Contextualizing the Institutional Repository within Faculty Research

DLF Spring Forum April 10, 2006
Austin, Texas

Deborah Holmes-Wong, Janis Brown, Sara Tompson,
Overview of presentation

• Background
  – Definition
  – Peer institutions and institutional repositories
  – Needs assessment rationale

• Methodology

• Key Findings
  – Interviews
  – Focus Groups

• Next Steps
Definition

... 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. **Clifford Lynch**
Institutional Repository Usage

• Not being used by faculty

• As of February 2006
  – 315 institutional repositories in production
  – Containing 850,000 Items
  – Or fewer than 3,000 items/institution

Source: Institution Archives Registry
Needs assessment rationale

- Published literature on faculty needs scarce
- Most assessments written post implementation
  - Focus on recruitment & training
  - Best example Foster & Gibbons 2005
- Published requirements often gathered indirectly (faculty Web sites, directories, deconstructing existing systems)
- Published requirements seem based on library needs not faculty needs
Goals

- Better understanding of how faculty disseminate research
- Determine faculty receptiveness to IR
- Develop use cases
- Gather high level requirements
- Determine features for first phase
Methodology for Needs Assessment

• Literature review
• Faculty office visits/interviews
• Focus groups/small group discussions
Office visits/interviews

- Identify and contact faculty to schedule meetings
- Interview 10-15 faculty from a variety of disciplines
- 2-3 person interview teams
  - Interview
  - Record
  - Listen/Ask follow-up questions
- Summarize interview incorporating feedback from interviewee
- Summarize all interviews
Goal of Questions

Determine

• How/where faculty publish
• Importance of PhD student work
• Degree that collaboration needs to be included in software
• Open Access requirements of funding agencies
• Nature/importance of supporting materials
Faculty info wanted

• Over a dozen places where faculty can submit biographies or their work
• Eight of them on campus
• All entered separately

• Global Directory Service/Campus directory
• Personal Web site
• Departmental Web
• School Web site
• SMARTS/GENIUS
• USC Experts Directory
• Blackboard
• Automated annual review system
• SF424 grants.gov
• NSF
• NIH
• Discipline-based repositories
• ISI Highly Cited Authors
• Who’s Who in …
Key Findings - Faculty

• Need a compelling reason for IR
• Research and teaching primary focus
  – Not cataloging
  – Not data entry
• Do not write grants to cover labor costs for digitization and data entry
Key Findings – Faculty

- Believe publishers responsible for archiving scholarly record
- Peer reviewed article still “gold standard” for most
- Interested in space for
  - Supporting materials
  - PhD student research
- Not as interested in Open Access for their articles
Major Impediments to Open Access IR

- Primacy of peer-review
- Responsibility for archiving recorded knowledge
- Publisher definitions of original research
- Access needs pre/post publication
- Faculty understanding of copyright
- Human Subjects Research--Privacy and confidentiality
- Supplemental materials
Services Suggested by Faculty

- High quality scanning of print only documents
- Digital dossiers for PhD students in Humanities
- Search engine optimization consultation (E-Resources as well as Google Scholar)
- Catalog technical reports/working papers in Homer (and OCLC)
- Copyright check/clearance
Focus groups

• Goals
  – Prioritize requirements
  – Determine feasibility of related projects

• Methodology
  – Give faculty use cases
  – Ask them to sort “compelling” “ok” “not needed”
  – Discuss how & why they selected what they did
Sample use case
One IR model – Electronic CV

- Harvests metadata from Web of Knowledge
- Faculty member reviews and validates (checklist)
- Creates CV for faculty
  - DOI link
  - OpenURL link for USC
  - Option to upload content
- Sends new citations to faculty
- Sends regular reports to faculty on frequency their items are downloaded.
Highest priority

- Secure storage
- Persistent URL
- Faculty control access (including editing and deleting)
- Multiple users access private space
- Commitment to preserve and if necessary change file formats
Sample Requirements

- Supports USC's authentication protocols for single sign-on
- Faculty control access to their content
- Faculty member can determine what others in group can do with content (add, modify, delete, view, export)
- External invitees can have the same roles/permissions as internal invitees.
- People can have different roles in the system including editor, author, reader, administrator.
- Faculty can invite individuals to see their space and documents.
- Faculty control whether document is public or private.
- Access of unpublished/private content can be limited to group or individual.
- Faculty can invite individuals to see their space and documents from USC or other institutions.
- Individuals can be from USC or other institutions.
- For private documents faculty control who has access.
Sample Requirements

Uploading files as easy as adding a file using a Web-based mail client

- Faculty can skip adding metadata.
- Minimal metadata fields for adding metadata. One screen ideal.
- Prepopulate metadata for items added based on file names, file content, faculty member info and or collection info.
- Digital rights on material are verified automatically before content is made available publicly.
- Work at any location using any computer available (MAC, PC, Linux).
- If adding requires multiple screens, then include information about how many more screens, steps are needed to complete process, ie route map like those found in shopping cart transactions.
Sample Requirements

URL/URN/URI do not change over time

Provide a mechanism to keep URL/URN/ URIs constant for documents, datasets, created by faculty member and his research associates over time.

For items that have been cited on servers that no longer exist at USC, where possible redirects from the previous URL to a persistent URL.

Provide a mechanism to keep outbound links within research documents constant over time even if the URLs change.

Provide a mechanism for securing permission to license and house content if need be to keep it linked to research.

Provide a mechanism to fix broken outbound links.

Provide a mechanism to find broken outbound links.
Implications for IR

- Institution needs to
  - Determine role of library/academic computing in supporting research life cycle
  - Determine which computing activities are supported centrally
  - Determine which are IR functions
  - Position the IR within research life cycle
Implications for IR

Library/academic computing needs to

- Balance permanence/persistence with need to correct information
- Balance faculty need for security, privacy with institutional imperative for access
- Balance rights management/HIPPA/IRB issues with desire for open access
Next Steps

• Develop technical requirements from functional requirements
• Implement basic ingest interface
• Develop key policies and procedures
• Identify of early adopters
Appendix: Questions

• Tell us about your research.
• How do you disseminate your research findings?
• Follow-up: Do you make pre or post prints available on a departmental or personal Web site or discipline specific site?
• How important are conference proceedings, working papers and/or technical papers in disseminating your research?
• What role do graduate and doctoral students play in the research process? What do you see as your role in disseminating their research?
• Please describe how you work with other authors when you collaborate.
• How do you link your finished research papers to the data sets and other research findings that you use in your research?
• Which funding agencies have you worked with?
• What is the relationship between your research materials and your curricular materials?