



Annual Report 2002-2003

**Initiatives, Publications, Contracts, and Grants,
and the
Activities of the Executive Director**

**Submitted to the DLF Steering Committee
by David Seaman**

EXECUTIVE SUMMARY

2002-2003 has been a lively and exciting first year for me at DLF. With help from Dan, and working closely with the CLIR and DLF staff, I have (I hope) effected a smooth transition from the previous directorship, maintaining the ongoing projects and publications, keeping up the high standard of the Forums, and moving us forward with new initiatives.

One of the great luxuries of the first few months of life as DLF Director is the “Grand Tour”, a whistle-stop journey around the country to meet the DLF directors, their staff, and to get a quick sense of their local realities. I benefited greatly from this flurry of visits, and (while I don’t intend to do a grand tour of that intensity every year!) I am committed to making personal visits to DLF members whenever I can, and I am always pleased to be asked to come to a campus to give a talk, take part in an event, provide an outside perspective, or simply to catch up with colleagues there. These visits are particularly important for the newer partner organizations, according to the feedback I have received, both to underscore for their staff the participatory nature of the DLF and to encourage them to put forward new initiatives.

A certain portion of my time goes to promoting the work of the DLF through writing and through speaking engagements, in order to communicate the results of DLF initiatives to a wider audience. On occasion I am invited to take part in various meetings germane to digital library activities, and I try to agree to attend the subset of those that are close to DLF initiatives and interests. Lists of these activities are appended, and the contacts I make during these meetings are already proving useful as we move forward. In particular I am pleased at our continued ability to attract non-DLF and non-US participants in our initiative working groups.

Communications continues to be a challenge, with numerous colleagues on many initiatives spread all over the country, and with the ongoing need to inform – and be informed by – the DLF directors and staff. The expansion of subscribers to the *dlf-announce* listserv has helped (now at 358 individuals), as have the recent development of new DLF “communities of interest,” the Digital Camera Operators group and the Developers Forum. With ongoing work on the website and the newsletter, I am continuing to look for ways to improve my communication with the membership, and to aid their communication with each other. I am particularly grateful for -- and encouraging of -- the feedback, comments, critiques, and encouragement I receive from individuals in the DLF across the course of the year.

Finally, the lively strategic planning session early in 2003 and the resulting work on the distributed library and the deep sharing of content has underscored for me the power and vibrancy of the federation of institutions and individuals that make up DLF. Whatever we eventually name it, the Distributed Library effort is for me the most exciting prospect (among many) as we move forward into 2003-2004, bringing us back to a foundation principle of the organization as expressed in the 1995 charter, drawing on much of the work, benchmarks, and standards we have fostered over the past seven years, extracting extra value from our existing digital collections, and giving us a large central challenge to engage and energize us.

1) DLF INITIATIVES

The following initiatives are ongoing or have seen activity this year.

THE DISTRIBUTED LIBRARY

This emerging set of initiatives cuts across the traditional areas of DLF activity – preservation, production, metadata, management, usability, and architecture – as befits an organization’s overarching strategic goal. Planning is well underway to build a unified sense of our holdings, including a finding tool that makes use of our investment in OAI and our collective expertise in building OAI harvesters, and both experiments and research into the deep sharing of our content for better library service provision designed to lead to more innovative teaching and scholarship.

PRESERVATION

Global Digital Formats Registry: [began January 2003] A central, trusted repository of information about the common digital file formats we are all ingesting into library repository systems. This information is central to any digital preservation strategy and it is wasteful for us all to run local mini-registries for this information. Participants include DLF members plus NIST, JISC, IETF, PRO, the Bibliotheque Nationale, and the National Library of Canada. Interest has been expressed from the British Library, and the Dutch and Australian national libraries. So far we have had two very effective day-long meetings, with use-cases being circulated. Consensus is growing on the shape and scope of the format registry and we expect to seek outside funding in 2003-2004. Contacts: Stephen Abrams, Harvard, and Mackenzie Smith, MIT.

Final Reports from the Mellon Electronic Journals Archiving Planning grants: [fall 2002-spring 2003] All available electronically from DLF website. These projects are being edited for consistency by Linda Cantara [Indiana University] and will be collected into a print volume with an introduction. [Fall/Winter 2003] <http://www.diglib.org/preserve/ejp.htm>.

