

To: Mary Case
University Librarian

From: Institutional Repository Task Force

Date: August 1, 2005

Re: Final report of the Library's Institutional Repository Task Force with recommendations for a campus committee

The Library's Institutional Repository Task Force (IRTF) was formed in March 2004 and charged with developing preliminary plans to establish an institutional repository at UIC, implement a pilot project, evaluate and modify the project, and then recommend a process for large-scale campus-wide implementation.

The Task Force believes that it has carried out its charge and this report summarizes its activities and recommendations.

Name:

The proposed name is: UIC Central Online Repository (UIC COR).

Definition:

An online repository of scholarly and institutional material generated by UIC faculty and researchers that will be cumulative, open, and accessible.

MIT, designers of the DSpace software, defined a university-based institutional repository as "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" (<http://dspace.org/introduction/irs.html>).

Summary of Activities and Decisions:

The IRTF created a Webpage that includes the charge, definition, goals, listserv, members, reports, minutes, and links to relevant websites (<http://www.uic.edu/depts/lib/staff/commwork/irtf/>).

IRTF evaluated a number of the most widely used repository software packages and evaluated established institutional repositories at several universities. The IRTF developed a list of criteria for evaluating institutional repository software and the university sites with institutional repositories, including flexibility, support, ease of entering data and files, range of file formats accepted, costs, and open source versus proprietary software. In addition, the software was required to be compatible with

national standards that have been established for institutional repositories. A detailed list of standards is in Appendix A.

Several task force members attended a workshop in the fall of 2004 that highlighted issues related to implementing institutional repositories and the various software options available.

After completing its evaluation, the IRTF recommended to the Library Steering Committee in December 2004 that the open source software of DSpace be the software of choice for UIC COR. DSpace is operational at MIT (<https://hpds1.mit.edu/index.jsp>) and at over 90 institutions worldwide <http://wiki.dspace.org/DspaceInstances>. The Steering Committee accepted the IRTF's recommendation.

In January 2005, the IRTF held two open sessions for UIC library faculty and staff to discuss plans for UIC COR and the decision to use DSpace software. The discussions at these sessions centered on the need to obtain university and faculty support for this project, and the need for faculty to understand the importance of an institutional repository.

DSpace software was loaded on a UIC library server in the spring of 2005. Several test files were loaded and functions were tested. Functionalities that were not working properly were corrected by library staff in the systems department with the help of the support network that is part of the DSpace federation. The electronic documents that were part of the pilot projects include a variety of formats.

The IRTF developed a list of appropriate and inappropriate material for UIC COR (Appendix B).

Pilot Projects were all appropriate material for UIC COR. The 36 items in the pilot project included:

1. Sample issues of Library-News, a regularly published electronic newsletter that highlights library collections;
2. Sample issues of ULIB, an internal electronic library staff newsletter;
3. Faculty Corner E-Books, an early library-sponsored electronic repository;
4. Recent articles by the Library of the Health Sciences librarians published in the *Journal of the Medical Library Association* (a publication with author-held copyright); and
5. University archival material, including records from the Office of Capital Programs, athletics (football), building photographs, and reports from the Office of Technology Management.

Several collections of material were identified that could be additional pilot projects for UIC COR. Some of these are not held by the library:

1. E-Theses from master students in CUPPA;
2. Urbana, COM, electronic graphics (<http://www.med.uiuc.edu/>), including links to two resources: The Internet Atlas of Histology and the Urbana Atlas of Pathology;

3. UIC political data;
4. Deans, Directors, and Department Heads memos in electronic format;
5. A database of masters projects, theses, and PhD dissertations in Public Health Nursing produced in Illinois (from a project with Professor Dorsch as PI in Peoria, Library of the Health Sciences);
6. UIC 2010 Strategic Thinking; and
7. UIC Library TV 2010.

