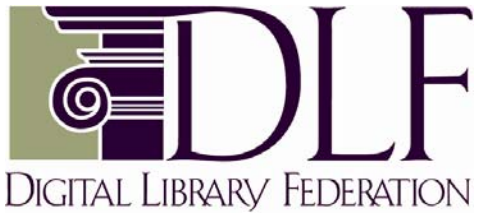


FALL FORUM 2005

OMNI CHARLOTTESVILLE HOTEL

CHARLOTTESVILLE, VA

NOVEMBER 7–9, 2005

Washington, D.C.

Printed by Balmar, Inc.

Digital Library Federation
Council on Library and Information Resources
1755 Massachusetts Avenue, NW, Suite 500
Washington, DC 20036
http://www.diglib.org

CONTENTS

ACKNOWLEDGMENTS

DLF Forum Fellowships for Librarians New To the Profession

The Digital Library Federation would like to congratulate the following for winning DLF Forum Fellowships:

- Timothy Donohue, Research Programmer, *University of Illinois at Urbana-Champaign*
- Melanie Feltner-Reichert, Digital Coordinator, *University of Tennessee*
- Tamara Lopez, "The Chymistry of Isaac Newton" Project, *Indiana University*
- Youn Noh, Digital Resources Catalog Librarian, *Yale University*
- Matthew Treskon, Library Assistant III, *The Johns Hopkins University*

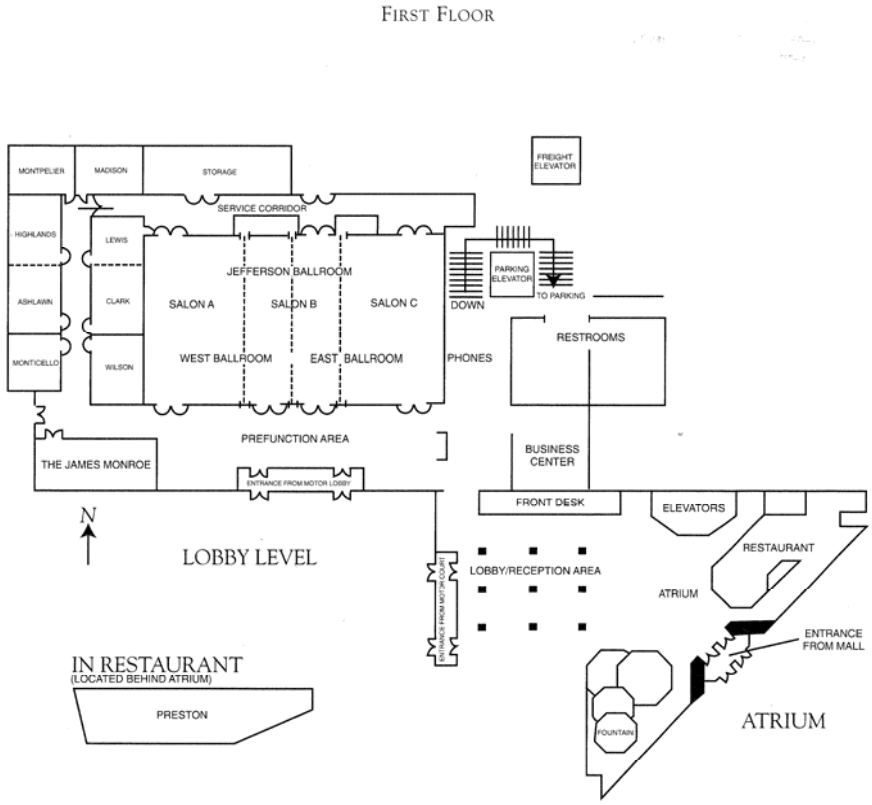
DLF Program Committee

The DLF would also like to thank the DLF Fall Forum 2005 Program Committee and Fellowship Selection Committee for all their hard work:

- Denise Troll Covey, *Carnegie Mellon University*
- Ann Lally, *University of Washington*
- David Seaman, *Digital Library Federation*
- Sarah Shreeves, *University of Illinois at Urbana-Champaign*
- Erin Stalberg, *University of Virginia*
- Jewel Ward, *University of Southern California*
- Perry Willett, *University of Michigan*

SITE MAP

Fig. 1. Site Map of the Omni Charlottesville Hotel



SCHEDULE AT A GLANCE

PRECONFERENCE: SUNDAY, NOVEMBER 6, 2005

9:00AM–4:00PM ***American West Partners Meeting****—for project participants* (Salon C, Lobby Level)

PRECONFERENCE: MONDAY, NOVEMBER 7, 2005

8:30AM–12:30PM ***DLF Aquifer Meeting****— for project participants* (James Monroe, Lobby Level)

9:00AM–11:30PM ***NDIIPP Technical Architecture Affinity Group Meeting****— for project participants* (Lewis/Clark, Lobby Level)

DAY ONE: MONDAY, NOVEMBER 7, 2005

10:30AM–1:00PM ***Registration***
(Prefunction Area, Lobby Level)

11:30AM–12:15PM ***First-time Attendee Orientation***
(Salon A, Lobby Level)

12:45PM–1:00PM ***Opening Remarks by David Seaman and Barrie Howard*** (Salons A and B, Lobby Level)

1:00PM–2:30PM ***Session 1: The National Digital Information Infrastructure and Preservation Program***
(Salon A, Lobby Level)

1:00PM–2:30PM ***Session 2: Metadata Strategies*** (Salon B, Lobby Level)

2:30PM–3:00PM ***Break*** (Prefunction Area, Lobby Level)

3:00PM–4:30PM ***Session 3: Digital Preservation*** (Salon A, Lobby Level)

3:00PM–4:30PM ***Session 4: Managing Digital Library Content and Code*** (Salon B, Lobby Level)

4:30PM–4:45PM ***Break*** (Prefunction Area, Lobby Level)

4:45PM–6:00PM ***Birds of a Feather (BOF) Session 1:***

1. *The DLF Electronic Resource Management Initiative: Phase 2* (Salon A, Lobby Level)

2. *OCKHAM: Digital Library Service Registries* (Salon B, Lobby Level)

3. *Archival Information Control* (Ashlawn/Highlands, Lobby Level)

4. *Preservation Digitization for Photographs* (Lewis/Clark, Lobby Level)

7:00PM–9:30PM ***Reception*** (Harrison Institute, University of Virginia)

Note: Round-trip transportation will be provided from the Omni hotel to the reception site on the University of Virginia campus.

DAY TWO: TUESDAY, NOVEMBER 8, 2005

8:00AM–9:00AM ***Breakfast*** (Atrium, Lobby Level)

9:00AM–10:30AM ***Session 5: Remodeling Digital Library Systems*** (Salon A, Lobby Level)

9:00AM–10:30AM ***Session 6: DLF Aquifer*** (Salon B, Lobby Level)

10:30AM–11:00AM ***Break*** (Prefunction Area, Lobby Level)

11:00AM–12:30PM ***Session 7: Dynamic Digital Environments*** (Salon A, Lobby Level)

11:00AM–12:30PM ***Session 8: Collaborative Metadata Aggregation*** (Salon B, Lobby Level)

12:30PM–2:30PM ***Break for Lunch*** (Individual Choice)

Note: There is an open-air, pedestrian mall just outside the Omni hotel where there are many restaurants for lunch choices.

2:30PM–4:00PM ***Session 9: NARA's Electronic Records Archives*** (Salon A, Lobby Level)

2:30PM–4:00PM ***Session 10: OAI for Digital Library Aggregation*** (Salon B, Lobby Level)

4:00PM–4:30PM ***Break*** (Prefunction Area, Lobby Level)

4:30PM–6:00PM ***Birds of a Feather (BOF) Session 2:***

5. *Open Archives Initiative Protocol for Metadata Harvesting: Best Practices for Data Provider Implementation and Shareable Metadata* (James Monroe, Lobby Level)

6. *Preservation Planning for Digital Objects and Repositories* (Ashlawn/Highlands, Lobby Level)

7. *METS Implementation and Profile Building* (Lewis/Clark, Lobby Level)

DAY THREE: WEDNESDAY, NOVEMBER 9, 2005

8:00AM–9:00AM ***Breakfast*** (Atrium, Lobby Level)

9:00AM–10:30AM ***Session 11: Digital Library Initiatives in South Africa and England*** (Salon A, Lobby Level)

9:00AM–10:30AM ***Session 12: Web Archiving Services*** (Salon B, Lobby Level)

10:30AM–11:00AM ***Break*** (Prefunction Area, Lobby Level)

11:00AM–12:30PM ***Session 13: Sustaining Digital Scholarship*** (Salon A, Lobby Level)

11:00AM–12:30PM ***Session 14: Archive-it: A Web Archiving Application*** (Salon B, Lobby Level)

12:30PM ***Adjourn***

POST-CONFERENCE: WEDNESDAY, NOVEMBER 9, 2005

12:30PM–4:30PM ***OAI Vendors' Panel—*** *for project participants* (Ashlawn, Lobby Level)

1:00PM–5:00PM ***METS Editorial Board Meeting*** (Monticello, Lobby Level)

1:00PM–5:00PM ***DLF Developers' Forum Meeting—*** *for project participants* (James Monroe, Lobby Level)

1:30PM–5:30PM ***CCS docWORKS Workshop—*** *for project participants* (Computer Classroom, Fourth Floor, Alderman Library, University of Virginia)

POST-CONFERENCE: THURSDAY, NOVEMBER 10, 2005

9:00AM–5:00PM ***METS Editorial Board Meeting*** (James Monroe, Lobby Level)

9:00AM–5:00PM ***DLF OAI Implementers' Workshop—*** *for project participants* (Computer Classroom, Fourth Floor, Alderman Library, University of Virginia)

PROGRAM WITH ABSTRACTS

PRECONFERENCE: SUNDAY, NOVEMBER 6, 2005

9:00AM–4:00PM ***American West Partners Meeting****—for project participants* (Salon C, Lobby Level)

PRECONFERENCE: MONDAY, NOVEMBER 7, 2005

8:00AM–12:30PM ***DLF Aquifer Meeting****— for project participants* (James Monroe, Lobby Level)

9:00AM–11:30PM ***NDIIPP Technical Architecture Affinity Group Meeting****— for project participants* (Lewis/Clark, Lobby Level)

DAY ONE: MONDAY, NOVEMBER 7, 2005

10:30AM–1:00PM ***Registration***
(Prefunction Area, Lobby Level)

11:30AM–12:15PM ***First-time Attendee Orientation***
(Salon A, Lobby Level)

12:45PM–1:00PM ***Opening Remarks*** (Salons A and B, Lobby Level)

"Welcome to the DLF Fall Forum 2005."

David Seaman, and Barrie Howard, Digital Library Federation

1:00PM–2:30PM ***Session 1: The National Digital Information Infrastructure and Preservation Program***
(Salon A, Lobby Level)

"Maintaining Archive Integrity During Inter-repository Transfer: Lessons Learned from the NDIIPP Archive Ingest and Handling Test."

