



DLF Quarterly Report

January-March 2006

David Seaman

These reports on DLF activities are designed to supplement the website sections on individual initiatives, reports at Board meetings, the Annual Report, and the regular postings of news on the DLF-ANNOUNCE listserv. In this report I am offering a short summary of ongoing work from a range of DLF initiatives, instead of concentrating on one only.

1) DLF Aquifer

See separate update circulated at Spring Board of Trustees Meeting.

2) DLF Services Framework

See separate update circulated at Spring Board of Trustees Meeting.

3) Service Registries/Ockham

How do you discover and use digital library services that go with a particular collection or collections? How does machine-to-machine interchange about such services take place? There is a growing awareness that the answer is one or more registries of services that can be queried – work that DLF's Ockham group has been doing with NSF funding for the past several years to improve the deployability of the National Science Digital Library (NSDL) in traditional libraries <<http://www.ockham.org/>>

On March 22nd and 23rd, 2006, the Ockham project hosted an international workshop on Digital Library Service Registries (DLSRs). Funded by the National Science Foundation's National Science Digital Library Program, and co-sponsored by the Digital Library Federation, JISC, and UKOLN, this workshop brought together twenty-three experts in digital library development and registry implementations from the US and Europe. The workshop focused on the concept of DLSRs, the current state of various registry projects in the U.S. and the U.K., the problems that service registries might solve

for digital libraries, the issues in codifying standards for digital library service registries, and the issues involved in using DLSRs.

The workshop was the initial step in a series of activities targeted at developing a sustained service registry for digital libraries across all digital library communities.

Information on the workshop, including presentations, a list of attendees, and a draft report can be found at <http://wiki.library.oregonstate.edu/confluence/x/xhY>

Repository Interoperability: Following up on this workshop, the Ockham project has applied for extended funding from NSF to continue work on the Ockham Digital Library Service Registry. In addition, Jeremy Frumkin gave a presentation on service registries and other needed infrastructure for a workshop on repository interoperability in NYC on April 20. This workshop, sponsored by Microsoft, the Mellon Foundation, CNI, DLF, and JISC, also brought together an international set of digital library experts to explore future directions of interoperability in relation to repositories. More information on this meeting can be found at <http://msc.mellon.org/Meetings/Interop>.

4) Registry of Digital Masters

There are several initiatives underway in support of the growth and development of The Registry of Digital Masters, a service hosted by OCLC to record books and other materials digitized as page images to preservation standards.

1. **Growing the registry:** The DLF working group continues to have regular meetings to discuss everything from registration projects currently underway to modifications necessary to the record guidelines that are required for new circumstances. DLF working group participants have added over 20,000 records to the registry through their direct efforts.
2. **Free access:** OCLC is in the final stages of putting a public face on the Registry for free searching. When finished in the next few months the Registry will be publicly accessible from the DLF web site.
3. **International exchange of records:** OCLC has engaged LIBER (the European Research Library Organization) to create a synchronized U.S. / European registry based on record exchanges between OCLC and EROMM, the European Registry of Digital Masters. On May 4-5, 2006, there will be a first meeting between LIBER & DLF libraries to harmonize registry practices and create an international best practice for registration.
4. **Google:** OCLC is engaged with both the Google digitization project, Open Content Alliance, and other mass digitization efforts exploring ways to efficiently record the output of these programs within the Registry.

5) Metadata Encoding and Transmission Standard (METS)

<http://www.loc.gov/standards/mets/>

The METS schema is a standard for encoding descriptive, administrative, and structural metadata regarding objects within a digital library, and has developed out of the DLF community. Use of METS is growing and it is being deployed and tested in various repository environments by a number of institutions who have received various grants from NDIIPP, IMLS, and NSF. This year has seen a re-grouping on and re-articulation of METS governance, strategy, activities, and promotion.

Re-defining the METS Board

They have solidified change in the structure of Board leadership to include Co-chairs for Administration and for Technical Issues <http://www.loc.gov/standards/mets/mets-board.html> .

Strategic Direction

They have a sharpened focus on several strategic directions to pursue as a Board:

- Building the METS schema into a viable, standard for metadata encoding and for transmission of content and metadata among our key constituents, i.e., creators and consumers of “library” or “scholarly” digital content
- Building and further developing a METS community

Key Activities in pursuit of strategic aims

- Develop means to discuss, develop implementation part of the METS community by establishing a periodic METS Implementation event at least once a year starting in June, 2006
- Complete technical documentation for METS implementers and those contemplating the use of METS
- Investigate, find funding and developer support for building METS tools as desired by the METS implementers
- Open the METS Board meetings to the public
- Identify small projects that test and prove the interoperability potential of METS / METS profiles.