Registry of Digital Masters (with OCLC): The Digital Registry Phase One Implementation Guidelines are complete and will soon be distributed to DLF members and the wider library community for use [November 3, 2003]. They use MARC 21 Format for Bibliographic Data elements and OCLC current cataloging system functionality. The Registry is designed for books and journals that exist in a page-image format, created in accordance with the DLF's *Benchmarks for Digital Reproductions*. The registry record will lead one to a use-copy (assumed to be web-based); the existence of the record gives one confidence that a preservation master exists; it does not guarantee that the master files are generally accessible.

Contacts: Susan Westberg (OCLC), Robin Wendler (Harvard), Rebecca Guenther (Library of Congress), Ericka Linke (Carnegie-Mellon), Glenn Patton (OCLC), Taylor Surface (OCLC), Bob Wolven (Columbia), Robin Dale (RLG), Lynn Marko (Michigan), Marty Kurth (Cornell).

PRODUCTION

Benchmarks for Digital Reproductions: This document recommends bit-density and resolution for page images of books and journals. Already being used as the desired specification for items added to the digital masters registry (December 2002 <http://www.diglib.org/standards/bmarkfin.htm>)

Digital Camera Operators Interest Group [May 2003] 18 participants attended a “Birds of a Feather” session for digital camera operators at the Spring 2003 DLF Forum and quickly affirmed that they need a way to exchange skills and advice, and that no such channel of communication existed for them. A listserv is in the process of being set up as a first step to support this production community.

Archivists’ Workbench: [2002] we funded a second meeting of the Archivists’ toolkit group (November 2002) as per the agreement of last year. This group is working on a grant application to continue their work (fourth quarter 2003) as a direct result of the DLF-funded meetings.

METADATA and MANAGEMENT

E-resources Management: This new DLF-sponsored project will develop a common, sharable, xml “record” for recording the content of license agreements, related administrative information, and internal processes associated with collections of licensed electronic resources. Contact: Tim Jewell, University of Washington. <http://www.diglib.org/standards/dlf-erm02.htm>

METS: version 1.2 open for comment at <http://www.loc.gov/standards/mets/>. DLF has provided travel and meeting space support for the international Editorial Board’s work this year; next year they anticipate needing some help with documentation. METS is being widely used in the library community, and has clearly met a need. Contact: Jerome McDonough, NYU.

Open Archives Initiative: With CNI, DLF provided funding last year and this year to support the development of OAI, resulting in the publication this year of *The Open Archives Initiative Protocol for Metadata Harvesting Protocol, Version 2.0* (2003/02/21), and a series of real-world applications, including the National Science Digital Library, work at Illinois to provide an OAI view of projects funded by IMLS, and a series of Mellon-funded OAI harvesting initiatives at various DLF partner institutions.

ARCHITECTURE

Ockham -- a think-tank group for architectural and technical issues confronting digital library developers. With support from DLF they have met in person several times (May 2002; September 2002; January 2003) and have a listserv or online discussion. The concrete result of this is a successful NSF grant to further their work, given to Emory University with Virginia Tech and the University of Arizona. [See *Grants*, below]

Courseware/Library Interactions Integration of digital library content into the courseware systems we are buying and developing is a general need across the DLF and beyond. In December I attended a meeting called by Don Waters to discuss OKI, the Open Courseware initiative, and related needs, and out of subsequent discussions with Don and Dale Flecker came the DLF grant to Harvard to examine the interaction of digital libraries and learning management systems. We will see ongoing work in this area, I am sure. I have also attended an IMS Global meeting, to begin to establish contact with some of the active parties in the courseware arena, and have met with Neil McLean, one of the authors of this year's IMS white paper on Digital Repositories. [See *Grants*, below]

Developers' Forum: this group of nearly 30 participants has met for half a day at both the fall 2002 and spring 2003 Forums. They are lively, free-flowing meetings, and the level of participation suggests that they are meeting a need. The precise nature and shape of this gathering is still shaping up – a listserv is being established to keep the conversation going and to help provide a more directed agenda for the next meeting. Contacts: Jerome McDonough (NYU) and Martin Halbert (Emory).