Session 1 (continued)

- Martha Anderson, Project Manager, Library of Congress, Moderator
- Clay Shirky, NDIIPP Technical Lead, NYU
- Michael Nelson, Old Dominion University
- Tim DiLauro, The Johns Hopkins University
- Keith Johnson, Stanford University
- Stephen Abrams, Harvard University

The Archive Ingest and Handling Test (AIHT), a practical experience with the proposed National Digital Information Infrastructure and Preservation Program (NDIIPP) architecture, completed its work in early summer 2005. Project teams from Harvard, Johns Hopkins, Old Dominion, and Stanford investigated diverse technologies and approaches ranging from the examination of new models of preservation with self-archiving object technologies to testing content repository technologies as platforms for preservation. Risk assessment, file evaluation, and validation tools were developed and used. A final phase of the project simulated change over time by testing file migration to different formats.

Surprising challenges to basic assumptions about file and whole archive transfer were brought to light. The tested approaches and lessons learned about the transfer, assessment, and management of a digital archive will be discussed by principal investigators from participating institutions and the Library of Congress. The University of Michigan digital library system was developed in the 1990s to process SGML files using Perl scripts.

This presentation will describe the transformation of our digital library system for XML/XSL and Unicode, and the lessons learned. Panelists will describe the original system, discuss the reasons for this major undertaking, and will cover topics such as the planning process, how production systems were maintained during the re-architecture, large-scale data conversion from SGML/ISO-Latin1 to XML/Unicode files, tools for conversion, processing, version control and debugging, and what they've learned in the process.

1:00PM–2:30PM ***Session 2: Metadata Strategies*** (Salon B, Lobby Level)

"A Taxonomic Approach to the Organization of Penn State Web Space."

Michael Pelikan, Pennsylvania State University

Session 2 (continued)

Penn State University has had broad experience with several search engines, and currently holds a university-wide license for the Google Search Appliance. While useful for general-purpose Web searches, the Google Search Appliance does not currently address critical search and retrieval issues in the Penn State Web environment.

The Taxonomic Tags group was formed as a university-level group with representatives from Information Technology Services, Finance & Business Administration, and the Penn State Libraries, to examine whether a taxonomic model, expressed as metadata tags and systematically applied across the university's Web pages could:

- permit specific pages to be the top hit for specific searches,
- make it easier to find specific pages from among the university's more than 1,000,000 public Web pages,
- remain useful amidst increasing adoption of content management systems across Penn State,
- remain useful over time as search engines continue to evolve, despite whether open source or commercial (and often proprietary) search algorithms are employed.

The Tags group has developed recommendations to address these issues. These recommendations include the development of a controlled vocabulary, along with synonyms, for Penn State departments, colleges, administrative units, etc. These terms would be incorporated both into Web pages, and into the university's LDAP system.

The group has recommended that a broadcast search mechanism be developed for the main Penn State Web search screen. Under this system, user search terms will be submitted both to a Web search appliance and to the university's LDAP system. The results will be combined, identified and presented to the user.

Members of the Tags group will present background on the project and update its progress to date. The Tags Group is highly interested in questions and comment, and will tailor the presentation to permit as much time for discussion as possible.

"Unpacking the Interpretation of METS Markup."

David Dubin, University of Illinois at Urbana-Champaign

Session 2 (continued)

Like most XML applications, METS, the Metadata Encoding and Transmission Standard, overloads a small number of generic syntactic relationships (e.g., parent/child) to represent a variety of specific semantic relationships. Human beings correctly infer the meaning of METS markup, and these understandings inform the logic and design of applications that import, export, and transform METS-encoded resources and descriptions.

However, METS's flexibility and generality invite diverse interpretations, posing challenges for processing across different METS profiles and local adaptations. Robust processing requires support in the form of a general software library for reasoning about METS documents. We describe the current state of development for such a library.

This METS interpretation software is an application of the BECHAMEL markup semantics framework (Dubin, et al., 2003). BECHAMEL applications translate properties and relationships expressed in conventional markup into logical assertions that unpack the overloaded XML-based syntax. The inference problems we aim to support include identifying inline and external storage objects, mapping storage objects to resources and descriptions, and correctly classifying the role of namespaces.

Another goal of explicating the interpretation of METS documents is to reserialize them in XML, directly asserting as many of the inferred facts as we can. In this way we hope to improve prospects for long term digital preservation.

Dubin, D., C. M. Sperberg-McQueen, A. Renear, and C. Huitfeldt. “A Logic Programming Environment for Document Semantics and Inference.” *Literary and Linguistic Computing* 18 (2003): 225–233.

“The Problem with Duplicates.”

Esme Cowles, University of California, San Diego

When harvesting large numbers of non-unique metadata records from several different institutions, duplicate records are inevitable. Typically, duplicates are identified by comparing globally-unique identifiers (such as ISBNs) or definitive metadata elements (such as creator and title statements). However, some disciplines (such as art) have neither unique identifiers nor definitive metadata, making the task of identifying duplicate records much harder.

Session 2 (continued)

The Union Catalog for Art Images (UCAI) project aggregated 920,000 metadata records for slides and digital images from six institutions, mapped all records to a common schema (VRA Core), and attempted to identify and merge duplicate records. Drawing on clustering techniques, controlled vocabularies, and string-comparison algorithms, the UCAI team developed tools and software to compare metadata records and identify duplicate records.

The experience of the UCAI team demonstrates the challenges of working with metadata created without common content standards, controlled vocabularies, or unique identifiers, and provides guidance for content producers and aggregators.

2:30PM–3:00PM ***Break*** (Prefunction Area, Lobby Level)

3:00PM–4:30PM ***Session 3: Digital Preservation***
(Salon A, Lobby Level)

"Preserving Digital Resources: Complexities and Emerging Solutions (A View from the NDIIPP Partners Early Work)."

- Joanne Kaczmerek, University of Illinois at Urbana-Champaign
- Patricia Cruse, California Digital Library
- Martin Halbert, Emory University
- Anthony Ramirez, University of Maryland
- Bill LeFurgy, Library of Congress
- Jim Tuttle, North Carolina State University
- Nan Rubin, WNET

The general purpose of this discussion panel is to present the challenges of digital preservation as experienced to date by institutions engaged in digital preservation projects through the NDIIPP initiative, and, specifically, to examine and discuss examples of solutions currently implemented or being considered by NDIIPP partners.

Discussion will open with remarks from the Library of Congress on specific challenges presented by the "digital preservation problem." After a brief introduction of NDIIPP partners, representatives from each project will provide concise examples of currently implemented or pondered solutions to the specific challenges they are encountering.

Session 3 (continued)

Challenges arising include varied technical issues, as well as issues related to workflow, rights clearance, and the economic sustainability of preservation activities. This discussion will also include consideration of the role the NDIIPP partnership model might play in developing preservation strategies and solutions. The session format is intended to foster discussion among panelists, and audience participation is encouraged.

Libraries and archives have traditionally played the role of "trusted repositories," assuming long-term responsibilities for assuring the integrity and authenticity of materials deposited with or collected by them. With the proliferation of digital resources, the role of a "trusted repository" takes on a new aspect, requiring libraries, archives, and other institutions to re-conceptualize their place in providing assurances of long-term digital preservation.

Recognizing the need for a coordinated approach to preserving digital resources, the Library of Congress launched a $99.8M national digital strategy effort through the National Digital Infrastructure Information Preservation Program (NDIIPP). Its mission is to "develop a national strategy to collect, archive and preserve the burgeoning amounts of digital content, especially materials that are created only in digital formats, for current and future generations." In Fall 2004, the eight projects participating in this proposed panel were awarded funding. More information about NDIIPP and its partners may be found at http://www.digitalpreservation.gov.

3:00PM–4:30PM ***Session 4: Managing Digital Library Content and Code*** (Salon B, Lobby Level)

"Flipping the Switch: Lessons Learned from a Major Digital Library Migration Project."

Jon Dunn, Mark Notess, and Ryan Scherle
(all Indiana University)

In 2005, the Variations2 digital music library transitioned from being only a research project to becoming the replacement for the heavily-used Variations online music service at Indiana University. The effort to bring this second-generation digital library into production included re-processing and checking approximately 10,000 digitized sound recordings, the creation of a new digital ingest tool, and the development of an access control mechanism to ensure appropriate copyright safeguards.

Session 4 (continued)

Moving nine years' worth of digitized recordings proved to be more than a simple matter of pointing the new tool at the old files. We had to retrieve, and in some cases locate, the original .wav files and re-encode them to support the superior capabilities of the new tool. We also moved the production file server from a tape-based system to hard disks. Re-encoding ran 24x7 for approximately two months. Subsequent error checking and clean-up took several months more.

The transition provided an opportunity to reassess and improve our audio ingest process. The new digitization tools were designed in consultation with the digitizing staff and fit much better with the digitization workflow, increasing throughput and improving quality.

The Variations2 access control mechanism limits out-of-library use of copyrighted materials according to a new access policy, based on student course enrollment. With this access mechanism in place, we are distributing the Variations2 client software to students to support home access of streaming audio and scanned score images.

This talk describes the lessons learned, and the surprises along the way, during the Variations to Variations2 migration.

"Organizing Project Code for Digital Library Applications."

Eric Stedfeld, New York University

Many digital library projects suffer a circuitous evolution, starting as a demo or proof of concept in a scripting language, then adding on requirements and expectations until the project reaches production status.

This can result in disorganized code that is difficult to maintain, update, extend, or scale, let alone transition into another coding environment such as Java. Even more challenging, programmers familiar with the original scripting language may have little background in the principles and methods of the new environment that lead to good programming practice.

This presentation provides some approaches for better structuring and maintaining such code, based on Java Guidelines and Patterns. The example application, a digitized collection of Colonial and Early American documents, utilizes servlets, JSP, JavaBeans, a database backend, and XML files generated with the METS Java Toolkit. The principles presented can help session participants make their application code more manageable, extensible and scalable, while saving time and reducing frustration in software development.

4.30PM–6.00PM ***Birds of a Feather (BOF) Session1:***

1. *The DLF Electronic Resource Management Initiative: Phase 2* (Salon A, Lobby Level)

   Adam Chandler, Cornell University

   This presentation will describe the scope of Phase 2 of the DLF Electronic Resource Management Initiative ("ERMI 2"), including timeline, objectives, and deliverables.

2. *OCKHAM: Digital Library Service Registries* (Salon B, Lobby Level)

   Jeremy Frumkin, Oregon State University
   Martin Halbert, Emory University

   The OCKHAM Project will hold a BOF session to explore Digital Library Service Registries (DLSRs) at the DLF Fall Forum. With the gaining popularity of metasearch tools, OAI-PMH available collections, the use of OpenURL resolvers, and the emergence of new efforts such as COinS, there is a growing need for registries to support access to these services. This BOF will focus on the concept of the DLSR, what functions a DLSR supports, and will examine current DLSR efforts, including the OCLC OpenURL registry, the JISC/IESR DLSR, and the Ockham Distributed DLSR.

   In addition, we will discuss how DLSRs might play a key role in enabling new digital library functionality. Combined with the concept of "Autodiscovery" (techniques for automatically finding machine-processable resources associated with a particular Web page), can we utilize DLSRs to lower the barriers to information integration while at the same time enabling greater and more complex information workflows? Can we create a "digital library dialtone" which makes connecting digital library services and content as easy as placing a phone call? Come to this BOF to find out!