Recommendations:

The IRTF recommends that the Provost:

1. Establish a campus-wide Institutional Repository Oversight Committee (IROC) to implement UIC COR.
2. Charge IROC to implement UIC COR using DSpace, the open source software developed by MIT; to review and revise policies and guidelines developed by the Library's Institutional Repository Task Force; to develop a business model and governance structure; to identify faculty to approach for the second phase of the pilot test; to develop a program to promote faculty buy-in; to promote UIC COR on campus; to design webpages for UIC COR; and develop a training program for faculty.
3. Support UIC COR by providing a server and a JAVA programmer to update, change code, problem solve, customize, add modules, as developed, and work with DSpace software developers, as needed. Most universities that have successfully implemented DSpace have the support of one FTE JAVA programmer.
4. Support holding a number of focus groups with faculty to make sure faculty needs and concerns are understood and met.
5. Appoint members to the IROC that have wide campus representation: a representative from the Provost's Office, from ACCC, from the Graduate College, and from the Library as well as representatives from such units as the humanities, the sciences, and the west side colleges.
6. Appoint a subcommittee within IROC comprised of Library staff: the JAVA programmer, select members from the IRTF, and public service and collection development representatives to:
 - a). develop guidelines for processing electronic documents and adding metadata;
 - b). review submitted material;
 - c). develop criteria for removing documents; and
 - d). develop criteria for acquiring software with deposited material (to access the material).

From the library staff meetings, the workshops IRTF members attended, and the evaluation of institutional repositories at a number of universities, the IRTF identified a number of issues that the IROC will need to address:

1. Training faculty to make sure only appropriate information is added to UIC COR;
2. The institutional repository requires extra work for faculty;
3. Material in the institutional repository is not peer reviewed;
4. An institutional repository is not a storage space for works in process; and

5. Long term preservation of digital objects is hard to guarantee.

These concerns might be ameliorated considering that:

1. The Library is a trusted partner in preservation;
2. The institutional repository is a way to increase faculty recognition;
3. Faculty do not have to keep their own data on their computers; and
4. Information from the institutional repository can be accessed from anywhere.

Website:

IRTF homepage: <http://www.uic.edu/depts/lib/staff/commwork/irtf/>

UIC COR homepage: <http://dspace.lib.uic.edu:8080/dspace/index.jsp>

Software:

DSpace (open source software developed by MIT): <http://www.dspace.org/>

MIT's Dspace: <https://dspace.mit.edu/index.jsp>

Listserv:

Public: cor@uic.edu

Task Force: irtf@uic.edu

Submitted by Task Force members:

Ann Weller, *Convener*

David Austin

Douglas Bicknese

Sara Blaszcak

John Cullars

Richard McGowan

Mircea Stefanu

Lisa Zhao

Appendix A Standards for UIC COR

University institutional repositories generally meet the following standards and DSpace is in compliance with these standards:

XML and Qualified Dublin Core (metadata standards)

OAI (Open Archives Initiative) – Protocol for Metadata Harvesting (OAI-PMH)

compliant software

A range of acceptable formats (PDF-A, ASCII, JPEG, doc)

Ability to batch load

Unicode compliant

Controlled vocabulary (e.g., LCSH, MeSH, AAT {Art and Architecture

Thesaurus}, TGN {Thesaurus of Geographic Names}, TGM {Thesaurus of Graphic Materials}. and other vocabularies appropriate to disciplines)

Storing software to access deposited material

- Standard questionnaire/template for depositing material
- Software support from external vendor
- Protecting deposited material from marketers
- Open access (non-restricted access to all)
- Unix-based system

Interoperability with other CIC institutions for data harvesting:

- Security/backup/authentication
- Search functionality
- Metadata, software, platform
- Search functionality

Appendix B

Examples of appropriate and inappropriate material to add to UIC COR

Examples of appropriate material for UIC COR:

- Pre-prints, post-prints, technical reports, data sets, working papers, presentations
- E-theses and e-dissertations
- Journal articles published under open access policy, including articles submitted to PubMed Central under the NIH guidelines
- University publications, policies, and procedures
- Course syllabi, course content material and course management software
- Training material
- Faculty CVs, personal faculty websites
- Video, audio material
- Non-western characters
- Computer programs
- Oral histories

Examples of inappropriate material for UIC COR:

- Non-scholarly material
- non-UIC related material
- Material in which the submitter does not hold the copyright
- Confidential data gathered, restricted by the IRB