Promotion

Articulate METS’ strategic directions by creating and approving a Mission Statement, a Position Statement of METS vis a vis other information standards, a Statement of Expectations for Board members, and METS Technical Documentation (the latter is in the final stages of completion – expected release date by METS Implementation Meeting in mid June of this year).

6) Text Encoding Initiative use in Libraries

The TEI in Libraries working group issued its original guidelines in 1999, providing a first draft of recommended common practices for library digital production in this use of this common full-text markup scheme. Both TEI and the use of it in libraries have developed since then and a new working group has been active over the last year with a mission to accomplish the following goals:

- 1) Update the current TEI in Libraries Guidelines (<http://www.diglib.org/standards/tei.htm>) to reflect the changes occurring within the text encoding world generally and within the TEI community specifically.
- 2) Further illuminate the different levels of encoding by offering clearer and more robust examples of implementation.

They have been successful in achieving both of these goals. After producing a substantially revised version of the Guidelines during the winter of 2005, the working group met in February 2006 to further edit and refine them. The current TEI in Libraries document (version 2.1) is the product of this meeting: <http://www.diglib.org/standards/tei.htm>.

The second major deliverable, now in the planning stage, is to produce a TEI in Libraries vendor schema in order to facilitate consistent document creation across institutions and repositories. This will be a distillation of the instructions and DTDs we now give separately to vendors who produce our encoded text; it will reflect the efficiencies and lessons learned from those of us doing this work, and will aid the interoperability of the files produced at various locations.

7) Photographic Preservation

Working with imaging engineers from Kodak, specialists from NARA, and with the large-scale production services at the Library of Congress, we are looking both at a definitive set of guidelines for preservation imaging of photographic materials, building on the *Technical Guidelines for Digitizing Archival Materials for Electronic Access: Creation of Production Master Files-Raster Images* NARA/DLF publication of last year <<http://www.diglib.org/pubs/dlf103/pugtech.html>> and on the creation of new, machine readable imaging targets.

These image-level and device-level targets and software that are being developed will allow institutions to verify the quality of the imaging being performed by contractors or in-house. Most scanning does not actually deliver the quality that is expected. Current standards are overly complex, not descriptive, and no comprehensive tools exist to test according to the standards. One other major advantage of the approach that is being developed is that it is intended to provide quality control in an automated process. As scanning projects get larger, it is becoming difficult to perform quality control on even a

small portion. The tools and process being developed will provide the means to certify the equipment and check quantitative characteristics on every image.

This single target will enable one to monitor and measure:

- 1) Visual and measured resolution (i.e., MTF or SFR) and color mis-registration
- 2) Color and neutral response characteristics,
- 3) One-dimensional illumination uniformity, and
- 4) Sampling rate and object size. (English and metric)



Each color and neutral patch is also documented with human and machine readable text (5) on the target itself with their L*a*b* values so future users of the digital image file can accurately reproduce the analytical colors and tones of the original object without having to search for the meaning of any particular patch. While not intended to be part of the finished delivered image or file, the target should be part of a repository's archived image file. Of course, input device color profiles can also be built from this information.

Director's Activities: presentations, site visits, and DLF meetings

1/5-1/6/2006: DLF Services Framework Working Group meeting to launch and plan Geneva Henry's year as Distinguished Fellow. Washington DC.

1/11/2006: Meet with Pat Stevens to discuss NISO/DLF joint work, especially on e-resource management.

1/12/2006: "Electronic Books, Digital Readers, and the Future of America's Libraries." Boston Athenaeum, Boston, MA.

1/16/2006: site visit to Cornell University, Ithaca, NY.

1/25-1/27/2006: Meet with JISC e-Framework and Information Environment teams, with Geneva Henry. London, UK.

1/26/2006: site visit to the British Library to introduce Geneva Henry and to discuss the DLF Services Framework.

2/8/2006: Presentation, PSP Round Table, as part of the AAP'S Professional and Scholarly Publishing Division's annual meeting. Washington DC.

2/9-2/11/2006: Aquifer Working Groups meeting, Washington DC.

2/13-2/14/2006: TEI in Libraries Working Group meeting, Washington DC

2/16-2/18/2006: Web-Wise Conference, Los Angeles, California.

2/22/2006: site visit to University of Tennessee, Knoxville, with Katherine Kott.

3/6-3/11/2006: Rare Book School, University of Virginia.

3/16/2006: JISC Annual Meeting, Birmingham, UK.

3/23-3/24/2006: Joint DLF/NSF Services registry meeting, Ballston, VA.

3/24/2006: Bibliographical Society of Virginia Spring Council Meeting. Harrison Institute/Small Special Collections Library, University of Virginia.

3/30/2006: Photographic Preservation working group meeting, with Don Williams (Kodak) and Michael Stelmach (LOC).