Fedora: A test bed repository for xml digital library objects. Funding this year has also gone to support a meeting of test implementers of the Fedora digital library repository software (October 2002). Further work is supported by an ongoing Mellon grant.

USABILITY

Dimensions and Use of the Scholarly Information Environment, the Mellon-funded DLF survey conducted by Outsell Inc. was our massive addition to what we collectively know about users' behaviors in our digital libraries. Equally massive was the task of publishing the results in meaningful ways on the web and (partially) in print. The entire 659-table set arrived as one MS Excel table, requiring lots of work to format it for delivery as separate tables, and we were very fortunate that Amy Friedlander stepped in to provide a framing introduction to this mass of data. The end result is CLIR and DLF's largest print publication ever.

COMMUNICATIONS

Standards: DLF webpage publications policy: we are now following our own recommendations for DLF electronic publications and the web pages. We are using XML, generating multiple output formats using style sheets (PDF, PDA, Web), and providing Dublin Core headers to web publications. The latter can be converted to OAI records in the future.

Website: The new website, with refreshed design, easier maintenance, and richer navigation, will go live in November 2003. It is xml-based and will meet basic Americans with Disabilities Act requirements. Both CLIR and DLF have recently moved their entire website and listserv operations from one commercial host to another, to enrich the web-based services we can offer in the future.

DLF-announce discussion list: this has gone from a surprisingly small list (one or two people per institution) to a group of hundreds, after a 2002 membership drive. It now also includes the steering committee; people engaged on projects; and Forum participants.

DLF Newsletter: this is about to be reactivated, to get us caught up with this and the lists of projects and publications it generates. Michael Pelikan (Penn State) is working with me as a guest editor for this double issue to put us back on track.

2) DLF Publications

The New Publications Process

All publications are issued electronically, from a core XML original. This XML file is created by DLF staff from the material that is submitted for publication (usually MS Word, MS Excel, or PDF documents). Once the XML file is created and parsed we have an easily re-purposeable “blueprint” document that creates the online derivatives quickly, usually through the application of various “re-shaping” stylesheets (web, PDF, ebook). This blueprint file can be made available to authors should they wish to have it.

Looking forward

Printed publications: over the course of 2003-2004, we are exploring the use of the DLF’s XML original in CLIR’s *InDesign* page layout program, for printed publications (rather than starting from an ASCII or MS Word surrogate). We are also working with CLIR to research print-on-demand rather than offset printing of a larger initial print run, to see what cost savings and service benefits there are.

The publications workflow: starting with the *DLF Newsletter*, we are examining the publications workflow to see what we can do to ease the burden of creating a standardized original from the proprietary formats that authors submit. We do not expect authors of DLF publications to have (or need to learn) any XML skills, although in some

cases they have them or are keen to work with us to acquire them. Instead, we are looking to the use of MS Word templates and basic guidance for layout, file-naming, formats submitted, etc.

Publications Issued, 2002-2003

The three items marked with asterisks below are currently available in print as well as in electronic formats. Others in this list are candidates for printed distribution in 2003-2004, especially if we can develop print-on-demand processes.

- The Mellon-Funded Archiving E-Journals series:
 - 1) Cornell University Library. *Project Harvest: Report of the Planning Grant for the Design of a Subject-Based Electronic Journal Repository*.
<http://www.diglib.org/preserve/cornellfinal.html>
 - 2) Harvard University Library. *Report on the Planning Year Grant for the Design of an E-journal Archive*.
<http://www.diglib.org/preserve/harvardfinal.html>
 - 3) MIT University Library. *DEJA: A Year in Review. Report on the Planning Year Grant for the Design of a Dynamic E-journal Archive*.
<http://www.diglib.org/preserve/mitfinal.html>
 - 4) New York Public Library. *Archiving Performing Arts Electronic Resources: A Planning Project*. <http://www.diglib.org/preserve/mitfinal.html>
 - 5) Stanford University Libraries. *LOCKSS: A Distributed Digital Archiving System - Progress Report for the Mellon Electronic Journal Archiving Program*. <http://www.diglib.org/preserve/stanfordfinal.html>
 - 6) University of Pennsylvania Library. *Report On A Mellon-Funded Planning Project For Archiving Scholarly Journals*.
<http://www.diglib.org/preserve/upennfinal.html>
 - 7) Yale University Library. *The Yale Electronic Archive: One Year of Progress: Report on the Digital Preservation Planning Project*.
<http://www.diglib.org/preserve/yalefinal.html>
- *Amy Friedlander. *An Introduction to "Dimensions and Use of the Scholarly Information Environment."* <http://www.diglib.org/pubs/scholinfor/>.
Note: at 893 printed pages, this item is by far the largest print publication issued by DLF or CLIR. It prints 158 of the 659 tables of data that are the results of the Outsell survey. The whole body of findings is available electronically [see publication below].

