07

WebWise

Stewardship in the Digital Age:
Managing Museum and Library Collections
for Preservation and Use

February 28–March 2, 2007
CONFERENCE PROCEEDINGS

INSTITUTE OF MUSEUM AND LIBRARY SERVICES
1800 M Street NW, 9th Floor
Washington, DC 20036
202-653-IMLS (4657)
www.imls.gov

Printed June 2007 in the United States of America

Written by Diane M. Zorich, except
where otherwise noted

The 2007 WebWise Conference on Libraries and Museums in the Digital World was held March 1–2, 2007, with pre-conference workshops on February 28, 2007. Conference activities took place at the Hyatt Regency on Capitol Hill and at Georgetown University in Washington, D.C.

The 2007 WebWise Conference was co-hosted by the Institute of Museum and Library Services, the J. Paul Getty Trust, and the Online Computer Library Center.

Library of Congress Cataloging-in-Publication Data not available at the time of printing

Dear Colleague:

The eighth annual WebWise conference, co-sponsored by IMLS, OCLC, and the J. Paul Getty Trust, was held February 28–March 2 at the Hyatt Regency Hotel in Washington, D.C. This year's theme was "Stewardship in the Digital Age: Managing Museum and Library Collections for Preservation and Use." The response from the library and museum communities was overwhelming—more than 400 participants, representing all types of museums and libraries nationwide, registered in the first few weeks after registration opened.

In order to extend the benefit of the conference, we asked Diane Zorich to prepare this summary. In addition, audio podcasts and PowerPoints of the main conference presentations are posted on the IMLS Web site at www.imls.gov/news/events/webwise07.shtm. Full-text papers of many of the presentations appear in the July issue of the online journal *First Monday,* which has been publishing WebWise papers in a special issue each year since the first conference in 2000 (see www.firstmonday.org).

Stewardship is a core responsibility of museums, libraries, and archives. Yet a study conducted by Heritage Preservation in partnership with IMLS found that current preservation and emergency preparedness practices in cultural heritage institutions are largely inadequate. *A Public Trust at Risk: The Heritage Health Index Report on the State of America's Collections* (www.heritagepreservation.org/HHI) called for immediate action to prevent the loss of millions of irreplaceable artifacts. IMLS has responded by making the centerpiece of its tenth anniversary year a new initiative called *Connecting to Collections.* The WebWise 2007 Conference was a key component of this initiative.

In just the past few years, museums and libraries have progressed from knowing virtually nothing about how to preserve digital assets to understanding that digitization is an important part of conservation and use. In addition to preserving the tangible objects in their care, there is an increasing awareness that institutions need digital repositories for collections that are physically vulnerable, on fragile or unstable media, or born digital.

How can cultural heritage institutions ensure the preservation of digital surrogates? How will they preserve the new forms of expression that exist only in digital form? How can they use digital technologies to document, track, and manage their collections more effectively? These questions and related issues were addressed at WebWise 2007 by leaders of national institutions who face these challenges on a massive scale, by technology and preservation experts who conduct research in digital preservation, and by library and museum professionals from a variety of institutions that are developing strategies to help their institutions meet immediate needs and prepare for the digital future.

Libraries and museums must position themselves to use digital technologies strategically in the stewardship of collections. We at IMLS believe that digital stewardship is an important part of the overall mission of libraries and museums in caring for their collections, and we, like you, want to be part of the solution.

Sincerely,

Anne-Imelda M. Radice

Anne-Imelda M. Radice, PhD
Director, IMLS

(Bolded words throughout the publication are included in the glossary.)

Project: Turtle Trails (see page 77).
Photo: "Bogue," a loggerhead sea turtle. North Carolina Aquarium.

Introduction

> The papyrus and the CD, the codex and the Web page; human ingenuity strives, not only to find a permanent medium for its culture, but to reclaim that which was fixed in a more fragile form.
>
> - Stuart Kelly

Stewardship is a concept with many components: guidance, preservation, responsibility, and care are among its most important aspects. For cultural heritage, stewardship means ensuring the products of society's creation—from the tangible (objects) to the intangible (language or ritual)—are cared for and made accessible to future generations.

Increasingly, this type of stewardship will occur in a digital realm. The 2007 WebWise Conference addressed this reality with its theme "Stewardship in the Digital Age: Managing Museum and Library Collections for Preservation and Use." Over the course of two and a half days (February 28–March 2, 2007), several dozen speakers addressed the state of digital preservation "readiness" in our nation's cultural institutions, emerging practices for preserving digital content, and strategies for using technology to promote and improve stewardship of our cultural heritage.

The presentations included overviews of the state of preservation today, summaries of community surveys, discussions of new tools, reports on cross-community and cross-organizational collaborations, updates on the activities of governmental agencies, and case studies about cultural stewardship. Two keynote addresses focused on access and preservation and highlighted underlying conference themes such as access, collaboration, holistic approaches to preservation, and community-based solutions.

Information about the WebWise conference and its presentations is available in several formats. Conference papers appear in the online journal *First Monday* (www.firstmonday.org), and podcasts and PowerPoint presentations are available at the conference Web site (www.imls.gov/news/events/webwise07.shtm).

This summary report provides another format, offering an abbreviated version of speakers' perspectives and discussions on digital preservation and access in the cultural heritage community, and highlighting the continuing challenges and opportunities cultural organizations face as they strive to be true stewards of our cultural legacy.

For eight years, the WebWise conferences have offered cultural heritage professionals a forum to discuss and learn about themes critical to the role of libraries and museums in the digital world. The 2007 WebWise Conference was hosted by three partners: the Institute of Museum and Library Services (IMLS, www.imls.gov), the J. Paul Getty Trust (www.getty.edu), and the Online Computer Library Center (OCLC, www.oclc.org). Each organization has extensive interest and experience in promoting cultural heritage using digital technologies.