3. *Archival Information Control* (Ashlawn/Highlands, Lobby Level)

   Stephen Davis, Columbia University
   Ellie Brown and Karen Calhoun, Cornell University

BOF Session 1 (continued)

A discussion about how to get control institutionally over finding aid creation and management, as well as the full lifecycle of archival collection information. Lee Mandell will also report briefly on the status of the Mellon-funded Archivists Toolkit project.

4. *Preservation Digitization for Photographs* (Lewis/Clark, Lobby Level)

   Erin Rhodes, U.S. National Archives and Records Administration
   David Seaman, Digital Library Federation

Last year, NARA produced a very well received electronic publication, *Technical Guidelines for Digitizing Archival Materials for Electronic Access: Creation of Production Master Files - Raster Images* (Steven Puglia, Jeffrey A. Reed, and Erin Rhodes, the U.S. National Archives and Records Administration, June 2004), subsequently issued as a print document by DLF (see http://www.diglib.org/pubs/dlfpubs.htm #nara-raster).

This report addresses a spectrum of considerations for digitizing a variety of textual and photographic records, including file formats, image capture, metadata, and quality assessment. In April 2005, a team of experts from across DLF and beyond, including Harvard, NARA, LC, Kodak, Chicago Albumen Works, RLG, RIT, and the Swiss Federal Institute of Technology, convened to build on this work to produce a guide for the digitization of photographs for preservation reformatting. This session allows you to hear about our plans, and have input into the process. You will also be able to preview a proposed new quality control target designed for image digitizing workflows in libraries and museums.

7:00PM–9:30PM ***Reception*** (Harrison Institute, University of Virginia)

Note: Round-trip transportation will be provided from the Omni hotel to the reception site on the University of Virginia campus.

DAY TWO: TUESDAY, NOVEMBER 8, 2005

8:00AM–9:00AM ***Breakfast*** (Atrium, Lobby Level)

9:00AM–10:30AM ***Session 5: Remodeling Digital Library Systems*** (Salon A, Lobby Level)

"Re-architecting a Digital Library System for XML/XSL and Unicode: Lessons Learned."

Phil Farber, Alan Pagliere, Chris Powell, John Weise, and Perry Willett (all University of Michigan)

The University of Michigan digital library system was developed in the 1990s to process SGML files using Perl scripts. At Michigan, this system provides access to over 20,000 texts, 250,000 images, and 850 archival finding aids. In addition, this system is used at 28 other institutions.

This presentation will describe the transformation of our digital library system for XML/XSL and Unicode, and the lessons learned. Panelists will describe the original system, discuss the reasons for this major undertaking, and will cover topics such as the planning process, how production systems were maintained during the re-architecture, large-scale data conversion from SGML/ISO-Latin1 to XML/Unicode files, tools for conversion, processing, version control and debugging, testing, and what they've learned in the process.

9:00AM–10:30AM ***Session 6: DLF Aquifer*** (Salon B, Lobby Level)

- Katherine Kott, DLF Aquifer
- Leslie Johnston, University of Virginia
- Sarah Shreeves, University of Illinois at Urbana Champaign
- Jon Dunn, Indiana University
- Martin Halbert, Emory University

DLF Aquifer, the Digital Library Federation distributed open digital library is in implementation. Since the DLF Spring Forum, project working groups have created a collection development policy, a DLF Aquifer MODS metadata profile to support service development and have selected a small set of digital collections to use as an initial test-bed. University of Michigan will begin harvesting metadata for the project soon.

The services working group has identified target audiences, developed use cases and surveyed DLF libraries to learn what is already known about digital collection use. Taking their cues from the services working group, the technology/architecture working group completed a draft of architectural principles and selected the "repository neutral" framework designed at Johns Hopkins as the DLF Aquifer "middleware layer".

Session 6 (continued)

This panel will review the accomplishments of the past six months and outline the phase I deliverables that will be demonstrated at the DLF Spring Forum in Austin next April.

10:30AM–11:00AM ***Break*** (Prefunction Area, Lobby Level)

11:00AM–12:30PM ***Session 7: Dynamic Digital Environments***
(Salon A, Lobby Level)

"A Format-registry Based Automated Workflow for the Ingest and Preservation of Electronic Journals."

Evan Owens, Chief Technology Officer, Portico

Portico (http://www.portico.org), with funding from The Andrew W. Mellon Foundation, Ithaka, and JSTOR, has developed an automated workflow for the ingest of publisher-supplied e-journal source files into a preservation repository. Electronic journals as a preservation challenge sit somewhere between traditional digitization projects and Web-harvesting projects in that the formats are known and controlled but by the content provider rather than by the archive. The workflow that Portico has developed builds on concepts that were developed in the preliminary work towards a Global Digital Format Registry (GDFR) and on the JHOVE tool set. The components of the workflow include package disassembly, format identification and verification, structure mapping, automated metadata harvesting, rule-based format normalization, and support for quality control and inspection. The system implementation uses a service-based architecture built upon a format registry and a tool registry with support for distributed and pluggable tools. This presentation will review the workflow and system design and discuss our experience in designing and building a system based on a format registry.

"WikiD—Applying Wiki Principles to Structured Data."

Jeffrey A. Young, OCLC Online Computer Library Center

Ward Cunningham describes a wiki as "the simplest online database that could possibly work". The cost of this simplicity is that wikis are generally limited to a single collection containing a single kind of record (viz. WikiMarkupLanguage records). WikiD extends the Wiki model to support multiple collections containing arbitrary schemas of XML records with minimal additional complexity.

Session 7 (continued)

WikiD is essentially a lightweight framework combining:

- Open-source implementations of various loosely-coupled open-standard protocols (e.g. OpenURL, SRW/U, SRW Update, OAI-PMH, RSS)

- An open-source version-controlled database.

- A set of bootstrap collections:

    - CollectionCollection — the master collection of all collections defined in WikiD
    - CollectionExternalSchemas — a registry of XML Schemas used to constrain the items in WikiD collections
    - CollectionWikiPages — the default collection that not only provides WikiD's conventional out-of-the-box wiki functionality but also acts as the user interface for the creation and maintenance of other collections.

- XSL Stylesheets to render collection-level open-standard protocol responses into HTML for human consumption. Automated processes can ignore the stylesheet reference and use the open-standard protocol responses directly.

Possible applications for WikiD include collaborative maintenance of registries, thesauri, taxonomies, reviews, and documentation. In addition to a standard set of features available for all collections, custom code (e.g. Java or XSL) can also be assigned to provide new types of Web services related to individual collections.

The WikiD project page can be found at http://www.oclc.org/research /projects /wikid/default.htm. A demo is running at http://alcme.oclc.org /wikid/. Instructions for creating a new collection can be found at http://alcme.oclc.org/wikid/DemoInstructions. A J2EE Web app distribution is in the works.

11:00AM–12:30PM ***Session 8: Collaborative Metadata Aggregations***
(Salon B, Lobby Level)

"Collaborative Metadata Aggregations: The Road to Shareable Metadata."

Session 8 (continued)

- Sarah Shreeves, University of Illinois at Urbana Champaign, Moderator
- Bill Landis, California Digital Library
- Trish Rose, University of California, San Diego
- Timothy Cole, University of Illinois at Urbana Champaign
- Jenn Riley, Indiana University

DLF emphasizes the role of collaboration to better understand how best to share our digital content and metadata. To this end it has supported the writing of the DLF-NSDL Best Practices for OAI Data Provider Implementations and Shareable Metadata in order to cope with the most common difficulties in the exchange of metadata between content providers and aggregators. This best practices work has drawn on the experiences of collaborative projects involving metadata sharing and has highlighted the importance of such projects to facilitate the dialogue between content providers and aggregators.

This panel will briefly highlight the experiences of several OAI based and non-OAI based collaborative aggregations and best practices building initiatives and will then turn to an open discussion of the issues facing these collaborative projects and initiatives and how they help foster more efficient mechanisms for sharing metadata.

12:30PM–2:30PM ***Break for Lunch*** (Individual Choice)

Note: There is an open-air, pedestrian mall just outside the Omni hotel where there are many restaurants for lunch choices.

2:30PM–4:00PM ***Session 9: NARA's Electronic Records Archives*** (Salon A, Lobby Level)

"Metadata Implementation Perspectives for the ERA System."

Quyen Nguyen, U.S. National Archives and Records Administration

With the advent of Information Technology, more and more records today are digital born. In order to continue to fulfill its mission in the computer age, the U.S. National Archives and Records Administration (NARA) has made the decision to develop the Electronic Records Archives (ERA) system.

Session 9 (continued)

The ERA system represents an endeavor undertaken by the agency to preserve digital records, and make those records accessible independently of hardware and software with which they are created. Metadata is an important element in such a system whose core functionality is digital preservation for long term access by the public.

This paper will discuss the potential issues and impact of implementing metadata in ERA from the perspectives of system architecture, data management, and software design.

We also present the information technologies that we are considering for the implementation of metadata within the ERA system such as XML, and Web services. By referring to the OAIS information model, we will look at different types of metadata, and how the system could support the creation and maintenance of these metadata automatically and manually via workflow. Meeting the security requirements with different levels of data and metadata classification also constitutes a challenge in the system design process. The decision of how to store metadata vis-à-vis records and record aggregates will significantly impact the software design object model, the storage size, and data replication. Metadata as well as data replication are critical to ensure the availability and safeguard of the archival records.

"Indexing and Search Implementation Perspectives for the ERA System."

Dyung Le, U.S. National Archives and Records Administration

With the advent of Information Technology, more and more records today are born digital. In order to continue to fulfill its mission in the computer age, the U.S. National Archives and Records Administration (NARA) has made the decision to develop the Electronic Records Archives (ERA) system. The ERA system represents an endeavor undertaken by the agency to preserve digital records over an indefinite period of time, and make those records accessible independently of the hardware and software with which they were created. The capability for indexing and searching of its assets is an important element in a system whose core functionality is digital preservation of electronic records for long term access by NARA and the public.

Session 9 (continued)

This paper will discuss the potential issues and impacts related to the implementation of Indexing and Search functionality in ERA from the perspectives of system architecture, software design, usability, and long term maintainability. We will present the information technologies represented by the major vendors of Enterprise Search COTS that we are analyzing for possible implementation of Indexing and Search services within the ERA system.

Meeting the information retrieval needs of the diverse, and potentially huge, ERA user community, given the resource limitations of ERA is a serious challenge. We will discuss options being considered by NARA to meet this challenge. Meeting the security requirements of a solution which must house records with different levels of classification or that contain sensitive information also constitutes a challenge for the Indexing and Search service. ERA is intended to exist for an essentially indefinite period of time, and its service oriented architecture provides the flexibility to evolve over time as technology changes, including changing out COTS products. There are no current or emerging standards (other than for metadata) governing the Enterprise Search arena. Hence there is a real danger of becoming locked into a particular Enterprise Search vendor's proprietary approach. The paper will discuss the related technical issues and possible mitigations.

