The Library Systems Office (LSO) is responsible for the development, installation, maintenance and enhancement of all information technology tools and systems within University Libraries.

The ITU is responsible for network infrastructure.
Hardware Environment

- Sun SunFire V880 (8 CPU) - Voyager
- Sun E3000 (4 CPU) - Voyager
- Sun E250 (2 CPU) - Library Web / MySQL
- Sun E250 (2 CPU) - SunRay Thin Client
- Sun E250 (2 CPU) - MasonLink+
- Sun E240 (2 CPU) - SunRay Thin Client
- Mac XServe G5 (2 CPU) - MARS
- Sun Sparc 20 - Intranet, mail relay
- Sun Ultra 2 (2 CPU) - EZProxy
- Locally built server (2 AMD Opteron) - Web Services
- 350+ Public and Staff workstations in 4 libraries
Software Environment

- Solaris 8
- Linux (SuSE, RedHat)
- BSD Unix
- Mac OS X Server
- Mac OS X (Panther)
- Windows 2000 Server
- Windows XP, Windows 2000
- Oracle, MySQL, PostgreSQL
- Apache / JBoss / Tomcat
- PHP, Perl, Samba
- Delphi
Milestones

• Online circulation system installed 1981
• First PC arrives 1982
• Card catalog replaced 1985
• NOTIS installed 1989
• CD-ROM network installed 1992
• First UNIX server installed 1993
• University website (www.gmu.edu) 1994
Milestones

• Voyager installed 1997
• E-Reserves system developed (OSCR) 1998
• SunRay ThinClient installation 2000
• MyLibrary@Mason 2001
• E-Journal Finder 2002
• Participant in LOCKSS project 2002
• OpenURL Link Resolver (MasonLink+) 2003
• MARS (Digital Repository) 2004
LSO Staffing

2 professional librarians
4 classified staff

• Wally Grotophorst - Associate University Librarian
• Lara Bushallow - Systems Librarian
• Pam Levy - Application Analyst (Voyager)
• Andrew Sikorski - PC Installation/repair
• Phat Le - PC Installation/repair
• Chandarari Chet - Systems Assistant
Mason has no student wages or GRA assistance in Systems. While other institutions do, they are not included in this graph.

Unlike Mason, most library IT units DO NOT host ILS system or provide sysadmin services.
National Rankings for Virginia Research Universities

Based on 2000 R&D Expenditures

**Top 100:**
- Virginia Tech: 51st
- University of Virginia: 58th

**Other Ranked Institutions:**
- Virginia Commonwealth Univ: 106th
- College of William & Mary: 157th
- George Mason University: 173rd
- Old Dominion University: 180th

Source: National Science Foundation, Fiscal Year 2000
LSO Services

- Voyager ILS (acquisitions, cataloging, circulation, ILL)
- Locally hosted web-based services (e-reserves, e-journal finder, ILL requests, MyLibrary@Mason, static web pages)
- Authenticated off-campus access to restricted content
- Mason campus directory
- OpenURL Link Resolver (MasonLink+)
- Institutional/Digital Repository (MARS)
- 220+ In-library public workstations and instruction labs
- 150+ Staff workstations, printers, scanners, etc.
LITA Top Trends - 2004

- Institutional Repositories
- Personal Search Software
- Federated Searching
- RSS
- RFID
- Copyright
- E-Resource Management

LITA is the Library & Information Technology Association (ALA)
Institutional Repositories

An 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.

— Clifford Lynch, CNI
University Libraries is building a digital repository service for Mason. Operational today, we anticipate an early 2005 roll-out.

Mason is one of the first universities to build this system using Apple OS X Server software and the XServe RAID platform.
“DSpace captures your data in any format--in text, video, audio and data. It distributes it over the web. It indexes your work, so others can search and retrieve your items. It preserves your digital work over the long term.”

-- from the DSpace “30 second elevator pitch”
Building On Our Strengths:
Digital Archiving, Preservation and Access


March 25, 2004

Wally Grotophorst, University Libraries, chair
Dan Cohen, Center for History and New Media
John Creuziger, Technical Systems Division
William Fleming, University Libraries
Polly Khater, University Libraries
Paul Koda, University Libraries
George Oberle, University Libraries
Lene Palmer, University Libraries
Angela Weaver, University Libraries
Thanks to compliance with OAI-MHP (Open Archives Initiative-Metadata Harvesting Protocol), items in our digital repository will be “found” by OAIster...and other metadata harvesting systems like Google Scholar and Yahoo’s Content Acquisition Program.
In response to an OAI-PMH request the system returns metadata about the contents of the collection...
Last month, we presented a session on Mason's work in the IR area at the SPARC / SPARC Europe meeting held in Washington, D.C.
Welcome to MARS

The Mason Archival Repository Service will launch soon...

News

Search

Enter some text in the box below to search MARS.

Search

Communities in MARS

Select a community to browse its collections.

University Libraries

Welcome!

We're building a archival repository system for Mason—a project of the Library Systems Office of University Libraries.