Dr. Anne-Imelda Radice, Director of IMLS, described her agency's goal as helping to build infrastructure that supports libraries and museums, and creating action agendas, high expectations, and collaborations that promote tangible outcomes. In celebration of the agency's tenth anniversary, IMLS is working with other federal agencies and private funders to spotlight U.S. collections through a "Connecting to Collections" initiative (www.imls.gov/about/collections.shtm). The 2007 WebWise Conference contributes to this major initiative by addressing the preservation challenges affecting our national collections as they are made accessible by digital means.

Dr. Ken Hamma, Executive Director of Digital Policy at The J. Paul Getty Trust, spoke of the many programs of the trust, and of the difficulty of speaking "across programs." Hamma talked about the opportunities we miss because we encase ourselves in organizational divisions—e.g., museums, libraries, archives—instead of conversing across these institutional lines to gain greater perspectives and possibilities. Users who seek our information do not care about our institutional or programmatic delineations. If we perpetuate them in the online world we risk making our collections less, rather than more, accessible. Hamma also reminded the audience that access and preservation are concepts embedded within our institutional traditions and grounded in our missions. The introduction of digital technologies into these arenas gives us new tool sets and methodologies to fulfill our institutional missions and express our core values.

Jay Jordan, CEO and President of OCLC, described his organization's long and rich tradition as a computer and library research organization. OCLC's activities are centrally tied to the WebWise conference theme, particularly in its support of standards and providing open source software protocols used for **harvesting*** collections information. It also offers tools such as a new terminology service (www.oclc.org/terminologies/default.htm) that fosters greater access to collections by promoting the use of consistent terminology

*Bolded words throughout the publication are included in the glossary.

in collections records. In addition, OCLC has made its WorldCat catalogue (WorldCat on the Web, *OpenWorldCat* at www.oclc.org/worldcat/open/default.htm) available at the point where people are likely to begin their search (i.e., a search engine), and then directs them to a local library where a related bibliographic resource can be found. In this way, OCLC is helping drive users directly to a library.

Radice, Jordan, and Hamma emphasized the importance of collaboration and partnerships to preserve cultural heritage and make it more accessible. Digital technologies clearly have an important role in this effort. As Dr. Radice noted in an introductory statement:

> How can cultural heritage institutions ensure the preservation of digital surrogates? How will they preserve the new forms of expression that exist only in digital form? How can they use digital technologies to document, track, and manage their collections more efficiently?

2007 Program Committee:

Liz Bishoff, University of Colorado, Boulder
Matt Burdetsky, Capital Meeting Planning, Inc.
Tom Clareson, PALINET
Ken Hamma, J. Paul Getty Trust
Jack Ludden, J. Paul Getty Trust
Elizabeth Lyons, IMLS
Amy Lytle, OCLC
Joyce Ray, IMLS
Taylor Surface, OCLC
Günter Waibel, OCLC

Stay tuned for 2008 details!

Information on the dates, location, and co-hosts of the 2008 WebWise conference will be available soon. To receive updates on this and many other IMLS activities, visit www.imls.gov to subscribe to our monthly e-mail newsletter, *Primary Source,* or subscribe to our RSS feed at http://www.imls.gov/rss.xml.

Previous WebWise Conferences:

2006: Inspiring Discovery: Unlocking Collections
Co-Hosts: J. Paul Getty Trust and Online Computer Library Center
Location: Los Angeles, CA

2005: Teaching and learning with Digital Resources
Co-host: University of Illinois at Chicago
Location: Washington, D.C.

2004: Sharing Digital Resources
Co-host: University of Illinois at Chicago
Location: Chicago, Illinois

2003: Sustaining Digital Resources
Co-host: Johns Hopkins University
Location: Washington, D.C.

2002: Building Digital Communities
Co-host: Johns Hopkins University
Location: Washington, D.C.

2001: The Digital Divide
Co-host: University of Missouri at Columbia
Location: Washington, D.C.

2000: A Conference on Libraries and Museums in the Digital World
Co-host: University of Missouri at Columbia
Location: Washington, D.C.

Project: Saving Nebraska's Treasures (see page 78).
Photo by Tom Sain.

Pre-Conference Workshops

February 28, 2007

This workshop introduced current recommended practices in digital preservation. It began with a tutorial that summarized the basic vocabulary, concepts, methodology, and tools used in the field of digital preservation. Following this introduction were presentations from three cultural heritage practitioners who discussed issues pertaining to the preservation of Web sites, of public television programming, and of digital art.

Priscilla Caplan, Assistant Director for Digital Library Services at the Florida Center for Library Automation (www.fcla.edu), provided the introductory presentation. She began by explaining that the rubric *digital preservation* really has three components: curation, archiving, and preservation.

Curation refers to "data curation," or the process of maintaining, managing, and promoting the use of data from creation through dissemination for discovery and use. *Archiving* is the specific data curation activity that involves making certain that information is properly selected, stored, and made accessible. It also ensures that the integrity of information, and the physical format encapsulating that information, remains secure and authenticated over time. *Preservation* is an archiving activity which ensures that specific items of data are maintained over time so that they can be accessed and understood regardless of changes in technology.

Successfully preserving digital collections requires a holistic, life-cycle management strategy that emphasizes creation, appraisal, documentation, and reuse. You cannot just preserve a digital object: You must preserve the entire digital ecosystem. Caplan offered a database analogy to illustrate this concept. When you export data values from a database, you gain access to information. But to truly preserve the context of that information you must preserve the database software's characteristics, its reports, forms, and query screens, its tools, and any other aspect of the database that makes the data values understandable.