2:20PM–4:00PM ***Session 10: OAI for Digital Library Aggregation***
(Salon B, Lobby Level)

"OAI for Digital Library Aggregation: An Update."

- David Seaman, Digital Library Federation
- Kat Hagedorn, University of Michigan
- Martin Halbert, Emory University
- Sarah Shreeves, UIUC
- Perry Willett, University of Michigan
- Tom Habing, UIUC

DLF in partnership with Emory, Michigan, and UIUC, is researching, designing, and prototyping a "second generation" OAI finding system, capitalizing on the lessons learned from the first wave of OAI harvesting and using as its raw material collections drawn from across the DLF membership. The aim is to foster better teaching and scholarship through easier, more relevant discovery of digital resources, and enhance libraries' ability to build more responsive local services on top of a distributed metadata platform.

Session 10 (continued)

This panel will update the DLF community on the progress of this work, and solicit feedback while we are still in medias res. The major deliverables will be described and demonstrated such as they can be, with particular emphasis on the first three, which are furthest along at this point:

1) Best Practices guidance for OAI use in libraries, with particular emphasis to the recommendation that we adopt MODS as the metadata schema to convey the richness of description that we are convinced we need to build OAI records that truly support innovative scholarship. The first version of the Best Practices document will be available online by the Forum and in print soon after, as grant deliverables. Emory University and the other DLF IMLS partners have also developed a curriculum series of OAI best practices training materials. These materials will be used to train staff and coordinate activities of DLF libraries interested in sharing metadata concerning their digital collections, and is intended to be shared with the larger digital library community.

2) A pair of portal prototype finding systems, informed heavily by feedback received from the grant-funded Scholars Advisory Panel, http://www.diglib.org/architectures/oai/imls2004/OAISAP05.htm. One portal, The DLF OAI Portal, offers a single place to access all OAI records (items and collections) from DLF institutions: http://www.hti.umich.edu/cgi/b/bib/bib-idx?c=imls;page=simple
The second, in production now, takes 330,000 MODS-based OAI records from four DLF institutions and is building a prototype service that reflects the service and functional desires of our scholarly team.

3) An Experimental OAI Registry at UIUC of use principally to builders of OAI services: http://gita.grainger.uiuc.edu/registry. The most significant recent additions to the registry are rich, human-generated collection descriptions for many of the DLF member OAI data providers, including description of select subsets. These data are browsable via the registry Web interface or as XML which conforms to the DC Collection Description profile.

4) A Survey of Digital Library Aggregation Services, version 2: as part of the grant, Martha Brogan is revisiting her 2003 Survey, and we will be publishing the results early in 2006.

As the grant progresses, we are also expecting to look at auto-characterization of data, Web services, and interfacing our prototype systems with Google. Evaluation will be a significant component later in the grant period.

4.00PM–4.30PM ***Break*** (Prefunction Area, Lobby Level)

4.30PM–6.00PM ***Birds of a Feather (BOF) Session 2:***

5. *Open Archives Initiative Protocols for Metadata Harvesting: Best Practices for Data Provider Implementation and Sharable Metadata* (James Monroe, Lobby Level)

   Sarah L. Shreeves, UIUC

   A working group made up of members of the DLF and NSDL and representing both service and data providers have been developing a set of best practices for OAI data provider implementations and shareable metadata (http://oai-best.comm.nsdl.org/cgi-bin/wiki.pl ?TableOfContents). Specifically the Best Practices for OAI Data Provider Implementations (http://oai-best.comm.nsdl.org/cgi-bin /wiki.pl?DataProviderPractices) offer guidelines and recommendations for a range of integral and optional pieces of the protocol (deleted records, sets, datestamps, descriptive containers). The Best Practices for Shareable Metadata (http://oai-best.comm.nsdl.org/cgibin/wiki.pl ?IntroductionMetadataContent) outline general guidelines for authoring metadata that is useful and effective within larger aggregations.

   These best practices are now beginning to go under public review and comment in preparation for publication in the coming year. This session is open to anyone who would like to discuss these best practices and ask questions or raise concerns with members of the working group.

6. *Preservation Planning for Digital Objects and Repositories* (Ashlawn/Highlands, Lobby Level)

   Taylor Surface, OCLC Online Computer Library Center
   Stephen Abrams, Harvard University

   Many of us are implementing or operating repositories of digital materials with an eye toward preserving the objects for the long-term. While there has been much theoretical discussion on practices, those of us with these repositories now face the very real challenge of providing these services. Come to this BOF to share your current practice for digital preservation planning and discuss opportunities for creating best practice.

BOF Session 2 (continued)

7. *METS Implementation and METS Profile Building* (Lewis/Clark, Lobby Level)

   Nancy J. Hoebelheinrich, Stanford University
   Brian Tingle, California Digital Library

The METS Editorial Board will hold a Birds of a Feather session on Technical Issues related to METS implementation and METS Profile Building. Members of the METS Editorial Board who have successfully written, registered and implemented METS profiles will be in attendance, as well as members of the METS community who are in the process of developing METS profiles (including Arwen Hutt from UCSD and Rob Wolfe from MIT/DSpace). Topics to be discussed include identifying the purpose and function that a METS profile can serve in local implementations of METS, how and whether local workflow should influence the development of a profile, and how profiles are designed to facilitate content and metadata sharing among members of the METS community. Lively discussion will be encouraged!

DAY THREE: WEDNESDAY, NOVEMBER 9, 2005

8:00AM–9:00AM ***Breakfast*** (Atrium, Lobby Level)

9:00AM–10:30AM ***Session 11: Digital Library Initiatives in South Africa and England*** (Salon A, Lobby Level)

"Herding Big Cats: An African Experience of Collaboration."

D. P. Peters, University of KwaZulu Natal

DISA, Digital Imaging South Africa, is a national collaborative digitization project, funded by The Andrew W. Mellon Foundation, to make available for international research, the repressed archival documentation of the apartheid era in South Africa. Some 70,000 pages have already been made available from http://disa.nu.ac.za.

In partnership with Aluka, a project of Ithaka Harbors Inc., the project has recently embarked upon a second phase, to make available research resources of neighboring regimes under the regional topic, Southern African Freedom Struggles, 1950–1994. This presentation will share an African experience of collaboration, in building partnerships with users to provide context, and with institutions to provide content.

Session 11 (continued)

The South African apartheid system can be equated on the level of crimes against humanity that history and responsible stewardship must prevent in the future. Digital technologies ideally serve this aim in the dissemination of information, but the process of building collaboration is not without pitfalls, beyond the organizational challenges.

The DISA project has recently focused on developing a common understanding amongst librarians, archivists, scholars and politicians, of its role in interpreting a sensitive and painful period of history.

Engagement of the scholarly community serves to build contextual layers in the information architecture, with descriptive essays linked to the archival resources by means of topic maps. The objective is to build a research resource for teaching and learning, stimulating curriculum development in this subject area. But archival concerns for a perceived loss of ownership must be juggled with access, and local heritage preservation with global cultural imperialism.

This presentation will investigate some of the benefits and pitfalls experienced in building new user communities in national and international collaboration.

"Digital Library Activities at Oxford."

Michael Popham, University of Oxford

The libraries of Oxford University have a long-standing interest in digital technologies. Large-scale digitization projects undertaken more than a decade ago such as "Early Manuscripts at Oxford" (http://www.image.ox.ac.uk/) are still attracting a growing number of users who wish to access selected items from our extensive holdings. However, by 2000 it had become apparent that the hitherto piecemeal and largely project-based approach to the selective digitization of material was not going to be sustainable in the longer term. Quite apart from the resources required to support and maintain dozens of separate Web sites built on a variety of applications and data standards, it was clear that digital surrogates were becoming acceptable to the scholarly community and greatly increasing public access to material that most readers would be unlikely to see first hand. A new approach was required.

Session 11 (continued)

In the summer of 2001, Oxford University Library Services established the Oxford Digital Library (ODL): a combination of services and technologies intended to develop, test, and implement the policies and standards that would underpin a University-wide framework for the digitization of library holdings. Thanks to a generous grant from the Andrew W Mellon Foundation, a Development Fund was established to create a testbed of core content for the ODL intended to be used by researchers, teachers, and the global community of learners.

The initial 4-year development phase of the ODL concludes in October 2005, and this presentation will outline the lessons we have learned to date, the implications for the way the ODL is likely to develop, and also look at the impact of such endeavors as the Oxford-Google digitization partnership.

This presentation will provide an update on digital library developments at the University of Oxford, outline the lessons that have been learned from the initial four-year development phase of the Oxford Digital Library, and discuss the likely impact of the Oxford-Google digitization agreement.

9:00AM–10:30AM ***Session 12: Web Archiving Services***
(Salon B, Lobby Level)

- Martha Anderson, Library of Congress
- John Tuck, The British Library
- Taylor Surface, OCLC Online Computer Library Center
- John Kunze, California Digital Library

Web archiving services emerging at a number of different institutions will enable librarians and other document selectors to extend their historic collection-building roles into the domain of web-based materials. Such services will allow curators to initiate and monitor web crawls relevant to specific topic areas, analyze and annotate harvested data, and search and browse local archives built from sites that may have been harvested multiple times.

1) "Introduction to Web Archiving" (Martha Anderson): a brief overview of the current landscape of challenges and opportunities of archiving web resources.

2) "Web Archiving at the British Library" (John Tuck): The British Library is lead partner in the UK Web Archiving Consortium (UKWAC) (www.webarchive.org.uk) and is a member of the International Internet Preservation Consortium (IIPC).

Session 12 (continued)

The focus of the presentation will be on collaborative working nationally and internationally. There will be specific reference to the challenges faced by UKWAC in areas such as permissions and legal deposit, software, and collection development and, in the case of IIPC, to current initiatives including progress on procurement for an automated smart crawler in conjunction with the Bibliothèque nationale de France.

3) "UIUC/OCLC's ECHO DEPository Project" (Taylor Surface): OCLC, as part of the ECHO DEPository NDIIPP project, is leading the development of a suite of open source web archiving tools named the Web Archives Workbench, which is based on an archival selection model developed at the Arizona State Library. OCLC will discuss the challenges facing state libraries in the collection of web information and review how the tools of the Web Archives Workbench help with those challenges.

4) "CDL's Web Archiving Service" (John Kunze): An overview of CDL's Web Archiving Service (WAS) and its approach to long-term preservation. The approach includes generating "dessicated data" (long-lived, low-tech derivatives for certain formats), defining service levels, assigning persistent identifiers, and replicating content at geographically distant locations.

10:30AM–11:00AM ***Break*** (Prefunction Area, Lobby Level)

11:00AM–12:30PM ***Session 13: Sustaining Digital Scholarship***
(Salon A, Lobby Level)

"Sustaining Digital Scholarship."

Bradley Daigle, Mike Furlough, Thornton Staples, and Madelyn Wessel (all University of Virginia)

Sustaining Digital Scholarship ("SDS") is a project at the University of Virginia Library that explores the complex technical, legal, institutional and policy issues arising for libraries in the development and formal collection of original digital scholarship. In the humanities, these born-digital scholarly efforts tend to look less like existing genres based on print-models (i.e., monographs, articles) and more like exhibitions, library collections, and thematic research archives. Such projects challenge us to develop consistent methods for production, delivery, rights management, access, and archiving of digital content of multiple media and content types. To sustain original digital scholarship, we assume that we must move beyond providing individual piecemeal solutions to define standard methods for collecting these projects by the library.

