 and late 1990. Very early on in the project we realized that changing the way in which educators approach the teaching of music, and affecting the way in which students learn, necessitated access to a collection of music from a range of musical styles and types, as well as several formats. Working with music faculty, we created custom digital collections for their use, including associated files in MIDI, Scorch and video formats. Using the Open Archives Initiative Protocol (OAI), records from the Maine Music Box database are also being harvested by the Sheet Music Consortium OAI Project, based at the University of California, Los Angeles (UCLA, Library of Congress, University of Indiana, and Johns Hopkins University. Implementation of OAI also enables us to contribute records, which are harvested for an IMLS project that is building a collection registry and item-level metadata repository of IMLS digitization projects at the University of Illinois at Champaign- Urbana.

b. University of Maine Museum of Art / Robert Carr Collection. URSUS catalog was enhanced with 2,400 records representing the principle collections of the

University of Maine Museum of Art. Within the collection, 300 artworks in the Robert Carr collection were photographed, and images linked to the bibliographic records. A digital gallery of the Carr Collection is also being implemented. Users can browse thumbnail images and bibliographic information by style of art, and link to a full image.

c. **Hudson Museum / William P. Palmer Collection** URSUS catalog was enhanced with 3,300 records representing artifacts in the Hudson Museum's William P. Palmer III Collection. Photographs of the artifacts are linked to the respective bibliographic records, which include research notes when available. A digital archive of the Palmer Collection images integrated with associated teaching materials for middle school grade levels will also be implemented.

d. **Previous digitization projects** using a combination of SGML, XML, and HTML encoded finding aids have included content from the William H. Cohen Archives and other distinctive collections such as the Maine Folklife Archive.

e. **Digital Object Management System.** With a growing number of discrete grant funded digital collections well underway or completed, we determined it was time to explore incorporating them into an aggregated database. All of them had descriptive metadata elements that could be drawn upon an mapped to the unqualified Dublin Core element set. We first examined the existing indices and and created a mapping grid that was subsequently programmed using using ASP.NET, the next generation ASP used by the current library web site. The feature-rich application environment and advanced services (such as XML services) provided by the .NET Framework facilitate building, deploying , and running public web based. A test version was created.

The full text search functions of the SQL Server were then integrated to the user interface. With Full-Text Search, end users can search through text stored in the database even when the text is contained in a formatted document.

The requirements of the Open Archives Initiative harvester were analyzed to ensure that the DOM database is suitably structured to receive output from the harvester. Usage of the .NET XML services to build a OAI data provider are currently underway.

Proper (standardized) database structures and programs have been implemented to harvest data from local collections into the DOM's database as future collections are identified and new content added.

Sharon Quinn Fitzgerald 12/3/04

Repositories

New Beginnings for Libraries

Presented on behalf of NECOP by
Marilyn Billings, University of Massachusetts
January 27, 2005

New England Council of Presidents NECOP

Voluntary organization of the six public New England land-grant research universities

- University of Connecticut
- University of Maine
- University of Massachusetts Amherst
- University of New Hampshire
- University of Rhode Island
- University of Vermont

Repository Task Force established Feb 2004

Institutional Repository

What is it?

"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."

Clifford Lynch, ARL Bimonthly Report 226

Services

- Intellectual output
- Teaching and learning materials
- Electronic portfolios
- Faculty legacy content
- Institutional archives, events, performances

Why an IR?

- Scholarly communication paradigm shift
- Library role as steward of scholarship
- Showcase for scholars and institution
- Wide dissemination and preservation

Core features

- Digital content in a variety of formats
- Community focus
- Institutional support
- Durable, permanent content
- Enhanced access

Core functions

- Material submission
- Metadata application
- Access control
- Discovery support
- Distribution of content
- Preservation, durability

Software solutions

- Open Source examples
  - DSpace
  - EPrints
  - Fedora
- Commercial examples
  - CONTENTdm
  - Integrated Library System vendors
  - UMI/ProQuest Digital Commons

Possible content

- Theses and dissertations
- Research and technical papers
- Co-operative Education publications
- Digitized collections, New England

Key Unresolved Issues

- Control and ownership
- Faculty participation
- Adherence to standards

Factors for Success

- Participation
- Simplicity
- Personal stake
- Creating a sense of community

Thank you for your attention!

- Contact information
  - Marilyn Billings, Coordinator, Five College ILS Dept & Special Projects Librarian, UMass
    mbillings@library.umass.edu