- *Dimensions and use of the scholarly information environment. A Data Set Assembled by the Digital Library Federation and Outsell, Inc.* All 659 tables published in HTML, PDF, and MS Excel. (November 2002). <http://www.diglib.org/pubs/scholinfor/>
- *Benchmarking digital reproductions of printed monographs and serials.* Version 1.0. <http://www.diglib.org/standards/bmarkfin.htm>
- *Daniel Greenstein and Suzanne E. Thorin. *The Digital Library: A Biography.* (September 2002). <http://www.clir.org/pubs/abstract/pub109abst.html>
- *Ed Jones. *The Exchange of Serials Subscription Information.* A NISO White Paper (co-funded by the DLF). <http://www.niso.org/standards/resources/SerialsWP.html>

3) DLF Contracts awarded and administered

1. **Martha Brogan. Review and Evaluate Digital Library Aggregation Services.** Includes research, travel, and extensive writing time. Delivered product: *Digital Library Aggregation Services: A Report to the Digital Library Federation.* 105pp. Awarded May 2003.
2. **Open Archives Initiative:** the joint DLF/CNI funding for OAI finished early in 2003 and since then some ongoing support has been provided through NSDL activities. OAI goes from strength to strength and Herbert von de Sompel is working on a “static repository” model that will further lower the barrier to publishing OAI records, by allowing an institution simply to put the records in an xml document and send it to a third-party gateway from which it can be harvested. This obviates the need to set up a harvesting space on one’s own server. [www.openarchives.org].
3. **Guidelines for the Cataloging of Cultural Objects:** The DLF joins the Getty Grant Program in sponsoring the Visual Resources Association (VRA) to review and evaluate existing data content standards and current practice in order to compile a manual that may be used to describe, document, and catalog cultural objects and their visual surrogates. The Guidelines have existed in draft since March 2003 – to be completed by Fall 2003. Contact: Diane Zorich, VRA. [http://www.diglib.org/standards/vrawork.htm].

4) Grants Received

Both of these were received after June 20, 2003, but are the result of DLF work and grant-writing done in the 2002-2003 period.

The OCKHAM Library Network: Integrating the NSDL into Traditional Library Services. National Science Foundation (National Science Digital Library)
\$425,000

Awarded to Emory University (lead institution). Principal Investigator, Martin Halbert, Emory University; with Ed Fox (Virginia Tech) and Jeremy Frumkin (University of Arizona). This award is not administered by DLF, but it goes to support the web services work of the DLF's Ockham Group. From the Executive Summary of the grant:

A two-year project to improve usage and access to the NSDL by learning communities through the existing national infrastructure of traditional libraries and their service programs. The NSDL has assembled an extraordinary set of resources and services in support of scientific education. However, there has been limited integration and deployment of NSDL services and content into the existing service programs and infrastructure of the traditional library community. Learning communities would realize many benefits from a coordinated set of networked services for dissemination of NSDL resources through traditional library protocols. This project has 4 goals: 1) to create a reference model for integrating the NSDL into traditional library services; 2) to create a set of networked middleware and test-bed services enabling traditional libraries to expand access to the NSDL through their local service programs; 3) to evaluate of the utility, usage, and impacts of the local library test-bed services on the participating campus communities by means of web log analysis, focus groups, and usability studies; and 4) to disseminate of results and growth of the network, which will initially include the four collaborating institutions of the project. This project will advance the dialog between librarians and researchers through the creation of an extensible framework for networked interoperation between the NSDL and traditional libraries.

