



*How Will I Get What I Need to Know When I Need It? Information
Access for the Future.*

The Digital Library Federation: Mapping our Digital Future

David Seaman
Executive Director
Digital Library Federation

Transportation Research Board 83rd Annual Meeting
Jefferson East, Washington Hilton
7:30pm-9:30pm, 10 January 2005
Sponsored by the Library and Information Science in Transportation (LIST)
Committee

Digital Library Federation

<http://www.diglib.org/>

- 33 members – major academic and national libraries, including The British Library; 5 allies (CNI; RLG; OCLC; LANL; JISC)
- Created in 1996 by directors of US research libraries; fills a need not simply met by larger library organizations: focuses exclusively on DL needs and strategies for large libraries
- Nimble, agile, collaborative
- Practical and strategic areas of activity



DLF Work -- background

- User Services and Tools
- Metadata Standards
- Resource Management
- Production Best Practices
- Digital Preservation
- Digital Library Architectures
- Acronymic Promiscuity: OAI; XML; TEI; METS; MODS; MIX; XSLT; SCORM; DCMI; ERMI;



Finding Order in Chaos (embrace the churn)

- New library/information disciplines still solidifying; new skills sets and work habits
- Tipping points -- when does a new item move from irrelevant to “surprisingly non-terrible” to indispensable and how do you know?
- Non-library arbiters of access to scholarship
- Ambition, Ignorance, and Lack of Money
- Seismic events are routine and continuing: Mosaic; Google; eBay; PDA; wireless; blogs

Trend: Virtual Learning

- Blackboard/WebCT/roll your own
- OpenCourseWare at MIT
- Flecker/McLean DLF report: *Digital Library Content and Course Management Systems: Issues of Interoperation*
<http://www.diglib.org/pubs/cmsdl0407/>
- Libraries often absent from virtual learning
- SAKAI – a collaboration imperative at work
- Libraries have an opportunity to be in the classroom like never before.

Trend: Digital Preservation

- National Digital Information Infrastructure Preservation Program (NDIIPP)
- Global Digital Format Registry (DLF)
- Digital Curation Centre (JISC UK)
- Digital Preservation Coalition (UK)
- Preservation metadata/tools (New Zealand Natl Lib)
- PADI (National Library of Australia)
- OCLC; RLG; DLF; CLIR
- Cornell (excellent online training guide)
<http://www.library.cornell.edu/iris/tutorial/dpm/>

Trend: Institutional Repositories (DSpace et al)

- “There is an growing interest in the more coordinated management and disclosure of digital assets of institutions — learning objects, data sets, e-prints, theses, dissertations and so on.” *OCLC Environmental Scan*, 2003.
www.oclc.org/membership/escan/research/default.htm
- Resistance from faculty to being seen as asset workers for an institutional content management system.
- Core question: how is the arrival of the Institutional Repository (and Open Access) tied in with changes in the faculty rewards system? How integrated into institution’s reflection of valuable and rewardable contributions?

National, trusted repositories

- British House of Commons Science and Technology Committee. *Scientific Publications: Free for all?* July 2004. (HC 399-I).
- The U.S. House Appropriations Committee draft report, concerning NIH-funded research – requirement to deposit in PubMed Central.
- **New Zealand:** “NZD\$24 million funding from the latest budget will allow New Zealand to ward off ‘digital amnesia’, and protect New Zealand's documentary heritage.”
www.natlib.govt.nz/bin/news/pr?item=1085888952



Trend: Open Access

- Exciting glimpse of a future where all scholarship is free, with rich metadata that allows virtual aggregations of content held in seamlessly integrated IRs across the globe, all controlled by libraries and leading to a productivity boom and a global public informed by the best data there is.
- New roles for libraries – publishing from IRs; customer support. New roles for publishers – add value not control access.
- Investment needed to add value to public content favors largest publishing operations – threatens smaller and society publishing? May lead to fewer publishers?



Trend: Open Access

- Most journals now allow self-archiving
- Now it is up to us – no legal barriers to having (most) scholarship published in traditional peer-reviewed journals also freely published from our Institutional Repositories
- So, how are our scholar-authors motivated to self-archive, learn metadata skills, publishing skills? Why should they do it? Should they be encouraged to do this for the general good? How to reward them?

Trend: Mass

- *Digital Opportunity Investment Trust (DO-IT)*: \$20 billion “digital gift to the nation.” Massive digitizing for the public good
<http://www.digitalpromise.org/>
- US Government Printing Office: government documents print/film conversion – 2.2 million items
- Carnegie Mellon’s Million Books Project
- Google Scholar; Google Print; Google Library. Massive digitizing of library material, in and out of print. Seven million books from Michigan alone.

Trend: Harvestable Metadata

- Open Archive Initiative www.openarchives.org
- No longer enough to build to local standards and put it on the web
- Need to push out simple metadata for others to grab and use in service-building
- OAIster: <http://www.oaister.org>
4,879,071 records from 396 institutions (Jan. 2005)



Trend: Malleability

- Every publisher is an island; we produce silos of data that plays badly with others. A good silo is a lovely thing – but not sufficient always.
- The need to have content that encourages local re-organization and creation of services, and that permits “beyond browsing and searching” engagement by individual users



Trend: Malleability

- Little ability to work with content or even metadata cross-publisher and cross-aggregator. Or cross-library.
- Content too often web-bound only – need to be where user is (mobile, nimble – XML)
- Digital couch potatoes versus rip/mix/burn
- We invite our users to visit sites and watch content channels (TV); they want to sample, re-use and re-package as a personal library, a classroom presentation (the music mix)



Major Force: Time

- Users are simultaneously over-whelmed with the time it takes to find relevant information in a “data silo” landscape, and (outside the sciences) under-whelmed by the lack of good material in their particular discipline.
- TIME: 39% of all respondents (60% faculty) report insufficient time as their major problem [*Dimensions and use*]



Major Need: discovery

- Much more content, and much richer, domain-sensitive, finding systems are vital, as is the ability to enrich, re-shape, re-package, annotate, and contextualize the data once one has found it.
- Respondents frustrated with finding information, determining its credibility, and analyzing it [*Dimensions and use*]
- Richer search and visualization tools (IBM’s WebFountain; Grokker’s Visual Search).
- Persistent naming (DOI et al)



Closing

- The transformation from isolation to integration is our central challenge and opportunity– with some enormous payoffs when we get it right.
- Innovative users need malleable content with which to innovate; need to learn to re-shape content in a mutable library.
- It is not sufficient simply to offer the current fragmented set of websites defined by a publisher or aggregator. This frustrates use, wastes time, and hogties library services.
- Mass, malleability, and the collaboration imperative.