Session 13 (continued)

SDS is a collaboration among the University of Virginia Library, NINES (Networked Interface for Nineteenth Century Electronic Scholarship), the Tibetan and Himalayan Digital Library, and the Virginia Center for Digital History. Pilot projects under SDS assume that: (1) the library will formally select, collect, preserve and distribute original digital scholarly projects through a digital library architecture based upon Fedora; (2) that intellectual property rights of those projects allow open access to the broadest extent possible; (3) that the library will strive to preserve the intellectual content, structures, and designs of the project; and (4) that the library will elaborate formal collection agreements with the scholars and possibly other institutions.

Mike Furlough will moderate and discuss the overall aims of the SDS project at Virginia; Thornton Staples will outline the project's theory of collection and content aggregation; Madelyn Wessel will review the policy and legal issues that the project raises for libraries and scholars; Bradley Daigle will discuss the implementation of pilot projects and expected outcomes.

11:00AM–12:30PM ***Session 14: Archive-it: A Web Archiving Application*** (Salon B, Lobby Level)

- Merrilee Proffitt, RLG, Moderator
- Michele Kimpton, Director Web Archive, Internet Archive
- Carolyn Palaima, University of Texas at Austin
- Cecile Jagodzinski, Indiana University
- Kathy Jordan, Library of Virginia
- Dan Avery, Internet Archive

Archive-it is a Web application uniquely designed for the needs of University and government institutions interesting in preserving Web content. The application allows organizations with limited infrastructure and technical staff to collect, catalogue, search and manage archived Web content through a Web interface.

The Internet Archive (IA), a nonprofit that manages the largest publicly available Web archive, developed Archive-it. IA currently provides these services to large institutions such as Library of Congress and the US National Archives. It is working with RLG and a handful of other organizations to make the same service available at a scale and cost that is broadly accessible. RLG member institutions participating in this pilot are Indiana University, the International Institute of Social History, Swarthmore College (with partner Haverford College), and the University of Toronto.

Session 14 (continued)

Other pilot participants working directly with the Internet Archive include the Library of Virginia, University of Texas at Austin, and North Carolina State Archives. The pilot run of the service is scheduled to conclude in November 2005, and the service is scheduled to launch in January 2006.

This presentation will include an overview of Archive-it and its major functions; pilot participants will give an overview of why they are interested in Web archiving, challenges they face in their own institutions regarding Web archiving, what they've learned so far using the Archive-it Web application, and how it's being applied in their institution. By the time of this panel, participants will be able to discuss their experience with Archive-it, challenges of Web archiving in general, and provide information and informed experiences to audience members.

POST-CONFERENCE: WEDNESDAY, NOVEMBER 9, 2005

12:30PM–4:30PM ***OAI Vendors' Panel— for project participants*** (Ashlawn, Lobby Level)

1:00PM–5:00PM ***DLF Developers' Forum Meeting— for project participants*** (James Monroe, Lobby Level)

1:00PM–5:00PM ***METS Editorial Board Meeting*** (Monticello, Lobby Level)

1:30PM–5:30PM ***CCS docWORKS Workshop— for project participants*** (Computer Classroom, Fourth Floor, Alderman Library, University of Virginia)

POST-CONFERENCE: THURSDAY, NOVEMBER 10, 2005

9:00AM–5:00PM ***METS Editorial Board Meeting*** (Monticello, Lobby Level)

9:00AM–5:00PM *DLF* ***OAI Implementers' Workshop— for project participants*** (Computer Classroom, Fourth Floor, Alderman Library, University of Virginia)

BIOGRAPHIES

A

Stephen Abrams is the Digital Library Program Manager at the Harvard University Library, where he provides technical leadership for strategic planning, design, and coordination of the Library's digital systems, projects, and assets. He is currently engaged in research and implementation of effective methods for archival preservation of digital objects. Mr. Abrams was the project manager for JHOVE, an extensible Java framework for format-specific object identification, validation, and characterization; the project leader and document editor for ISO/TC171/SC2/WG5, the joint working group that developed the PDF/A standard (ISO 19005-1); and is leading efforts to establish a Global Digital Format Registry (GDFR).

Martha Anderson is the Library of Congress NDIIPP program officer for the California Digital Library Web at Risk project. She is the manager of the LC Web Capture team in the Office of Strategic Initiatives and chairs the Content Management Working Group for the International Internet Preservation Consortium. During 2004 and 2005, she served as the project manager for the NDIIPP Archive Ingest and Handling Test(AIHT). Previously she coordinated the production of American Memory.

Dan Avery is Senior Crawl Engineer at Internet Archive. Prior to joining the Internet Archive, Dan worked as a fast food cashier, theme park ride attendant, university recycling coordinator, gopher developer, VP of a software company, graduate student in sociology, software engineer, and research scientist, but not all at the same time. He has a B.A. in interdisciplinary studies from the University of Virginia.

B

Ellie Brown is Head of Program and Project Management in the Division of Rare and Manuscript Collections at Cornell University Library. She has oversight for RMC's archival processing, Encoded Archival Description project, and has overall responsibility for the management of all Divisional projects including exhibitions and public programs.

Brown (continued)

Before coming to Cornell in 2001, she worked at the National Archives of Canada as Project Coordinator in the establishment of a digital resources program and as an archivist in the Documentary Art and Photography Section and the Manuscript Division, with extensive experience in textual and media collections.

C

Karen Calhoun is the Assistant University Librarian for Technical Services at Cornell University Library. Karen has been with CUL since 1997. Previously Karen worked for OCLC and the University of Oregon. She has an M.S. in library and information science and an M.B.A.

Adam Chandler is Information Technology Librarian within the Central Technical Services department of the Cornell University Library, where his responsibilities include creation of new automated technical services processes, participation in library-wide technology initiatives and management of the department's computers. His year 2000 assignment, to explore how to build a database to help manage the library's administrative metadata for electronic journals, led to his meeting Tim Jewell and creating the "Web Hub for Developing Administrative Metadata for Electronic Resource Management." That path led to his becoming a member of the ERMI steering group.

Currently, Adam is technical lead for Cornell's implementation of III's ERM stand alone module. His ERMI standards involvement continues with work on the mapping of license terms to ERM systems; also, he has a growing curiosity about the potential relationship between electronic resource usage statistics and ERM systems.

Timothy W. Cole is Mathematics Librarian and Professor of Library Administration at the University of Illinois at Urbana-Champaign. He is principal investigator for an IMLS-NLG project to build a collection registry and OAI metadata repository for digital content created or developed under the auspices of IMLS, as well as for a CIC collaborative metadata harvesting service based on OAI. He is past chair of the NSDL Technology Standing Committee and a former member of the OAI Technical Committee. He has published widely on OAI-PMH, metadata, and the use of XML and SGML for encoding sci-tech journal literature, and has spoken about OAI at multiple venues including OAI4, the IMLS Web-Wise Conference, ALA Annual Meeting, JCDL, and the Open Archives Forum.

Esme Cowles is a software engineer at the UC San Diego Libraries, currently working for the Union Catalog for Art Images metadata project and developing the VRA Core 4.0 XML schema. Previously, he has worked on the Libraries' electronic resources database (Sage), the main UC San Diego Web site and campus map, and the PCASSO online medical records project.

Patricia Cruse is the Director of the Digital Preservation Program at UC's California Digital Library. She is responsible for overseeing all activities related to digital preservation, which includes the establishment of a UC libraries Digital Preservation Repository (DPR) and evaluating methods for the persistent management of content to support the research, teaching, and learning at the University of California. In addition, Ms. Cruse is overseeing CDL's National Digital Information Infrastructure Preservation Program (NDIIPP) grant, which will build a Web-archiving service to enable libraries to collect and manage Web-published content.

D

Bradley Daigle works at the University of Virginia Library as Associate Director, overseeing a unit that digitizes the library's special collections and rare materials—Rare Materials Digital Services (RMDS). This unit is unique within our library in that it focuses on research and commercial requests to digitize materials and maintains its own donor-funded collection building projects. Since many research efforts centered in the library draw upon special collections materials, RMDS often supports many of the digital projects undertaken by researchers in the library. Given the unique nature of special collections material and its widespread demand for reuse, Bradley frequently engages with issues of intellectual property and digital rights management.

Before that he was also Project Supervisor for the Virginia Heritage Project where my duties included overseeing the training and implementation of new technology for this NEH funded project for marking up manuscript finding aids in XML-enhanced Encoded Archival Description (EAD). He provided technical support and production guidelines for fifteen state institutions to ensure the long-term codification of EAD standards.

Bradley received a master's degree in literature from l'Université de Montréal in 1996 and a MLS from the Catholic University of America in 1999. He has been active in both the digital library community as well as regional archival associations—presenting papers on collaborative digitization projects in Virginia as well as digitization best practices.

Daigle (continued)

He is leading a statewide team of Special Collections staff to establish statewide standards and best practices for creating digital objects from special collections materials.

Stephen Davis (BA, MA in German Literature from Yale, MS in Library Service/Information Science from Columbia) currently directs Columbia University's Library Digital Program. He got his start in library computing as an analyst in the Library of Congress's Network Development & MARC Standards Office, and then spent over a decade as Director of Columbia's Library Systems Office. He has written and spoken on digital library management, standards, and metadata.

Tim DiLauro is the Digital Library Architect in the Library Digital Programs and Digital Knowledge Center of the Sheridan Libraries at Johns Hopkins University. Since 1982, he has worked for JHU as a Programmer, Systems Programmer, and Sr. Systems Programmer, with a network programming and management component. He has been with the Sheridan Libraries since 1990. He has also worked as a consultant for several companies with Internet businesses. Since 1995, his project work has focused on designing systems to improve and simplify user access to information, including the development of access gateways and Web proxies. His current work deals with the integration of multiple repositories with multiple services to support digital collections, learning, publishing, and preservation.

David Dubin is a Senior Research Scientist on the staff of the University of Illinois Graduate School of Library and Information Science. Together with Professor Allen Renear, David leads the Electronic Publishing Research Group at GSLIS. His current research activities focus on issues in digital preservation (as part of the UIUC ECHO-DEP project led by John Unsworth and Beth Sandore) and the semantic analysis of markup languages (in partnership with the University of Bergen's MLCD project). In addition to his research activities, Dr. Dubin teaches classes in information processing, quantitative research methods, and electronic publishing technologies.

Jon Dunn is Associate Director for Technology in the Digital Library Program at Indiana University, overseeing the development and management of software systems to support IU's digital library collections. Prior to joining the Digital Library Program, he worked in the Cook Music Library at IU from 1994–1998 as Technical Director for the Variations digital music library project.

Dunn (continued)

He is currently serving as Project Director for IU's IMLS-funded Variations3 digital music library and learning system development project and chairs the DLF Aquifer Technology/Architecture Working Group.