Caplan discussed several aspects critical to successful archiving and preservation:

- Availability—You must have access and control over the material you archive. Some of the ways to gain control of digital materials are through Web harvesting, by having departments place their materials in an institutional repository, by licensing digital resources from suppliers, or by negotiating deposit agreements with owners of materials.
- Identity—You must know what a resource is, which is why descriptive metadata is critical for digital preservation.
- Understandability—A resource must be understandable to an archive's designated community. If the context or documentation for the resource is lost, understandability is threatened. To prevent this from occurring, you must archive the resource and any documentation or contextual information that makes the resource understandable.
- Authenticity—An object must be what it purports to be, not just what someone says it is. Authenticity is confirmed when both the content and the source of the content can be verified. To ensure authenticity, you must have (and maintain) a complete event history and chain of custody for the resource, as well as possess documented proof of the fixity (see below) of the resource.
- Fixity—A resource must not be accidentally altered or deleted. Insecure storage, data transmission errors, or media degradation jeopardize the fixity of resources. Methods for ensuring fixity include good storage management, media refreshment, and running programs that calculate and review **checksums** (to detect transmission errors).
- Viability—The digital resource must be readable from digital media. This quality is jeopardized when media degrade or become obsolescent. Media refreshment and migration are two methods used to ensure viability.

- Renderability—A resource must be "displayable" or otherwise made usable. Digital objects cannot be rendered when their file formats become obsolete. To circumvent the obsolescence issue you must either maintain or emulate old hardware and software, or reformat the digital object into current renderable formats.

Even when all these aspects of archiving and preservation are successfully addressed, other factors may interfere with preservation efforts. Karen Coyle, in her report for the Library of Congress entitled "Rights in the PREMIS Data Model,"[1] outlined several rights scenarios that affect preservation. These scenarios vary from broad legal mandates (e.g., a state legislature requiring that a state archive preserve all state agency records) to a limited set of legal strictures placed on preservation activities (such as those outlined in 17 U.S.C. 108 of the US copyright code). In between are a host of other possibilities: The copyright owner may or may not give you permission to preserve his/her materials, or the materials may be in the public domain so anyone can preserve them, or you may be the copyright owner and thus can do what you wish.

Similarly, economic incentives often affect preservation activities. Brian Lavoie, in his paper entitled "The Incentives to Preserve Digital Material,"[2] describes a triad of economic interests among those who have the right, those who have the need, and those who have the ability to preserve. Problems arise when these interests are not aligned: for example, when the beneficiaries (those who have the need) are not the rightholders.

Moving from theory into practice, Caplan talked about practical tools and methodologies that are pushing the preservation of digital collections forward in a constructive manner. First among these is the Open Archival Information System (OAIS, http://ssdoo.gsfc.nasa.gov/nost/isoas) reference model. This model provides a common vocabulary of concepts that facilitate description and comparison of archives. It tells you *what to do* when preserving a digital collection, but not *how to do it*.

A second development in digital preservation is the creation of "trusted digital repositories." These are organizations or organizational entities that provide reliable, long-term access to managed digital resources for a particular community. Many universities, for example, are creating campus-based repositories for the digital materials created by their faculty and staff, and some states have designated statewide repositories for official state collections. As more of these repositories emerge, the issue of trust has grown in importance. What are the attributes of a trustworthy repository? A number of initiatives are under way that address this question and provide tools for assessing digital repositories. (See WebWise speaker Robin Dale's discussion of these initiatives on page 34.)

A third trend in the field is the growing number of preservation methods that now exist. Version **migration**, **emulation**, **software engineering**, and **format standardization** are but a few of the digital preservation methods available. Dr. Kenneth Thibodeau, in an article entitled "Overview of Technological Approaches to Digital Preservation and Challenges in Coming Years,"[3] details the many different technological ways to preserve digital collections.

New developments are also taking place in the area of preservation metadata, defined as "the information a repository uses to support the digital preservation process." The PREMIS data dictionary[4] developed by RLG and OCLC is becoming a de facto standard in this area, defining core preservation metadata pertaining to objects, agents, events, and rights.

---

1 www.loc.gov/standards/premis/Rights-in-the-PREMIS-Data-Model.pdf.

2 www.oclc.org/research/projects/digipres/incentives-dp.pdf.

3 In *The State of Digital Preservation: An International Perspective*. Conference Proceedings, July 2002. Council of Library and Information Resources. www.clir.org/pubs/reports/pub107/thibodeau.html.

4 *PREMIS Data Dictionary version 1.0*. Maintained by the Library of Congress at www.loc.gov/standards/premis.

When developing preservation metadata, cultural heritage professionals must consider the most significant properties of what they are planning to preserve. What is the essence or property that you need to retain? For example, when you preserve or can peaches, you are preserving their flavor, not the fruit as it looks when it is plucked from the tree. Caplan offered a more visual example in the following excerpt from A.A. Milne's *Winnie the Pooh*:

this take
If is shall really to
flying I never it.

Piglet getting bounced along by Kanga,
Winnie the Pooh, A.A. Milne, p. 103

this take
If is shall really to
flying I never it.

...transcribed as linear text

Caplan's PowerPoint slide presentation, slide 27.

Clearly, the significant property of this passage —the essence of the information—cannot be captured by the text alone. A strict linear interpretation of the text not only garbles the meaning of the passage, but it fails to convey the dimension of movement one might feel when bounced around by a kangaroo.

Lastly, Caplan addressed the developments taking place in building a preservation infrastructure. While significant inroads have been made in areas of standards and best practices, Caplan identified the following areas that need further work:

- File format registries that provide standard categories for describing digital file formats
- Environment registries that describe various hardware and software environments
- Technology watch services that track technologies and advise us when their obsolescence is likely
- Distributed networks of trusted repositories
- Large-scale storage management technologies

Caplan emphasized that there is no "one size fits all" when preserving digital collections. Everyone must keep abreast of developments in the field and make choices appropriate to local circumstances. In response to an audience question about what a user can do now to prepare for using a repository service in the future, Caplan strongly advised the audience to actively curate their data, identify what is to be preserved, record preservation metadata, gather files and documentation, consider and document the significant properties of the material, and then store it in the safest way possible until a trusted repository is chosen.

In addition to the publications mentioned above, Caplan also recommended the following resources:

- The Library of Congress's "Sustainability of Digital Formats" (www.digitalpreservation.gov/formats) for current assessments of digital formats.
- The IMLS/NISO publication *A Framework of Guidance for Building Good Digital Collections* (www.niso.org/framework/Framework2.html) for insights on format usage and guidance on building digital collections.
- The chapter "Appraisal and Selection" in the Digital Curation Centre's publication, *Digital Curation Manual* (www.dcc.ac.uk/resource/curation-manual/chapters/open-source) for practical advice about selecting what to preserve.

