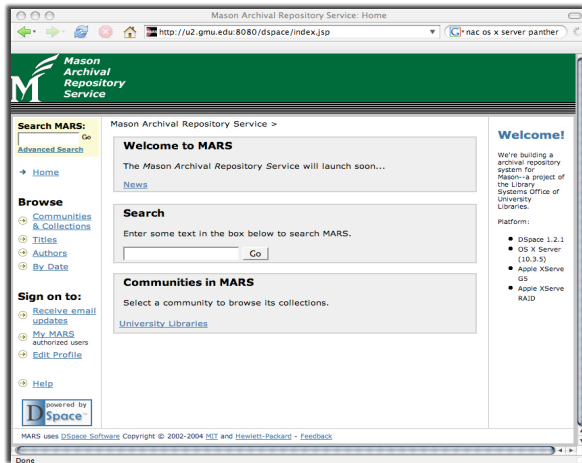




Leveraging Institutional Repository Technology to Address Archiving, Preservation and Access Objectives Across the University

Wally Grotophorst, Associate University Librarian - Systems
John G. Zenelis, University Librarian / Associate Vice President, IT



Executive Summary:

University Libraries recognized a range of digital preservation needs across campus--many specific to the libraries but not exclusively. Bringing together stakeholders and staff expertise, a university-wide Task Force was formed, led by the Library Systems Office.

The group created a website to pull together readings and useful links, identified candidate platforms and evaluated implementation options. DSpace was selected as the platform for several reasons:

- system architecture met needs beyond open-access publishing
- open source, OAI compliant, persistent web-accessible object identifiers
- increasing popularity in the academic community

Final report was shared with senior University administrators. Library was given responsibility for the initiative and new funding to cover startup hardware. A new Librarian position was authorized to assist in the management of this new library service.

1. Libraries recognition of multiple needs for digital preservation, archiving and access in the libraries and across the university.

- A previous commitment to the Center for History and New Media to explore preservation of objects in their ECHO project archives (www.chnm.gmu.edu).
- Several departments had expressed need for a publication system for working papers, individual research, etc.
- Multiple departments had approached library staff with need to preserve digital assets
- Library had archiving needs of digital objects in Special Collections (film, video, oral histories, images, scanned documents).
- Library needed to improve access and streamline workflows in processing digital objects
- A need to archive and preserve digital assets of the library (e.g., MARC database; purchased digital content, etc.)

2. Campus-wide Task Force was formed by University Libraries to explore issues and recommend solution(s)

Bringing together stakeholders and staff expertise we formed a task force drawn from library functional areas (Library Systems, Special Collections, Cataloging and Reference), departmental representation (Center for History and New Media); campus IT (database / web services expertise) and Copyright Office.

Task force built a website (<http://silo.gmu.edu/da>) of readings and relevant links; formalized system requirements; developed a "tiers of service" model; evaluated three candidate systems (Fedora, Greenstone and DSpace); recommended DSpace and developed a pilot project test plan.

Lessons Learned Thus far...

- Not all digital preservation needs can be met by DSpace, e.g., it is not very good at archiving websites.
- Mac OS X works as a server platform and XServe RAID works well too. We believe this platform offers a way for libraries to build a low-cost, high-performance DSpace system while reducing the level of UNIX system administration expertise required.
- Don't allow your IR system to be viewed as a "cost-free" digital storage service
- DSpace can serve as a "digital warehouse" of web-addressable objects...giving new options for system integration

"Institutionalizing" IR

The task force report was shared with the VP/CIO of the University (IT), the Provost (Academic) and the University Budget Group (Funding). The IR concept was approved by all and the library was given responsibility for what was viewed as a "university project."

Library has selected Apple Mac OS X Server and Apple XServe hardware for our production system.

A new librarian position has been added to staff to manage the service and provide outreach/audience-building activities working closely with our Faculty Advisory Group. Service is administered in the Library Systems Office.

3. Tiers of Service

Just as we determined that DSpace could be used to satisfy a variety of digital preservation, archiving and access needs, we also realized that no one level of service would fit every audience. Based on work done at Duke University, we adopted the following tier-based approach:

- **Archived** - Materials of significant and widespread value; complex, normalized metadata; commitment to periodic migration.
- **Preserved** - Materials have enduring value, but not enough to merit significant investment currently; basic metadata, supplied by content submitters; commitment to preserve in current format, but not migrate. The bulk of the repository's content will possibly merit this level of service.
- **Stored** - Materials not owned or managed by Mason, but which have long term value to Mason scholarship; mirrors of e-journals, other web sites, datasets, CD-ROMs, working papers, and so on. No commitment to migrate or preserve unless explicitly negotiated.