Platform:

- DSpace 1.2.1
- OS X Server (10.3.6)
- Apple Xserve G5
- Apple Xserve RAID

MARS uses DSpace Software Copyright © 2002-2004 MIT and Hewlett-Packard - Feedback
Are we ahead of the curve with this service?

Yes, but not by much...
On Monday (12/6/2004) this search on Google Scholar revealed there are already 1,500+ self-archived articles on various Mason servers.
Google has developed an algorithm that makes a calculated “guess” at what it thinks is scholarly material and makes that available to the end user via the Google Scholar interface.

Google has made arrangements with publishers (e.g., database vendors) to “crawl” and index their content, then provide information at the abstract level to the end user...pulling this content out of the “invisible” or “deep” web.

What version of content do you see? It depends on the source of the version you select and your relationship with its owner:

Full text of articles from open access journals and preprint repositories, as well as preprints on the web. For others, you should be able to get access to the full text if you or an institution you belong to has a subscription for the content. Regardless of the source, you should be able to see an abstract for any article, with the exception of those that are offline and referenced in citations only.

Interesting features:

- Citation counts
- Google requires content providers to include abstracts (at a minimum) in displayed “match” content
- Links can trigger a business transaction with the publisher
- Yahoo will likely enter this space soon (e.g., CAP program arrangements with NPR, Northwestern, OAIster, etc.)
- Exposes copyright-violations
Millions of researchers, scholarly writers, students, and librarians use EndNote to search online bibliographic databases, organize their references and images, and create bibliographies and figure lists instantly.

Instead of spending hours typing bibliographies, or using index cards to organize their references, they do it the easy way—by using EndNote!
University Libraries (George Oberle and Kevin Simons) is leading an joint effort (with Technology Across the Curriculum, DoIT, and the ITU Support Center) to support use of EndNote software across campus.

- site license for all students, faculty & staff
- training sessions
- online support center
The online support center is provided by the Library Systems Office.
“The Network is the Database”

- Library Catalogs
- Citation Databases
- FullText Databases
- E-Journals
- Aggregated Content
- Open Web
- “Deep Web”
- Digital archives
- Institutional Repositories
“Noise” in the system

- Many sources of information
- Many different user interfaces
- Content / format varies across sources
- Overlapping coverage
- Undocumented gaps in coverage
- Easy to miss content or be flooded with irrelevant information
- Must deal with inevitable duplication
The Problem

• Figuring out where to begin a search is a challenge for many users, particularly undergraduates.

• Selecting the “optimal” resources requires experience with organizational schemes imposed by the discipline, by libraries and by content providers.

• There typically is no “one best path.”
Solutions?

- Cross-database searching (e.g., federated searching, metasearching, or broadcast searching)
- Content linking
Federated Searching

- For the user, federated searching is a simple search box that retrieves content from a wide variety of databases and e-resource collections.
- Who does that sound like...?
But it’s not like Google™

- Google sends its robots to collect data from millions of web pages in advance, so the user is actually searching a cross-file index, not the content that created the index.

- Users connect to the full-text pages when they click on a hyperlink in the results set.
A Federated Search...

- Executes a cross-file query across citation and full-text databases that do not share a common thesaurus or index.

- A different search protocol may required for each source.
Search protocol challenges

- Some use variations of Z39.50 protocol which predates the web.
- Some use XML to identify the data elements being used.
- Many leave it to the search engine vendor to figure out how to access the content (screen scraping).
Myths and Realities of Federated Searching
Myth # 1
No Stone Left Unturned

- Not all federated search engines can search all data sources.
- Most can handle Z39.50 interfaces and popular (free) databases.
Myth #2
De-duplication really works

• True de-duping is virtually impossible. The search engine would have to download all results and compare them on multiple data elements (often poorly documented)

• The limiting factor is the way in which data sources return results -- 10 to 20 at a time. Completing a true de-dupe would take hours as a typical search might return hits in the thousands...
• When attempting to relevancy-rank citations, the only words you have to work with are those that appear in the citation. Often, the search word doesn't appear.

• The abstract and full-text data, as well as the indexing that content providers use to relevancy-rank their content, are unavailable to federated search engines.
Myth #4
E-Resource Management is Simplified

- A federated search engine searches databases that update and change an average of 2 to 3 times a year. A system accessing 200 databases is subject to 400 to 600 updates per year--better than two a day.

- When updates are not made, access to content is lost and expensive resources go unused.

- Maintenance of this software layer is expensive.
Our take...

- University Libraries believes that federated search engines *show promise* but we do not feel they are ready for anything more than experimental use.

- We are *monitoring developments* in this area closely and positioned to expand our use of the technology as it becomes useful to do so.
Mason is an active participant in the WRLC’s MetaLib project.
Content Linking
Federated searching tackles information overload on the front end--attempting to improve the efficiency of the discovery process.

Content linking focuses on the back end--you’ve found a useful source but how do you know that it is the most complete version available? And is there “related” information that you might be missing?
OpenURL

- This service relies on the OpenURL standard (Z39.88-2004)
- OpenURL is a syntax for encoding metadata into a URL so that it might be passed between systems via the HTTP protocol.