- Roy Rosenzweig's article "Scarcity or Abundance? Preserving the Past in the Digital Era" in *The American Historical Review (*www.dcc.ac.uk/resource/curation-manual/chapters/open-source) for insights into questions that future scholars, librarians, and preservationists must address in today's digital world.

Following Caplan's digital preservation overview, a panel of three practitioners discussed specific preservation projects of interest to libraries, museums, and archives.

Valerie Glenn, the first speaker in this panel, spoke about "Preserving Government and Political Information: The Web-at-Risk Project."[5] This project, funded by the National Digital Information Infrastructure and Preservation Program (NDIIPP), is a collaborative effort of the California Digital Library, the University of North Texas, and New York University. Together, these partners are creating a distributed approach to preserving our nation's political cultural heritage as it appears on the Web. They are doing so by using Web harvesting to capture government and political materials at risk of disappearing.

Glenn, who was a curator on the project, gave several examples of at-risk materials. The numerous Web sites and blogs that emerged in the aftermath of Hurricane Katrina (to help direct aid and provide real-time information for the affected area) are rapidly disappearing. One Web-at-Risk project partner has collected a "cybercemetery"[6] of defunct government agency Web sites, including such historically important sites as the "The National Commission on Terrorist Attacks Upon the United States" (more familiarly known as the "9-11 Commission").

The goals of the Web-at-Risk project are threefold: to capture political and government materials in danger of disappearing, to capture particular political events or moments in time, and to build a collection of similar or related materials. To this end, the project is building tools that help libraries capture, curate, and preserve Web-based government and political information. Project curators are developing collection plans[7] to address acquisition and selection issues such as the depth and breadth of the capture (e.g., Should you capture an entire site, or just a homepage?), presentation and access issues, and descriptive metadata needs. They also are testing capture tools[8] and harvesting services.[9]

The long-term goal of Web-at-Risk is to develop infrastructure and tools to build collections of Web-based political and governmental activities. But Glenn acknowledged, in response to an

---

5 For more information, see the Web-at-Risk project wiki at http://wiki.cdlib.org/WebAtRisk/tiki-index.php, the NDIIPP project description at http://web3.unt.edu/webatrisk, and the California Digital Archive's project-specific Web site at www.cdlib.org/inside/projects/preservation/webatrisk.

6 See Cybercemetery, University of North Texas Libraries at http://govinfo.library.unt.edu.

7 As of this writing, 15 collections plans are available online at http://wiki.cdlib.org/WebAtRisk/tiki-index.php?page=WebCollectionPlans.

8 The harvesting tools tested in this project were:

- *Heritrex,* the Internet Archive's open-source, archival-quality Web crawler (http://crawler.archive.org).
- *HTTrack,* an offline browser utility that allows downloading of a Web site from the Internet to a local directory, recursively building all directories, and getting HTML, images, and other files from the server to a local computer.
- *Web Curator Tool,* designed for libraries and other collecting institutions, a tool that allows nontechnical users to manage the harvesting process (http://webcurator.sourceforge.net/).

9 The harvesting services reviewed for this project were:

- *ArchiveIt!,* an Internet Archive subscription service that lets institutions build, manage, and search their own Web archive via a user-friendly Web application (www.archive-it.org).
- *OCLC Digital Archive,* a service that allows users to archive assets by item-by-item harvesting and submission of Web pages/Web-based documents, or by batch archiving (www.oclc.org/digitalarchive).
- *Web Archiving Service,* an archiving service developed by the Web-at-Risk partners (see project wiki at http://wiki.cdlib.org/WebAtRisk/tiki-index.php for more details).

audience question, that the issue of responsibility remains unaddressed in preservation projects of this sort. Who is responsible for undertaking these preservation efforts? Some states have assigned responsibility for archiving state Web sites to interstate agencies. However, at local levels, there is no consensus about who is responsible for archiving these materials, and it is not always intuitive as to who should shoulder this responsibility.

Mary Ide, the second panelist and Director of the WGBH Archives (www.wgbh.org/resources/archives), spoke about "Public Television: Preserving Digital Programs" from three perspectives: production, distribution, and collecting institution. Public television programs are created in myriad analog and digital formats. These formats present problems in terms of playback equipment, hardware obsolescence and replacement issues, and the high levels of expertise needed to use various equipment. Because most public television stations create programs in *both* analog and digital formats they must run parallel systems, which complicates already extensive production and workflow processes.

In 2004, WGBH, in partnership with the Public Broadcasting Service (PBS), Channel 13/WNET, and New York University, received an NDIIPP grant for a project entitled "Preserving Digital Public Television."[10] The project is developing a long-term preservation environment for digital public television programs by creating a design infrastructure, technical specifications, and a functional model repository. Common descriptive metadata standards also are being developed and reviewed. The project will adhere to the OAIS framework and use DSpace (www.dspace.org), the open software digital repository system.

For the WGBH Archives, several preservation concerns remain. Chief among them are issues about appraisal and selection, the demand by stations for more production elements and supporting materials, and sustaining the onerous cataloging and logging processes necessary to ensure appropriate levels of access. In addition, there are issues about integrity, copyright, and access, and the need for metadata standards and transcoding technologies. Capping all of this is a concern about the huge storage requirements needed to accommodate the very large size of public television programs and a need for global repositories to preserve these programs in the long term.

Ide concluded by stating that the ideal video preservation format is not yet here, but it is on the horizon. She cited projects such as "The Digital Video Preservation Reformatting Project"[11] (under the auspices of the Dance Heritage Coalition) as an example of a digital preservation project that is helping move the broader community toward better preservation practices.