A Study of the Interoperation of Learning Management and Library Information Systems. Andrew Mellon Foundation
\$36,000

This award is granted to and administered by the DLF. Project Investigator: Dale Flecker, Harvard University. From the grant proposal:

This study will examine the interaction of digital libraries and learning management (courseware) systems. Learning management systems are increasingly ubiquitous in higher education, and there is an explosion of internet-accessible collections of digital resources relevant to the teachers and students. A natural context for such resources is course web sites. As these sites are increasingly produced through formal learning management systems, the issue of the interactions of such systems with external repositories and discovery systems becomes important. The IMS Digital Repositories Specification provides a good beginning for the understanding of the functions of discovery (search) and object access for

learning management systems. Deliverables from the process will be: 1) a paper describing the interactions between library oriented information repositories and learning management systems; 2) some core use cases of such interactions; 3) a list of functional requirements to allow the interoperation of such systems.

5) Director's Activities

Publications written

David Seaman. "Deep Sharing: The Case for the Federated Digital Library." *EDUCAUSE Review* (July/August, 2003).

David Seaman and Jerry George. "The Digital Library Federation in America: A Portrait." *Zeitschrift fur Bibliothekswesen und Bibliographie*. Vittorio Klostermann, Frankfurt am Main, May 2003.

David Seaman. "Electronic Resource Management: A New Digital Library Federation Initiative." *CLIR Issues 30* (November/December 2002).
<http://www.clir.org/pubs/issues/issues30.html#ermi>

David Seaman. "Update from the DLF." *CLIR Issues 29* (September/October 2002).
<http://www.clir.org/pubs/issues/issues29.html#dlf>

Site Visits to DLF Partners and Allies (The Grand Tour)

CLIR is not listed on here as I am located in their offices; the Library of Congress is listed twice because I visited once while Winston Tabb was still there and then later on to meet Beacher Wiggins.

July 2002

1. Library of Congress (Winston Tabb) 07/26/2002

August 2002

2. Stanford University 08/25/2002
3. Research Libraries Group 08/27/2002
4. University of California, Berkeley 08/28/2002
5. California Digital Library 08/29/2002

September 2002

6. CNI 09/10/2002
7. North Carolina State University 09/13/2002
8. Columbia University: 09/24/2002

DLF Annual Report 2002-2003

9. New York Public Library	09/25/2002
10. New York University	09/25/2002
11. NARA	09/30/2002

October 2002

12. University of Southern California	10/03/2002
13. University of Washington	10/04/2002
14. University of Pennsylvania	10/07/2002
15. Princeton University	10/08/2002
16. Yale University	10/10/2002
17. Carnegie Mellon University	10/11/2002
18. University of Texas, Austin	10/15/2002
19. Harvard University	10/22/2002
20. Cornell University	10/23/2002
21. Penn State University	10/24/2002
22. Dartmouth College	10/28/2002
23. University of Michigan	10/30/2002

November 2002

24. University of Illinois Urbana-Champaign	11/12/2002
25. University of Chicago	11/13/2002
26. Library of Congress	11/18/2002
27. University of Minnesota	11/21/2002

December 2002

28. University of Virginia	12/11/2002
29. OCLC	12/12/2002
30. Indiana University	12/13/2002
31. University of Tennessee	12/16/2002
32. Emory University	12/17/2002
33. Johns Hopkins University	12/19/2002

February 2003

34. Los Alamos National Laboratory	02/19/2003
------------------------------------	------------

Other sites visits to DLF members

- 11/25/2002. Digital humanities pre-planning meeting with Ed Ayers (Dean of Arts and Sciences at the University of Virginia). Charlottesville, VA
- 12/03/2002: Digital Humanities planning meeting with the University of Virginia Provost, University Librarian, and Dean, at the request of the Dean. Charlottesville, VA
- 01/28/2003: New York Public Library: UL think-tank meeting. New York City
- 03/28/2003: CLIR Sponsors Symposium. Washington, DC.
- 03/09/2003-
03/12/2003: University of Washington Library's Mellon-funded Digital Library Planning Retreat. Sleeping Lady, WA
- 04/22/2003: CLIR/CDL's DAI Institute, San Francisco.
- 04/29/2003: CNI Task Force Meeting, Washington DC.