F

Jeremy Frumkin is the Gray Family Chair for Innovative Library Services at the Oregon State University Libraries. He received his MLIS from Florida State University in 1996, and has previously worked at OCLC, Inc. and the University of Arizona Main Library. Mr. Frumkin is currently a co-Principal Investigator on the OCKHAM grant project, funded by the National Science Foundation, and is involved in a number of digital library research initiatives. The founding chair of LITA's Open Source Systems Interest Group, Mr. Frumkin has long been a proponent of open source software in libraries, and his current work focus is around open, distributed digital library systems and workflows.

Mike Furlough is the Director of Digital Research and Instructional Services at the University of Virginia Library. He is responsible for developing new technology service programs to support digital scholarship and manages the activities of several of the Library's groundbreaking digital centers. Mike is active in the American Library Association, EDUCAUSE, and is currently working on the DLF Aquifer initiative on the services working group.

H

Thomas G. Habing is a Research Programmer at the Grainger Engineering Library Information Center at the University of Illinois at Urbana-Champaign where for the past eight years he has worked on various digital library projects. In addition to his technical support for various ongoing OAI-PMH related projects at UIUC, including being the developer of the UIUC OAI Registry, Tom is a technical lead for the Library’s NDIIPP ECHO DEPository grant project. Before the OAI-era, Tom was a lead developer on the Library’s NSF funded Digital Library Initiative (DLI I) project, and the CNRI funded DLib Test Suite projects. Prior to returning to the Midwestern, U.S. in 1997, Tom was a Senior Computing Methods and Technology Engineer for The Boeing Company in Seattle, Washington, where he had been employed since 1986 doing systems analysis, programming, and graphical user interface design.

Kat Hagedorn is OAIster / Metadata Harvesting Librarian at the University of Michigan Libraries. She currently manages the OAIster project, a search gateway for OAI harvested records leading to digital objects, initially Mellon-funded in 2001–2002. She is also responsible for DLXS Bibliographic Class and co-coordinates the processing of Text Class materials. Her previous experience is in information architecture (with the Argus Associates firm) and ontology and taxonomy consulting (with the Food and Agriculture Organization in Rome). She graduated with an undergraduate degree in Biological Sciences from Cornell University and got her MLIS at the University of Michigan in 1996.

Martin Halbert is Director for Library Systems at Emory University. He is currently a principal investigator on the NSF-funded Ockham Project, on DLF's IMLS-funded work to research, design, and prototype a "second generation" OAI finding system, and on two Mellon-funded metadata harvesting initiative projects. He also serves as executive director of the MetaScholar Initiative, a consortium of thirty institutions working to aggregate metadata for scholarly portal services. Martin serves as the chair of the LOCKSS sub-committee on Institutional Access Integration, and has there studied issues of low-cost library server networks and associated integration issues. He has served as editor of several library publication projects, and currently supervises a university library division of fourteen professional staff.

Nancy J. Hoebelheinrich is Metadata Coordinator for the Digital Services Group at the Stanford University Libraries / Academic Information Resources. In that capacity, Nancy coordinates metadata services for Stanford Libraries' digital production activities, digital repository development and implementation, and educational technology services. She has been a member of the METS Editorial Board since 2002 and is currently serving as co-chair. Nancy has been active in a number of information and educational technology specification efforts including that of PREMIS (for preservation metadata), and several of IMS Global specifications related to packaging, repository and resource list interoperability. She is currently involved with the IEEE Learning Technology Standards Committee's RAMLET project, and continues to monitor various groups working on practices related to the use of digital rights expression languages.

J

Cecile Jagodzinski is the Director of Collection Development and Digital Scholarship at the Indiana University Libraries.

Jagodzinski (continued)

She formerly held positions at Illinois State University, the American Medical Association, and Northwestern University Law Library. She holds an M.L.S. from the State University of New York at Buffalo and a Ph.D. in English from the University of Illinois at Urbana-Champaign. Along with her work on collection management issues in research libraries, she pursues her research interests in the history of the book and of libraries.

Keith Johnson is the Product Manager for the Stanford Digital Repository (SDR). In this role he is responsible for continually defining the digital preservation services being built and offered by the SDR via discovering, monitoring, and reconciling user needs, emerging best practices in the digital preservation field, and the institutional priorities of Stanford University. Keith brings deep experience in the commercial content creation process having spent not only the majority of his career in pre-press and publishing, but also significant parts in television, advertising, and classical music performance. Keith recently worked for Scholastic Inc., where he was responsible for developing collaborative electronic print and media publishing workflow systems and corporate digital archiving services.

Leslie Johnston is the Director of Digital Access Services at the University of Virginia Library, where she manages digital library program components supporting the collection, management, and dissemination of digital content. Previously, she served as the Head of Instructional Technology and Library Information Systems at the Harvard Design School, where she managed the implementation of instructional technology projects for faculty and coordinated information systems and new media projects for Design Library.

Prior to that, Ms. Johnston worked as the Academic Technology Specialist for Art for the Stanford University Libraries, Systems Project Coordinator at the Historic New Orleans Collection, and as Database Specialist for the Getty Research Institute. Ms. Johnston also served for many years on the Board of Directors of the Museum Computer Network, and was founding editor of ESpectra, the MCN news portal for the cultural heritage information management community.

K

Joanne Kaczmarek is the Archivist for Electronic Records, for the University of Illinois at Urbana-Champaign. In this capacity, she is working to develop a self-sustaining records management model into which the ongoing management of electronic records can be strategically and somewhat seamlessly positioned. Prior to her current position, she was the Project Coordinator on the University of Illinois Cultural Heritage Metadata Repository, an initial OAI-PMH Mellon Foundation project focused on testing the extensibility of the protocol. Her professional interests include: exploring ways to manage electronic records as part of larger information systems, extending digital preservation activities across institutional boundaries, and identifying and redefining the shifting roles of information professionals.

Michele Kimpton has been a Director at the Internet Archive for three years. In her role, she works closely with national libraries, archives and universities to provide technical expertise and services in Web archiving. She has developed partnerships with several of these institutions to collaborate on Web archiving activities, including co-founding the International Internet Preservation Consortium.

Prior to the Internet Archive, Michele worked in the high-tech-industry, mainly for-profit sector, for the last 20 years. Before coming to the Internet Archive she was one of the co-founders of an online digital imaging company, which was subsequently bought by one of the larger photo imaging companies. For the last ten years of her career she has worked primarily in technical management and business development. She has worked and lived in both Europe and Asia during her career.

Katherine Kott is the director of the DLF Aquifer digital library initiative. Her professional career has included experience in academic library systems, and in technical and public services. Most recently, Kott was the head of cataloging and metadata services at Stanford University, where she is based. Before coming to Stanford, she led the implementation services department at a major ILS vendor, coordinating the installation of systems at a wide range of libraries, including consortia.

John Kunze is a preservation technologist for the California Digital Library and has a background in computer science and mathematics. His current work focuses on archiving Web sites, creating long-term durable digital references (ARKs) to information objects, and specifying lightweight (kernel) metadata. Prior work includes major contributions to the standardization of URLs, Dublin Core metadata, and the Z39.50 search and retrieval protocol.

Kunze (continued)

In an earlier life he designed, wrote, and ran UC Berkeley's first campus-wide information system, which was an early rival and client of the World Wide Web. Before that he was a BSD Unix hacker whose work survives in today's Linux and Apple systems.

L

Bill Landis, currently metadata coordinator at the California Digital Library, received his archival training and MLIS from the University of Michigan in 1994. He served as the first production coordinator for JSTOR in 1996-1997, and as manuscripts librarian in Special Collections and Archives at the University of California, Irvine from 1998-2004.

Dyung Le is Director of System Engineering, responsible for all technical aspects of the Electronic Records Archives (ERA) program at the U.S. National Archives and Records Administration (NARA), and chair of the ERA Engineering Review Board. The ERA system will be a comprehensive, systematic, and dynamic means for preserving virtually any kind of electronic record, free from dependence on any specific hardware or software; and to provide workflow automation for the Record Life Cycle. The ERA Project Office leads the development of this system and sponsors and collaborates in research and exploratory development initiatives in partnership with government agencies, associations, national archives in other countries, and universities. Prior this position, Mr. Dyung Le served as Senior Director, Server Engineering at MicroStrategy, Inc. where he directed the MicroStrategy Business Intelligence Server development team. He also served as Development Manager at Digital Equipment Corporation, where he led the development of Video-on-Demand project and as Program Manager for DARPA funded supercomputing program. Mr. Le is a holder of a M.S. and B.S. in Electrical Engineering and Computer Science, and a B. S. in Applied Mathematics from the Massachusetts Institute of Technology, Cambridge, MA.

Bill LeFurgy is Digital Initiatives Project Manager with the Library of Congress Office of Strategic Initiatives. He manages projects for the National Digital Information Infrastructure and Preservation Program (NDIIPP). These efforts include overseeing advanced digital preservation projects, guiding research on improved methods for long-term management of digital materials, and working with government agencies, private corporations, professional organizations, and other stakeholders.

LeFurgy (continued)

Prior to joining LC in 2002, Bill served as Deputy Director of Modern Records Programs with the National Archives and Records Administration. He worked for the National Archives over the course of 12 years in areas such as electronic records archiving, records appraisal, and management. He worked extensively with many federal agencies in developing strategies for the long-term management and preservation of permanent records, particularly those in electronic form.

O

Evan Owens is Chief Technology Officer of Portico (www.portico.org), an electronic archiving service for electronic scholarly journals with initial funding provided by the Andrew W. Mellon Foundation, Ithaka, the Library of Congress (NDIIPP), and JSTOR. Prior to joining Portico, he worked for the University of Chicago Press as IT Manager and as Electronic Publishing Manager of the Journals Division.

P

Alan Pagliere is University of Michigan Digital Library Production Services Programmer. Pedal steel guitarist, fallen linguist, and former beekeeper. Current work includes designing, programming and maintaining the software which is used by the UM Digital Library and 29 other institutions for developing web access to text data (monographs, journals, structured and page-image-based) and finding aids.

Carolyn Palaima is the Project Director for LANIC (lanic.utexas.edu) at the Teresa Lozano Long Institute of Latin American Studies, University of Texas at Austin. She has worked on Internet-related projects since LANIC's inception in 1992. LANIC works closely with the Latin Americanist Research Resources Project and in partnership has launched a series of Web-based resources including the Latin American Periodicals Tables of Contents and the Latin American Open Archive Portal. LANIC is currently a partner on the Internet Archive pilot project "Archive-it." Carolyn holds a B.B.A. in Finance and an M.A. in Latin American Studies from the University of Texas at Austin.

Michael Pelikan is Technology Initiatives Librarian at Penn State, and Librarian to the Penn State School of Information Sciences and Technology (IST). Prior to becoming a librarian, Pelikan worked for many years in public radio as a reporter, program director, production director, and independent producer/recording engineer. He has also worked in public television as a production engineer and in higher education IT management. A Chicago native, Pelikan was raised in Connecticut and lived 14 years in Alaska.

D. P. Peters is Project Manager of DISA: Digital Imaging South Africa http://disa.nu.ac.za/), a non-profit making initiative for cooperation among research libraries and archives in Southern Africa, sponsored by the Andrew W. Mellon Foundation. Dr Peters currently provides overall leadership to the project in supporting the development of digital library services across national institutions, while providing a central digital library service component.