Richard Rinehart was the final panelist, whose presentation, "The Romance of Lost Causes: Preserving Digital Art," addressed the distinctive characteristics and needs of "born-digital" art. Rinehart, a Digital Media Director and Adjunct Curator at the University of California Berkeley Art Museum/Pacific Film Archive and a digital media artist, brings the unique insights and perspectives of a creator, curator, and preservationist to this area. He defined digital art as fine art created "primarily using computation technologies and rendered using the same computation technologies. To illustrate just how complex digital artwork can be, Rinehart presented several examples of born-digital art. One particular work—*Ouija* (http://ouija.berkeley.edu) by Ken Goldberg—consists of a Web site, Flash files, custom software, custom-built robotics, a Ouija board, a tent, and myriad other components.

Rinehart echoed Priscilla Caplan's preservation mantra to identify "the essence of what you want

10 For a brief description of the project, see www.digitalpreservation.gov/partners/pdf/project_ebc.pdf.

11 See the Dance Heritage Coalition's summary and report on this project at www.danceheritage.org/preservation/digital.html and www.danceheritage.org/preservation/DigitalVideoPreservation1.pdf.

to preserve." For digital art, this statement can be framed specifically as "Do you want to preserve art or keep it alive?" Preserving the technology that renders the work is a misguided and impractical notion. It would be impossible to preserve the myriad technologies now used to create digital art. Even if you could, this notion is based on a false premise. The technology is *not* the work. As Goldberg's *Ouija* vividly shows, digital art is more than the technology that renders it.

Because digital art is variable in performance and form, any attempt to preserve it must embrace that variability. Music offers an apt analogy. A musical score preserves a musical composition, but it is the performance of the piece itself—which can be played by different instruments, with differing interpretations—that renders the music. Similarly, the essence of digital art is in the performance. The dilemma is how to preserve the performance, which is really behavior rather than form.

Rinehart discussed the work of an NEA-funded project, "Archiving the Avant-Garde,"[12] that addressed this very issue. Project members initially looked at computer code as a possible preservation format, since it seemed similar in function to a musical score. However, computer code, even today, is too platform dependent. In the end, they settled on the equivalent of western musical notation for digital art, creating the Media Art Notation System (MANS), a notational language used to create documentation ("scores") for digital works. MANS is an adaptation of the MPEG 21 standard, which encodes complex digital objects so they can be used across complex digital devices.[13] It is not intended for use by the artist, but rather for use by the preserving institution, which must work with the artist to determine the essential qualities of the work that should be preserved in the re-creation of future performances of the work. The preserving institution then use MANS to encode this information and, once encoded, it can be ingested into various systems.

Codifying the behavior of a work (rather than the form) is a concept foreign to most collecting institutions, which operate on the premise of preserving an "original." Digital art forces us to rethink this approach. It deflates the traditional art historical notion of a master work, such as the *Mona Lisa* or the *Venus de Milo* that must be preserved "as is."

Future showings of digital art will always be a "remix"—an interpretation of the original work. These works may be rendered according to the artist's original intent, according to a curator's view of the artist's intent, or according to the wishes of the artist's estate. Rinehart calls these alternate forms the "parallel memories" of a work. Borrowing on Caplan's earlier "peach preservation" analogy, Rinehart concluded that when it comes to digital art, we must not just preserve the peach, we must preserve the orchard.

12 See *Archiving the Avant-Garde: Documenting and Preserving Digital/Variable Media Art* at www.bampfa.berkeley.edu/about_bampfa/avantgarde.html.

13 The project partners rejected encoding schemas such as METS (Metadata Encoding and Transmission Standard) because these schemas presume that only digital objects are being described. Digital artworks frequently have nondigital components that also must be encoded to successfully render the work.

Left: Shadow Bag, 2005, by Scott Snibbe. Digital art is is variable in performance and form. To preserve a digital artwork we must ask ourselves, "What are the essential qualities of the work that need to be recreated in future performances of the work?"

This workshop demonstrated how collecting institutions can enhance their investments in digital information by making this information available for harvesting via the Open Archives Initiative Protocol for Metadata Harvesting (OAI-PMH)[14] model for resource sharing. The workshop opened with an introduction to the OAI-PMH model and the benefits it offers cultural organizations for sharing their metadata. Following this overview were presentations of metadata harvesting efforts undertaken at a museum and a research institute. A final group of "in the trenches" practitioners then spoke of how they perceive metadata harvesting will work in their own institutions for both local and collaborative projects.

Dr. Ken Hamma of the J. Paul Getty Trust opened the workshop by providing background on data-sharing projects and the problems they presented at the Getty and at other museums. While projects such as AMICO (www.amico.org) and ARTstor (www.artstor.org) demonstrated the usefulness of having aggregated databases of collections images and information, contributing to these databases was an extremely onerous process, and the quality of metadata and images being collected was inconsistent.

OAI-PMH provides a means for making this process easier and better. As a generic exchange model, it recasts the contribution process from one requiring participation in a collaborative project (for example, "us and ARTstor") to one requiring participation in a networked environment ("us and the network").

Sarah Shreeves, Coordinator of the Illinois Digital Environment for Access to Learning and Scholarship (IDEALS, www.ideals.uiuc.edu) at the University of Illinois at Urbana-Champaign gave an overview the OAI-PMH model and the importance of searchable metadata in her talk entitled "Search Interoperability, OAI-PMH, and Metadata." Shreeves, who is both a librarian and works on a campus aggregator project, began by discussing why we have to share our metadata and "push it out" for harvesting. As the OCLC "Perceptions Report"[15] shows, we can't expect users to know about our collections or to come to our Web sites. If they find us at all it will be through search engines.

Sharing our metadata by making it available for harvesting by **aggregators**[16] has benefits for both users and institutions. Users can access subject-specific resources from multiple institutions through a single source. They also receive an array of services built around the content they are seeking. Institutions receive greater exposure, a larger user base, and benefit from the multiple synergies that emerge when distributed collections are brought together.