Presentations Given

- 11/14/2002-11/16/2002: *The E-ssential Library: A LAMA Institute*. Naples, Florida:
- a) "Creating & Sustaining Leading Edge Services: Introduction." (with Lorcan Dempsey, OCLC).
 - b) "Creating & Sustaining Leading Edge Services: Advanced I: Aggregation and Malleability: A Behind-the-Scenes Look at How it Works." (with Lorcan Dempsey, OCLC)
 - c) "Creating & Sustaining Leading Edge Services: Advanced II: The library and the classroom/the library and the researcher/The library and the global user." (with Lorcan Dempsey, OCLC)
- 11/19/2002. "Digital Preservation and Archiving: Making sure you can always go back with what is going forward." With Kevin Guthrie, JSTOR. Society for Scholarly Publishing Educational Seminar. American Geophysical Union, Washington DC,. <http://www.diglib.org/presentations/ssp.htm>
- 12/6/2002: "The Digital Library Federation: A Briefing." CNI Task Force Meeting. San Antonio, http://www.diglib.org/presentations/cni-dlf1202_files/frame.htm

DLF Annual Report 2002-2003

01/23/2003 “Repurposing Electronic Materials: New Forms and New Uses.” *Making E-books Work: Librarians and Publishers Exchange Needs and Solutions*. ALA ebooks pre-conference, Philadelphia

02/05/2003 Adding Value to Electronic Publications: What do Librarians and Users Want?” Association of American Publishers, PSP Annual Conference: Smart Content: New Ways To Add Value. Mayflower Hotel, Washington, DC.

02/25/2003. “The Digital Library Federation and Trends in Digital Libraries.” Robert Allen’s Digital Library class at the University of Maryland Library School .

03/10/2003: “Transformations in Librarianship and Scholarship.” University of Washington digital library and scholarship retreat. Sleeping Lady resort, WA.

04/07/2003 “From Isolation to Integration: Reshaping the Serials Data Silos.” Plenary session, with Lynne Brindley and Jay Jordan. *United Kingdom Serials Group Annual Conference*, Herriot-Watt University, Edinburgh.

04/14/2003: With Professor Christopher Mulvey: “The African American Research Library.” *The African Atlantic Conference*, King Alfred’s College, Winchester, England.

05/02/2003. “Digitization, Copyright And Fundraising: How To Make It Work For You.” Maryland Library Directors Meeting, Hood College, MD.

05/13/2003. “Business models for Libraries and Scholars in a Digital Age: The Role of Linking and Aggregation.” NISO-sponsored *OpenURL Day*, New York City.

05/15/2003. “The DLF Today, and the Case for DODL.” DLF Forum, New York City.

05/29/2003. Moderator, Plenary Session. Society for Scholarly Publishers Baltimore.

06/01/2003. “Experiences Using the Open Archives Initiative (OAI).” ACH/ALLC International Conference, Athens, Georgia.

06/05/2003. “What do Librarians and Users Want?” *International Group of Publishing Libraries*. Chaired by Ralph Eubanks. Library of Congress.

Representation of DLF at other meetings

09/09/2002: NSF Middleware digital rights management strategy meeting

09/16/2003-

09/18/2002: European Digital Library Conference, Rome

DLF Annual Report 2002-2003

- 09/18/2002: Rome National library: Internet Archive/national libraries meeting with Michelle Kimpton and the representatives of a range of European national libraries
- 09/23/2002: Southern Area Library Association, Ashville, NC
- 12/18/2002 OKI courseware meeting, Mellon Foundation, NYC,
- 02/20/2003 IMS Digital Library Task Force. Meeting with Neil McLean. Vancouver.
- 02/20/2003: IMS Global's Open Technical Forum. Vancouver.
- 03/17/2003-
03/18/2003: IMLS Workshop, Opportunities for Applied Research on the Creation, Management, Preservation and Use of Digital Content. Washington, DC
- 04/09/2003: Meeting with Richard Boulderstone, Director of e-Strategy, British Library, London.
- 04/10/2003: Meeting with Oxford digital library staff, Oxford.
- 04/11/2003: Meeting with UKOLN metadata group – University of Bath.
- 04/30/2003: NSF/Humanities meeting (Lee Zia/Abby Smith) Washington DC
- 05/10/2003: ACLS Annual Meeting, Philadelphia
- 05/15/2003: OAI Steering Committee, New York City
- 05/30/2003: Alliance for the 21st Century Library planning meeting, Washington DC