Her remit encompasses in a wide range of research, technical, legal and financial functions, including copyright law and heritage management policy, knowledge organization, computer systems and user interaction related to cultural heritage preservation and digital library management. She is responsible for the planning, development and management of scholarly digital resources, maintaining the highest level of technology awareness to inform and implement national policy, to advance the vision and to direct the project's mission and goals.

Dr. Peters holds a PhD in Information Studies from the University of Natal in Pietermaritzburg (1999). She is a founder member of the South African Preservation and Conservation Group (SAPCON), and currently serves on the national executive of the South African Society of Archivists (SASA). In addition, she is Editor of the South African Archives Journal. A member of the South African Museums Association (SAMA) and the Library and Information Association of South Africa (LIASA), she is a wide ranging specialist in cultural heritage management.

An early adaptor to the use of digital information technologies to promote preservation through enhanced access, Dr. Peters has a personal research interest in digital preservation, and a commitment to the objectives of enhanced access. She is a leading authority in the southern African region in the use of digital technologies in higher education, and plays an active role in research institutions across the country in developing policies, strategies and guidelines in support of digital library development and electronic resource management.

Peters (continued)

Dale Peters has consulted for national and international bodies, including UNESCO, The International Federation of Library Associations and Institutions (IFLA), the International Council on Archives (ICA) and is a research consultant on digital technologies and specifically on the development role of digital libraries in the African context.

Michael Popham took up the post of Head of the Oxford Digital Library in the summer of 2003, transferring from the role of Project Manager for the Oxford e-Science Centre based within Oxford University's Computing Services. Between 1996–2002, Michael was Head of the Oxford Text Archive which hosts the Centre for Literature, Languages, and Linguistics of the UK's national Arts and Humanities Data Service (AHDS). Michael holds first and higher degrees in English and American Literature, and an M.Sc. in Computer Science. He has previously worked as a technical author and as a management consultant, and has served on the committees of several professional bodies, notably as Chair of the British Computer Society's Electronic Publishing Specialist Group (1996-2000). Married with three small children, he lives in a picturesque village just north of Oxford, with fantastic views but no discernable public transport.

Chris Powell is the Coordinator of Encoded Text Services and the Humanities Text Initiative at the University of Michigan's Digital Library Production Service. She is responsible for the online delivery of encoded texts and finding aids through DLXS, Michigan's suite of tools for mounting searchable collections of digital library content.

Merrilee Proffitt has been a Program Officer at RLG since 2001. She serves as a liaison with RLG member community on a variety of programmatic endeavors within RLG's research resources community, and is involved in digital library standards work (such as the Text Encoding Initiative, the Metadata Encoding & Transmission Standard, and Encoded Archival Description). Merrilee has been directly involved with RLG's Primary Resources community, and chaired the working group that authored the award winning "RLG Best Practice Guidelines for Encoded Archival Description." Recently, Merrilee has become interested in end user issues, primarily because of her involvement with the RedLightGreen project. Prior to her work at RLG, Merrilee managed digital library projects at UC Berkeley.

R

Anthony Ramirez is the Research Project Manager for the R. H. Smith School's effort to preserve confidential digital records from the Dot Com era. Anthony received his undergraduate degree in Industrial and Labor Relations from Cornell University in 2004, and plans to pursue law school in the fall of 2006.

Erin Rhodes is a Digital Imaging Specialist in the special media labs at the National Archives and Records Administration, where she works on imaging projects in support of the agency's digital initiatives. Prior to coming to NARA, she worked at the University of Chicago Library and with the Colorado Digitization Project, and received an Advanced Certificate in Preservation Administration from the Preservation and Conservation Studies Program at UT-Austin in 1998. She is interested in digital preservation, metadata, repository design, and digital imaging for preservation reformatting.

Jenn Riley is the Metadata Librarian with the Indiana University Digital Library Program. In addition to METS Navigator, the subject of her joint DLF Spring Forum 2005 presentation, she currently works on a number of other digital library projects, including the NSF-funded Variations2 Digital Music Library, the NEH-funded Ethnomusicological Video for Instruction and Analysis, and the IMLS-funded IN Harmony Indiana Sheet Music collaborative project. She holds Bachelor's and Masters' degrees in music in addition to an MLS.

Trish Rose has had a wide variety of experience in library, museum, and academic settings in which she has provided computer systems support, developed classification systems, and overseen project management and execution of digital initiatives. In the past two years, Trish has been heavily involved in the development of metadata standards for the visual resource community both as a member of the VRA Core 3.0 development team and as an advisory committee member for the Cataloging Cultural Objects guidelines. Currently, Trish is an image metadata librarian at the University of California, San Diego working on a Mellon-funded research and development project called UCAI. UCAI, which stands for Union Catalog for Art Images, is developing the prototype for a shared cataloging utility for art image metadata.

Nan Rubin has more than twenty years of experience managing technical and communications projects for public radio and television, including more than five years engaged in technology planning at Thirteen.

Rubin (continued)

Projects have included introducing program scheduling software at Thirteen, and being a core member of the Special Task Force that oversaw restoration of Thirteen's broadcast signal after the destruction of the World Trade Center.

She planned and coordinated the creation of the Thirteen Tape Archives and has been instrumental in assisting Thirteen in developing an initiative for the long-term preservation of its program assets, both analog and digital. As Project Director of this Library of Congress-funded effort, she is responsible for managing all aspects of project work, overseeing tasks and activities, convening meetings, coordinating communications, making presentations, maintaining relationships with other organizations, and representing the project to the public.

S

Ryan Scherle is a software analyst/programmer in the Digital Library Program (DLP) at Indiana University. He is currently involved in migrating the DLP's collections to a new technical infrastructure based on a Fedora repository. Previously, Ryan spent several years developing server software for the Variations2 Digital Music Library project. Before that, he worked as a web developer and independent consultant. Ryan is pursuing a combined Ph.D. in Computer Science and Cognitive Science. His research focuses on leveraging the context of a user's request to guide automatic selection of information resources.

David Seaman is Executive Director of the Digital Library Federation. Prior to that, he was the founding director of the Electronic Text Center at the University of Virginia Library (1992–2002), a humanities digital library of texts and images. His published work includes studies of Chaucer, and he speaks and writes frequently on various aspects of humanities computing.

Clay Shirky chaired the Technical Committee of the NDIIPP during the design of the program and in its first year of existence. With Martha Anderson of LC, he designed and chaired the Archive Ingest and Handling Test (AIHT). He teaches networking at NYU's graduate Interactive Telecommunications Program.

Sarah L. Shreeves is Coordinator for the Illinois Digital Environment for Access to Learning and Scholarship (IDEALS) at the University of Illinois at Urbana-Champaign (UIUC).

Shreeves (continued)

Her experience with the Open Archives Initiative Protocol for Metadata Harvesting is grounded in both the IMLS DCC project and the Mellon funded OAI Metadata Harvesting Project (2001–2002) at UIUC where she worked as a graduate assistant and project coordinator.

Prior to coming to UIUC, Sarah worked for nine years in the MIT Libraries in Boston. She has a BA in Medieval Studies from Bryn Mawr College, an M.A. in Children's Literature from Simmons College, and an M.S. in Library and Information Science from UIUC.

Thornton Staples is currently the Director of Digital Library Research and Development at the University of Virginia Library and is the Project Director for the Fedora Project. Previous positions include: Chief, Office of Information Technology at the National Museum of American Art, Smithsonian Institution; Project Director at the Institute for Advanced Technology in the Humanities, University of Virginia; and Special Projects Coordinator, Academic Computing at the University of Virginia.

Eric Stedfeld serves as Information Technology Specialist in the Digital Library Development area of New York University's Division of Libraries. He received his M.S. in computer science at NYU in 2002. For several decades previously Eric designed and produced numerous videos, audiovisual presentations, interactive multimedia projects and websites for corporations, television networks, publishers, museums, and government agencies. In his current capacity at NYU Eric's particular focus includes metadata and software standardization efforts, application architecture development, implementation of new technologies, and dynamic multimedia delivery and presentation.

Taylor Surface is Manager, Digital Collection Services at OCLC. His team creates and operates services supporting the management and preservation of library-owned digital content. These services include OCLC's Digital Archive service, the Registry of Digital Masters, and OAI harvesting.

In addition, the team is the primary partner of University of Illinois (UIUC) on the Library of Congress funded ECHO DEPository project developing the Web Archives Workbench.

T

John Tuck is Head of British Collections at the British Library since October 2002. Previously worked in various roles at the John Rylands University Library, University of Manchester (1977–1997) and was then Deputy to the Director of University Library Services and to Bodley's Librarian in the University of Oxford (1998–2002).
Tuck (continued)

Current responsibilities at the British Library include overall responsibility for the Web Archiving Programme and involvement in two major digitization projects, one for 19th-century British newspapers, the other for 4,000 hours of archival sound recordings.

Jim Tuttle is the Geospatial Data Librarian and project coordinator for the North Carolina Geospatial Data Archiving Project at North Carolina State University Libraries. The joint project of the NCSU Libraries and the North Carolina Center for Geographic Information and Analysis will focus on collection and preservation of digital geospatial data resources from state and local government agencies in North Carolina.

W

John Weise is a librarian and programmer at the University of Michigan. He coordinates Image Services within the Digital Library Production Service, which hosts about 90 image databases containing over 200,000 digital images for campus and public access. Weise deploys and maintains these databases in addition to handling the programming of the Image Class component of Michigan's DLXS digital library software.

Madelyn Wessel is Special Advisor to the University Librarian and Liaison to the General Counsel, focusing on a broad range of library system legal issues including intellectual property, copyright, licensing, and special issues arising in the area of digital scholarship. Ms. Wessel is an adjunct professor at the Curry Graduate School of Education and admitted to practice in Virginia, Massachusetts, New Hampshire, and Oregon. Ms. Wessel served as Deputy and later Chief Deputy City Attorney for Portland, Oregon from 1989-2001. Ms. Wessel holds a BA from Swarthmore College and a J.D. from Boston University.

Perry Willett is the Head of the Digital Library Production Service at the University of Michigan. DLPS is responsible for digital content creation at the University of Michigan libraries, and software development for DLXS, the digital library system. In addition to his work in digital libraries, he has served as a bibliographer and reference librarian at Indiana University, SUNY-Binghamton and Columbia University.

Y

Jeffrey A. Young graduated Beta Phi Mu with an M.L.S. from Kent State University. He has been at OCLC since 1987 and works as a software architect in the Office of Research. His focus is on registries, Web services, and support and integration of protocol standards such as OAI and OpenURL. His work on WikiD was recently integrated into OCLC's Open WorldCat and DeweyBrowser services to support user-contributed content. The potential of WikiD as a general-purpose collaborative registry platform is currently being explored.

APPENDIX

WHAT IS DLF?