However, the potential of shareable metadata is only possible if the metadata is of good quality. Users must get *meaningful* results from harvested metadata. Institutions must provide metadata that is understandable outside of a local institutional context. For example, in a local database documenting a map collection, individual records do not need to have the word "map" in their descriptive metadata because their presence in a map-only database makes this self-evident. However, when a record from this database appears outside its local (i.e., institutional) context, this association is stripped away and no one may realize the record describes a map.

Shreeves offered numerous "real-world" examples of harvested records that illustrate the confusion that results when records are searched outside of their local environment. In

14 www.openarchives.org/OAI/openarchivesprotocol.html

15 *Perceptions of Libraries and Information Resources.* 2005 (www.oclc.org/reports/2005perceptions.htm). This report reveals that more than 84 percent of users reach library resources via search engines; only 1 percent begin an information search from a library's homepage.

16 Some OAI-PMH service providers are OAIster, an aggregator of digital resources (www.oaister.org); CIC Metadata Portal, which aggregates metadata describing the information resources of CIC libraries (http://cicharvest.grainger.uiuc.edu); and IMLS Digital Collections and Content, which aggregates information on digital resources developed by IMLS grantees (http://imlsdcc.grainger.uiuc.edu).

one instance it was impossible to determine what a record was describing; in another, a title, drawn from a larger local collection, totally misrepresented the underlying work.

Shreeves discussed two basic models for sharing metadata on networks. The first model is a federated search where the search occurs over multiple databases, the results are sorted, duplicates are removed, and the remaining results are returned to the user. The second model is data aggregation, where an aggregating service provider collects metadata from different places, pools it together, and the user searches from the pooled resource.

OAI-PMH is an example of the second model. Shreeves described OAI-PMH as a "plumbing tool" that simply moves metadata back and forth between data providers and service providers. It consists of a set of rules that define communication between two systems (such as FTP and HTTP) and facilitate the aggregation of metadata (like a union catalogue). Although OAI-PMH requires, at a minimum, the use of simple Dublin Core (http://dublincore.org), it supports and even encourages use of other community-developed metadata schemas. It also allows the data provider to control what it wants to share, how to share it, and with whom. By offering this level of control in a straightforward exchange model, OAI-PMH lowers the barrier for data providers so they can share their data with minimal effort.

OAI-PMH moves records that *describe* digital objects: It does not move the objects themselves. Aggregation and delivery of digital objects —digitized texts, images, data, software, etc.—is the next frontier in resource discovery research. The Mellon Foundation is funding a new initiative, called Open Archives Initiative: Object Reuse and Exchange (OAI-ORE, www.openarchives.org/ore) to facilitate this process. The goal of OAI-ORE is to devise a way to represent digital objects and their various parts, as well as develop services that can access and ingest them for use beyond the borders of local institutions.

Shreeves summed up her talk with the statement that "OAI-PMH is easy, metadata is hard." The tool for "pushing out" our metadata is there. We now have to create good shareable metadata to ensure better interoperability and thus more exposure and use of our collections information.

For more information about OAI-PMH and shareable metadata, Shreeves recommends the following resources:

- Tennant, Roy. "Bitter Harvest: Problems and Suggested Solutions for OAI-PMH Data and Service Providers." www.cdlib.org/inside/projects/harvesting/bitter_harvest.html.
- OAI-PMH Best Practices: http://oai-best.comm.nsdl.org/cgi-bin/wiki.pl?OAI_Best_Practices.
- Use of Multiple Metadata Formats with OAI-PMH: http://oai-best.comm.nsdl.org/cgi-bin/wiki.pl?MultipleMetadataFormats.

Erin Coburn, Manager of Collections Information at the J. Paul Getty Museum, moved from theory to practice when she spoke about "Sharing Images and Data: The Museum Perspective." Coburn outlined her institution's learning curve with shareable metadata and how it has rethought the entire process of sharing collections information.

A catalyst for their rethinking was the AMICO Library. To participate in this initiative, museums had to contribute a certain amount of information in a certain way, using AMICO-based tools. Museums such as the Getty were putting an extraordinary amount of time into this process. Ironically, after having expended time and energy creating AMICO records of their holdings, they rarely went back to their own collections management systems to add the newly updated information identified during the AMICO contribution process. Nor did they uniformly update their records in the AMICO Library with any new changes added into their collections management systems. These problems overwhelmed the process, making this model for data sharing unsustainable.

The AMICO project ceased operation and the library was taken over by various distributors[17] who approached the Getty about continuing to participate in the library. Realizing that they couldn't keep up with AMICO, let alone many AMICO-like projects, the museum needed to find a different model for sharing its collections metadata. This new model had to reduce overhead expended on contributing to union catalogues and service providers, reduce labor and delivery costs, ensure mechanisms for updating information, and include a link that would bring people back to the Getty's "home" content. The museum did not want to create a new standard ("there are already so many to choose from…"), so the Getty identified existing standards that would work for it. It chose *Cataloging Cultural Objects* (CCO)[18] as its data content standard, the *Categories for the Description of Works of Art* (CDWA)[19] as their conceptual framework, and the OAI-PMH model as the mechanism for making its records harvestable using open standards like HTTP and XML. The result, christened "CDWA Lite,"[20] is an XML schema that describes core records for art and cultural materials that can be harvested through OAI-PMH.

Development of CDWA Lite was a collaborative effort. The Getty Museum worked with ARTstor on creating the XML schema, and the Getty Museum and the Getty Research Institute provided the data (on paintings and images of tapestries, respectively). ARTstor was the service provider, and various collaborative departments at the Getty (the Web group, the Research Institute's Information Systems Department) served as internal service providers. Once the data were entered, ARTstor began to harvest them.

The CDWA Lite schema provides a very minimal set of information—what would be considered "core" museum documentation (also referred to as "tombstone" data). The Getty selected fields where the data were consistent and compliant with CCO. Italso included a link back to the Getty's Web site, where layers of contextual materials are available for users. Coburn displayed an example of why this link is important by taking the audience from a harvested record of a Gustav Courbet painting to the place on the Getty Web site where there are direct links to an exhibit of Courbet's works, his biography, publications on the artist, images and records of other Courbet works in the Getty collections, etc. Offering this link in the harvested record points users back to the museum's Web site, where they can find a wealth of context and information about the work and its creator that they may not have known existed.