Sample OpenURL:

http://www.mysrv.org/menu?
id=doi:10.111/12345&
genre=article&
aulast=Weibel&aufirst=Stu&ISSN=35345353
&year=2001&volume=14&issue=3&spage=44
&pid=2829393&
sid=OCLC:Inspec
How does it work?
We’ve done a search in EconLit and found an interesting citation. We click on MasonLink.
Title: Saving, Dependency and Development

Author(s): Kelley, Allen C.; Schmidt, Robert M.

Affiliation: Unlisted; Unlisted

Publication: Duke University, Department of Economics, Working Papers: 95-01

Year: 1995

Abstract: It appears that the widely-observed finding in the literature showing little or no relationship between population growth (and dependency) and saving requires modification based on panel and cross-section estimation of aggregate country data. First, while it is consistent with an update of the hybrid (Keynesian, life-cycle) Leff-type model for the 1960s and 1970s, this is not true of the 1980s, when there is a "jump" in the importance of demography. Second, in a more clearly interpretable life-cycle framework (a la Mason), the role of demography increases smoothly and systematically
Citation Information

Article Title: Saving, Dependency and Development
Source: Duke University, Department of Economics, Working Papers; Issue: 95-01; Date: 1995

MasonLink+ is unable to make a direct link to this item.

If a journal article, try the Search E-Journal Finder link below.

Explore Related Information

Other Internet Resources
- Metacrawler - Title Search
- Google - Journal Title Search
- Search All The Web for Journal Info
- Google - Search Author
- Google - Article Title Search
- DAISter - Search Article Title
- All The Web - Author Search
- Publist - Search Journal Info
We find the “working paper” via OAIster.

Title: Saving, Dependency and Development

Authors: Kelley, Allen C.
         Robert M. Schmidt

Note: It appears that the widely-observed finding in the literature, showing little or no relationship between population growth (and dependency) and saving requires modification based on panel and cross-section estimation of aggregate country data. First, while it is consistent with an update of the hybrid (Keynesian, life-cycle) Leff-type model for the 1960s and 1970s, this is not true of the 1980s, when there is a “jump” in the importance of demography. Second, in a more clearly interpretable life-cycle framework (a la Mason), the role of demography increases smoothly and systematically over the 30 years, although the form which it takes (life-cycle versus lifetime level) also varies systematically over time. Overall, however, demographic factors accounted for a major portion of changes in saving across countries.
MasonLink+ is an OpenURL link resolver. It allows the user to move from a citation in one vendor’s database to the full-text version of the same content in another. It enables the university to get more value and use from the content we pay to license. It offers the library a tool (and development platform) to deliver related but remote content.
RSS is an XML-based format for distributing and aggregating Web content. The format can be used to “feed” RSS readers or other websites (e.g., portals, blogs, etc.).
University Libraries is currently using RSS to build dynamic content on some web pages.
In the future we expect to begin pushing other content to our users via RSS:

- New book and e-content lists
- Library News and Announcements
- Research / Reference Tips from Liaison Librarians

RFID
Radio Frequency Identification

University Libraries is not contemplating a conversion to RFID security in the near future but will consider the technology when we implement our next generation checkpoint control system.

RFID is expensive for larger libraries.

e.g., RFID tags cost $0.85 each
What’s Ahead?
By 2010...

We will have to expand and improve facilities, increase systems office staff and improve the skill level and IT literacy of library staff.
We have a solid record of identifying, adapting and deploying advanced technologies...

...but to do so on a larger scale and with capacities to satisfy a more demanding audience, we will require enhanced resources--both staff and equipment.
Current facilities for mission-critical LSO systems are inadequate. We recommend construction of a secure server area in the next structure built for library use.
In the interim, during 2005 we intend to co-locate LSO staff in Fenwick and Johnson Center to new facilities in Fenwick. This will:

- improve staff efficiency
- offer cross-training opportunities
- separate staff from areas where servers and storage systems are housed

Given staffing levels within ITU and the skill sets required, we do not believe relocating our servers to an ITU-hosted area would be beneficial to the university.
Hardware Funding

By 2008 all servers currently in use will have reached or exceeded their expected service lifetime. We will need to find a funding stream to replace them.

SunFire V880  $35,000
Three Sun E250’s  $24,000
Additional storage for MARS  $20,000+
Also true for our 200+ public workstations

As we move toward 2010, we should focus more attention on providing robust, secure wireless connectivity to our in-house users and recalibrate the number of library-provided workstations we feel we must offer.
In Summary

• Library should maintain an in-house IT operation (e.g., resist “outsourcing” either to other university operations or third-party services). The requirements of Library IT support are too specialized for generic IT assistance.

• We need to encourage central IT (ITU) to fully develop core technologies that can benefit the library (e.g., LDAP for authentication, high-speed access to off-site backup, etc.)
In Summary

• Staffing and support for Library Systems should be expected to increase between now and 2010.

Additional staff will enable University Libraries to embark on the sorts of initiatives one associates with libraries of research universities.

With workloads reduced, LSO staff can assume a greater role in staff development efforts--which can improve overall technology skills across professional staff.

• Improving LSO facilities must be an important component of new construction initiatives.
Questions?