What is DLF? The Digital Library Federation (DLF) is a leadership organization that pioneers the use of electronic-information technologies to extend library collections and services. Through its strategic partners and allies, DLF provides leadership to libraries broadly by

- identifying standards and "best practices" for digital collections and network access
- coordinating leading-edge research, development, and delivery
- incubating projects and services that libraries need but cannot develop individually

How does DLF operate? The DLF consists of an Executive Director, a small staff, an Executive Committee, and a Board of Trustees on which each partner institution is represented. The bulk of its work on many initiatives is performed by working groups of its partners and others in the scholarly, library, and computing communities. DLF brings together experts from across disciplines and industries. The Council on Library and Information Resources (CLIR) is the administrative home to DLF.

DLF Strategic Partners contribute annually to the DLF's operating budget and pledge funds over five years to its capital fund. Each member institution has a seat on the Board and the responsibility to help direct the organization. The current partners are as follows:

Bibliotheca Alexandrina
British Library
California Digital Library
Carnegie Mellon University
Columbia University
Cornell University
Council on Library and Information Resources
Dartmouth College
Emory University
Harvard University
Indiana University
Johns Hopkins University
Library of Congress
Massachusetts Institute of Technology
New York Public Library
New York University

WHAT IS DLF? (continued)

North Carolina State University
Pennsylvania State University
Princeton University
Rice University
Stanford University
University of California, Berkeley
University of Chicago
University of Illinois at Urbana-Champaign
University of Michigan
University of Minnesota
University of Pennsylvania
University of Southern California
University of Tennessee
University of Texas at Austin
University of Virginia
University of Washington
U.S. National Archives and Records Administration
Yale University

DLF Allied Partners are organizations working in proximate areas. A senior officer from each allied organization sits on the DLF Board of Trustees "with voice but without vote."

Coalition for Networked Information (CNI)
Joint Information Systems Committee (JISC)
Los Alamos National Laboratory Research Library (LANL)
OCLC Online Computer Library Center, Inc. (OCLC)
RLG

DLF Staff maintain the Executive Director's Office and are responsible for achieving program goals and priorities, facilitating and supporting DLF initiatives, managing communications, and administering finances and the work of the governing Board. The central office staff includes:

Executive Director: David Seaman (dseaman@clir.org)
Program Associate: Barrie Howard (bhoward@clir.org)
Administrative and Communications Associate: Christie Hartmann (chartmann@clir.org)

What does DLF provide?

Leadership and support for new research, standards development, and project start ups. Notable successes include OAI, METS, the Registry of Digital Masters, and the emerging Electronic Resources Management Initiative (ERMI).
WHAT IS DLF? (continued)

A semi-annual Forum to report on developments, standards, and projects, to plan new areas of collaborative endeavor, and to allow members to share experiences and find new colleagues.

E-mail listservs to exchange information, announce initiatives, identify resources, and stimulate discussion.

A Web site (http://www.diglib.org/) to provide public access to information about activities, resources, developments, and DLF itself.

Periodic newsletters and a pair of online databases to provide access to digital collections available from DLF members, and digital library documentation (policies, strategies, working papers, standards, and technical documentation).

Publications for reporting on research and conferences, the progress of initiatives, and on members' digital-library services, collections, projects, and challenges.

Multiple partnership opportunities, a sense of community and shared vision, and an opportunity to collaborate with a rich array of digital library practitioners and theorists.

For more information, contact:

David Seaman
Executive Director
Digital Library Federation
Council on Library and Information Resources
1755 Massachusetts Avenue, NW, Suite 500
Washington, DC, 20036
dlf@clir.org
www.diglib.org

RECENT DLF PUBLICATIONS

Acquiring Copyright Permission to Digitize and Provide Open Access to Books. Denise Troll Covey. Washington, D.C.: Digital Library Federation, Council on Library and Information Resources, 2005. http://purl.oclc.org /dlf/trollcovey0509.

What are the stumbling blocks to digitization? Is copyright law a major barrier? Is it easier to negotiate with some types of publishers than with others? To what extent does the age of the material influence permission decisions? This report, by Denise Troll Covey, principal librarian for special projects at Carnegie Mellon University, responds to many of these questions. It begins with a brief, cogent overview of U.S. copyright laws, licensing practices, and technological developments in publishing that serve as the backdrop for the current environment. It then recounts in detail three efforts undertaken at Carnegie-Mellon University to secure copyright permission to digitize and provide open access to books with scholarly content.

A Kaleidoscope of Digital American Scholarship. Martha L. Brogan, with assistance from Daphnée Rentfrow. Washington, D.C.: Digital Library Federation, Council on Library and Information Resources, 2005. http://www.diglib.org/pubs/2005brogan/.

This report will be useful to anyone interested in the current state of online American literature resources. Its purpose is twofold: to offer a sampling of the types of digital resources currently available or under development in support of American literature; and to identify the prevailing concerns of specialists in the field as expressed during interviews conducted between July 2004 and May 2005. Part two of the report consolidates the results of these interviews with an exploration of resources currently available. Part three examines six categories of digital work in progress: (1) quality-controlled subject gateways, (2) author studies, (3) public domain e-book collections and alternative publishing models, (4) proprietary reference resources and full-text primary source collections, (5) collections by design, and (6) teaching applications. This survey is informed by a selective review of the recent literature.

Electronic Resource Management: A Report of the DLF Initiative. Timothy D. Jewell, Ivy Anderson, Adam Chandler, Sharon E. Farb, Kimberly Parker, Angela Riggio, and Nathan D. M. Robertson. Washington, D.C.: Digital Library Federation, 2004. http://www.diglib.org/pubs/dlfermi0408/.

RECENT DLF PUBLICATIONS (continued)

As libraries have worked to incorporate electronic resources into their collections, services, and operations, most have found their existing Integrated Library Systems to lack important functionality to support these new resources. An earlier study (Jewell 2001) determined that a number of libraries had begun developing local systems to overcome these shortcomings, and the DLF Electronic Resource Management Initiative (ERMI) was organized to aid the rapid development of such systems by providing a series of interrelated documents to define needs and to help establish data standards.

Digital Library Content and Course Management Systems: Issues of Interoperation. A Report of a study group funded by the Andrew W. Mellon Foundation, and co-chaired by Dale Flecker, Associate Director for Planning & Systems, Harvard University Library, and Neil McLean, Director, IMS Australia. Washington, D.C.: Digital Library Federation, 2004. http://www.diglib.org/pubs/cmsdl0407/.

An ad hoc group of digital librarians, course management system developers, and publishers met under the aegis of the Digital Library Federation to discuss the issues related to the use of digital library content in course management systems. The size, heterogeneity, and complexity of the current information landscape create enormous challenges for the interoperation of information repositories and systems that support course instruction. The group has created a checklist of things that operators of digital content repositories can do to help ameliorate the complexities of such interoperation. It also explored through the means of use cases the utility of tools which help instructors gather information resources from various distributed information repositories for teaching purposes, and created a model of how the group envisions the interaction of users, tools, and information repositories in the future.

A Survey of Digital Library Aggregation Services. Martha L. Brogan. Washington, D.C.: Digital Library Federation, 2003. http://www.diglib.org/pubs/brogan/.

This report, commissioned by the DLF, provides an overview of a diverse set of more than thirty digital library aggregation services, organizes them into functional clusters, and then evaluates them more fully from the perspective of an informed user. Most of the services under review rely wholly or partially on the Protocol for Metadata Harvesting of the Open Archives Initiative (OAI-PMH).

RECENT DLF PUBLICATIONS (continued)

Each service is annotated with its organizational affiliation, subject coverage, function, audience, status, and size. Critical issues surrounding each of these elements are presented in order to provide the reader with an appreciation of the nuances inherent in seemingly straightforward factual information, such as *audience* or *size*.

Archiving Electronic Journals: Research Funded by the Andrew W. Mellon Foundation. Edited, with an Introduction, by Linda Cantara. Washington, D.C.: Digital Library Federation 2003. http://www.diglib.org/preserve/ejp.htm.

Increasingly, scholarly journals are published electronically. What does it take to keep them accessible electronically in perpetuity? Can the property rights of publishers, the access responsibilities of libraries, and the reliability assurances that scholars need be reconciled in agreements to create archives of electronic journals? These series of studies from seven major libraries examine various aspects of the challenges of archiving electronic journal content.

An Introduction to *Dimensions and Use of the Scholarly Information Environment*. Amy Friedlander. Washington, D.C.: Digital Library Federation, Council on Library and Information Resources, 2003. http://www.clir.org/pubs/reports/pub110/contents.html.

659 Data Tables for *Dimensions and Use of the Scholarly Information Environment.* Washington, D.C.: Digital Library Federation, Council on Library and Information Resources, 2003. http://www.diglib.org /pubs/scholinfo/.

We know from anecdotal evidence that users' expectations of libraries are changing as they find more information directly from the Web, but anecdotes are an insufficient basis for developing new library services. DLF and CLIR commissioned Outsell, Inc. to conduct a large-scale study to give us a much more reliable picture of user behaviors. Published here are the 659 data tables that record the responses to 35 groups of questions asked of 3,200 undergraduates, graduate students, and faculty members from academic institutions ranging from small liberal arts colleges to the largest public and private research universities. Accompanying them is a summary of the findings and 158 selected data tables; it should be viewed as an entry to the much larger data set of 659 data tables provided above.

RECENT DLF PUBLICATIONS (continued)

The Digital Library: A Biography. Daniel Greenstein & Suzanne E. Thorin**.** Washington, D.C.: Digital Library Federation, Council on Library and Information Resources, 2002. http://www.clir.org/pubs/abstract /pub109abst.html.

Digital libraries, once project-based and largely autonomous efforts, are maturing. As individual programs have grown, each has developed its own personality, reflecting the circumstances of its creation and environment, and its leadership. This report from CLIR and the DLF draws on the results of a survey and case studies of DLF members to reveal how these influences have molded a range of organizational forms that we call the digital library. The report is written by Daniel Greenstein and Suzanne Thorin. Greenstein, formerly the director of the DLF, is now university librarian and director of the California Digital Library. Thorin is the dean of university libraries at Indiana University. *Section One* of the report examines three stages of digital library growth: the young digital library, the maturing digital library, and the adult digital library. *Section Two* of the report presents case studies of digital library development at six institutions.

For full lists of DLF and CLIR publications are available from http://www.diglib.org/pubs/dlfpubs.htm, and http://www.clir.org /pubs/pubs.html.

DLF-ANNOUNCE LISTSERV

The Digital Library Federation (http://www.diglib.org/) is a consortium of thirty-nine libraries and related agencies that are pioneering the use of information technologies to extend, share, and manage their collections and services. Through its members, DLF provides leadership for libraries broadly by:

- identifying standards and best practices for digital collections and networked access
- coordinating research and development in the use of information technology
- incubating projects and services that libraries need but cannot develop individually

The best way to keep up with DLF's initiatives, Forums, calls for collaboration, and news is to subscribe to the DLF-ANNOUNCE discussion list, available to all members and selected guests.

To subscribe:

1. Post an e-mail to the following address: listserv@www.diglib.org;
2. Leave the Subject line blank and remove your signature block from the body;
3. Type: subscribe dlf-announce [your first name] [your last name];

The DLF listserv will send a welcome message to you; and

You may leave the list at any time by following steps 1 through 3, substituting the word *signoff* for *subscribe*, mentioned in step 2.