Coburn noted that CDWA Lite is one solution for making museum metadata sharable. Before using it, however, she suggested that museums consider several of what she termed "reality checks." First, in order to use CDWA Lite, you must have data. Fortunately, the CDWA Lite elements map to information that nearly all museums collect about their objects. Second, getting information out of your local system into XML in accordance with the CDWA Lite schema can be a challenge, but some system developers are working on software solutions to make this process easier. Third, making CDWA Lite XML records available for harvesting can be difficult, but OAICat (a software protocol used for harvesting collections information, www.oclc.org/research/software/oai/cat.htm) is being modified to work with CDWA Lite more seamlessly so this effort will soon be less onerous. Finally, museums must address the broader question of "How good are our data?" Coburn displayed records on three different paintings located in three separate museums, all of which represented the theme of "St. Francis Receiving the Stigmata."

---

17 The AMICA Library (www.davidrumsey.com/amica); Catalog of Art Museum Images Online (CAMIO, http://camio.rlg.org); and H.W. Wilson Art Museum Image Gallery (www.hwwilson.com/Databases/artmuseum.htm).

18 *Cataloging Cultural Objects: A Guide to Describing Cultural Works and Their Images* (CCO, www.vraweb.org/ccoweb/cco/index.html).

19 *Categories for the Description of Works of Art* (CDWA, www.getty.edu/research/conducting_research/standards/cdwa).

20 CDWA Lite (www.getty.edu/research/conducting_research/standards/cdwa/CDWA Lite.html).

Because of differences in the way each museum described its particular painting, it is impossible to find all three works without using multiple terms and undertaking multiple searches. This problem is not one that CDWA Lite or metadata harvesting is designed to correct. Good data is still the museum's responsibility, and the entire community needs to work on consistency in how it describes its objects.

Dr. Murtha Baca, Head of the Getty Vocabulary Program and the Digital Resource Management Department at the Getty Research Institute (GRI, www.getty.edu/research/institute), discussed the larger issue of shareable metadata and CDWA Lite for nonbibliographic items in her talk entitled "Shareable Metadata for Nonbibliographic Materials: Implications for Libraries and Archives." Beginning with an overview of the typology of standards, Baca noted that MARC[21] may not be the most appropriate metadata schema for libraries and archives with nonbibliographic materials such as art objects, architecture, and other cultural works.

Baca outlined emerging trends in metadata creation and spoke about an increasing movement toward "cross-cultural" use of standards in various collecting institutions. She cited as an example The Morgan Library and Museum (www.morganlibrary.org), which concerned that a strict application of the AACR[22] standard would lead to nonsensical displays in its online database, used non-AACR cataloging codes (like CCO and DACS[23]) in its library system.[24]

Baca next discussed the harvesting project undertaken by the GRI's Library Photo Archives, using its image collection of tapestries.[25] This project was similar in overall design to what Erin Coburn discussed earlier in her work at the Getty Museum. However, the GRI made different decisions about the data it contributed: its offered more than "core" information to the harvesting, and they contributed as many images of a tapestry as were available (the museum offered only one image per painting). It also had content issues that the museum did not have. For example, the Photo Study Collections use different, nonstandard, locally developed metadata schemas, and its metadata records are a hybrid of work and image records. In addition, some cataloging decisions that work locally did not translate well in a union environment.

Baca cautioned about the problems that occur when you "dumb down" metadata for harvesting. She advised cataloguers to use the most appropriate schema possible to express their data and to be aware that there is usually some loss of granularity and/or context when mapping to another schema. She also echoed Sarah Shreeves's earlier admonition about considering how information "translates" outside of your database.

Using examples from several well-known cultural collections databases,[26] Baca demonstrated how these problems play out. A search of "theatre" does not identify materials recorded as "theater"; a search of works by the artist "Giambologna" does not retrieve objects catalogued under "Bologne Jean de" (or other variants of his name). The solution to these problems is to include all variants of a name or term, but this is extremely labor-intensive.

---

21 Machine Readable Cataloging (MARC) is a data interchange standard administered by the Library of Congress at www.loc.gov/marc.

22 Anglo American Cataloging Rules, Second Edition (www.aacr2.org/us/products/aacr2.html).

23 *Describing Archives: A Data Content Standard (DACS)* (www.archivists.org/catalog/pubDetail.asp?objectID=1279).

24 *Corsair:* The Online Research Resource of The Pierpont Morgan Library (http://corsair.morganlibrary.org).

25 The Getty Research Institute's Study Images of Tapestries (www.getty.edu/research/conducting_research/digitized_collections/tapestries.html).

26 RLG Cultural Materials (http://culturalmaterials.rlg.org/cmiprod//web/workspace.jsp?MENU=1); Experience Art LACMA: Collections Online (http://collectionsonline.lacma.org/); the National Gallery of Art (www.nga.gov/search/index.shtm), and the Getty Research Institute's Selected Special Collections Finding Aids (http://archives.getty.edu:8082/cgi/f/findaid/findaid-idx?cc=utf8a;c=utf8a;tpl=browse.tpl).

Service providers and aggregators could help by using controlled vocabularies and thesauri as search assistants. They also could add value to their offerings by providing services such as vocabulary mapping, query expansion, vocabulary-assisted searching, user-added metadata, metadata enhancement, etc. But the solutions don't lie solely with service providers. Baca believes data providers should be held to higher standards: Service providers should demand "pre-washed" metadata that uses vocabularies and adheres to data standards. Data providers also need to consistently use appropriate standard schemas in their local systems. Baca cited Roy Tennant's article (suggested by Sarah Shreeves, see page 16) for details on the work needed to make shareable metadata and harvesting more efficient and effective.

