



**Mass, malleability, and the  
collaboration imperative:  
trends for the 21st-century library**

David Seaman  
Executive Director  
Digital Library Federation

Simon Fraser University  
Vancouver, Canada  
February 23, 2005

# 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 1995 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***

- *Dimensions and use of the scholarly information environment* <http://www.diglib.org/pubs/scholinfo>
- IMS –repository/courseware integration
- Distributed single collection of our own material

## ➤ ***METADATA STANDARDS***

- OAI-PMH (The Open Archives Initiative)
- METS (Metadata Transmission Standard)

# DLF Work -- background

---

- ***RESOURCE MANAGEMENT***
- XML format for license content (ERMI)
- Registry of Digital Masters
- ***PRODUCTION***
- Production benchmarks and good practices
- ***PRESERVATION***
- NDIIPP
- Global Digital Format Registry

# 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; wikis; blogs; Google Scholar/Print

# 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 Production

---

- Regularized production within the institution – from ad hoc projects to continuous process
- Regional production centers
- DLF/OCLC Registry of Digital Masters
- Special collections materials a focus of this activity
- A very long tail – surprising usage for materials that have no use in print when locked in academic libraries <http://www.wired.com/wired/archive/12.10/tail.html>
- Strong library/faculty partnership opportunities

# Trend: Harvestable Metadata

---

- Open Archive Initiative <http://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: 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.  
<http://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 tied in with changes in the faculty rewards system? How integrated into institution’s reflection of valuable, rewardable contributions?

# 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.
- New roles for libraries – publishing from IRs; the public face of scholarship for an institution; customer support. New roles for publishers – add value not control access.
- Investment needed to add value to public content favors largest publishing operations. May lead to fewer small or society publishers?

# Trend: Open Access

---

- Most journals now allow self-archiving (including all of Elsevier, Nature, etc)
- 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

*(not if but when and by whom)*

---

- *Digital Opportunity Investment Trust (DO-IT): \$20 billion “digital gift to the nation.”*
  - Digital Library Federation, the American Library Association, the Association of American Universities, the Association of Public Television Stations, the Association of Research Libraries, the George Lucas Educational Foundation, and EDUCAUSE endorse it, and senior personnel from eBay, Google, IBM, the Internet Archive, RealNetworks, and 3Com all in planning. (<http://www.digitalpromise.org/>)
- Whatever happens with it, its arrival on the scene in 2001 spurred us on to think about what we would do in the face of a massive public investment in digital content, tools, evaluation, and learning systems.

# Trend: Mass

---

- US Government Printing Office: print documents conversion – 2.2 million items
- Carnegie Mellon's *Million Books Project*
- Library of Congress and a group of international libraries from the US, Canada, Egypt, China and the Netherlands to make one million books digitally available on the Internet (Dec 2004)
- Google Scholar; Google Print. Massive digitizing of library material, in and out of print (Harvard, Oxford, Stanford, Michigan, NYPL).

# 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)

# Aquifer

---

- DLF Strategic Goal – a *Distributed Open Digital Library*: <http://www.diglib.org/aquifer>
- New level of interdependence
- Two-phase Finding System, initially OAI
- Digital Object Sharing for richer library services and better scholarship
- New infrastructure and data creation needs – what are the characteristics of sharable content?

# Closing

---

- Need to think strategically and focus on our core mission to advance pedagogy and scholarship
- Any library that can be replaced by Google, should be.
- 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.

# Closing

---

- Standards abound, and we are getting better at applying them in ways that work across institutions
- The days of competing on access are over – context, services, convenience, cataloging, research skills, long-term thinking are our edge.
- Collaboration is not just a nice thing – it is a survival mechanism
- Managing digital content over time is a tough business – and we are equipped to do it.
- Mass, malleability, and the collaboration imperative.