Baca concluded by stating that metadata is one of an institution's biggest investments. Creating consistent, standards-based metadata ("also known as cataloging") is onerous, but it opens up a wealth of opportunities for sharing information. "Do it once, do it right," and you will be able to repurpose your metadata in ways that expose your collections to greater audiences and uses. She emphasized that good descriptive metadata records do not need to have a lot of information. "Core" records can be extremely useful if they are done well.

The second half of the workshop focused on how OAI-PMH is being considered by three very different art institutions: the Courtauld Institute of Art, the Metropolitan Museum of Art, and the Princeton University Art Museum.

Günter Waibel, Program Officer in the OCLC Programs and Research division (www.oclc.org/research), introduced this half of the workshop and spoke briefly about how OCLC Programs and Research is working to assist institutions with the more technical aspects of sharing their metadata. It is, for example, working to leverage schema transformation software[27] and provide a tool that will transform data from collections management systems into CDWA Lite XML. It also is working to create an open source museum version of OAICat software for sharing or aggregating data.

RLG Programs, a Division of OCLC, is tackling the problem of collections sharing at the institutional level. It has created a Museum Collection Sharing Working Group to address the human and local cataloging obstacles affecting the harvesting of collections information. The three speakers in this panel are members of this Working Group. Each panelist is planning OAI-PMH implementations at his or her home institution, and their experiences offer insights about local issues encountered with data harvesting.

Barbara Thompson, Witt Librarian at the Courtauld Institute of Art in the UK (www.courtauld.ac.uk), was the first presenter. She discussed various "experiments" that she had conducted mapping material from the Courtauld Gallery and Library Collections to the CDWA Lite XML schema. The Courtauld is a teaching academy with collections of painting and sculpture, prints and drawings, and two photographic libraries. It has many areas where they it pull data, including an online art and architecture Web site (www.artandarchitecture.org.uk), spreadsheets, and images in the Witt Library.

Thompson conducted her mapping exercise using four data sets: uncatalogued materials, two sets of data from the art and architecture Web site, and one set of data from a spreadsheet. She mapped several works by George Frederick Watts and archival documents from the life of Sir Robert Witt to familiarize herself with the CDWA Lite schema, the process of mapping, and the issues that could emerge from using diverse data sets.

Walking the audience through the various steps of her mapping exercise, Thompson identified some of the problems and anomalies she encountered along the way. There were questions about mapping terminologies (particularly how to incorporate terms embedded in strings), differences that emerged when mapping data from images of works versus the works themselves,

27 OCLC Metadata Scheme Transformation Services (www.oclc.org/research/projects/mswitch/1_schematrans.htm).

and identification of some local cataloging that would not translate appropriately in a union environment. Thompson ultimately concluded that CDWA Lite catalogued raw data with a great deal of richness, although she did not find a metadata set that catalogued provenance history to her satisfaction. However, she felt that being able to convert her spreadsheet data to XML offered a unique opportunity for her institution not only to share data via harvesting, but also to pull information together from dispersed sources within a department or across an institution.

Cathryn Goodwin, Collections Data Specialist at the Princeton University Art Museum (www.princetonartmuseum.org), described how she envisions using CDWA Lite for OAI harvesting across her university campus and beyond. Goodwin described Princeton as a decentralized organization with multiple aggregation models for information around campus. The University Library is developing an XML database of digital collections created within the library itself. The university also has created a widely used software application[28] that functions as a database for delivering images and content for teaching, but it is "fed" by content providers and thus duplicates existing data sets and images available elsewhere.

For Goodwin, the key advantage of using CDWA Lite is that campus aggregators can get information from one another, rather than data providers having to feed each aggregator on a one-by-one basis. At the moment, all dialogue at the university about the potential of OAI harvesting is focused on using it within the university itself. However, Goodwin also wants to position the museum so that it can contribute to cross-institutional sharing in the future.

An implementation of CDWA Lite is currently being developed by the university's database applications development office. This office is creating a photo services implementation that works with the museum's collections management system (CMS).

Michael Jenkins, Manager of Met Images at the Metropolitan Museum of Art,[29] made the final presentation, addressing how CDWA Lite was being used at his institution to gather information from dispersed information resources.

Internal collections sharing is a critical but challenging endeavor at the Metropolitan. The museum has 20 implementations of its collections management system,[30] and has used cumbersome software programs to pull data out of these systems when they are needed for publishing or other projects. CDWA Lite immediately appealed to the museum because it offered a schema that made it easier for everyone to contribute information from their systems to joint institutional efforts.

One of these efforts was a new partnership with ARTstor. In March 2007, the museum announced an agreement to make 2000 high resolution images available for ARTstor's digital academic publishing project.[31] Having a method (CDWA Lite) for contributing data from the museum's multiple databases to this project made it more palatable for all involved. ARTstor assisted the Metropolitan in this effort by installing an OAI server at the museum and writing an application that produced CDWA Lite XML records from the museum's numerous databases. The XML records were refreshed nightly, guaranteeing up-to-date information would be available when harvested.

Jenkins feels the collaboration with ARTstor provides a useful model for the community. By providing the museum with the technical utility to offer records for harvesting, ARTstor "lowered the bar" for the museum, making participation much easier. The museum now also uses CDWA Lite internally (without OAI) as a delivery container to publish information. Individual curatorial departments map their collections

28 *Almages*t (www.princeton.edu/~almagest/opensource).

29 Metropolitan Museum of Art Collections On line (www.metmuseum.org/Works_of_Art/collection.asp).

30 Gallery System's "The Museum System (TMS)" (www.gallerysystems.com/products/tms.html).

31 ARTstor's "Images for Academic Publishing" service (www.artstor.org/info/news/iap_announce.jsp).

information to CDWA Lite and then contribute these records for projects within the museum.

Jenkins concluded that CDWA Lite and OAI have improved both the efficiency and efficacy of the museum's collections services. They will be important components of the museum's collection-sharing strategy as it moves into future collections-sharing endeavors.

The summary of this Workshop (#3) was written by Dr. Joyce Ray, Associate Deputy Director for Library Services, IMLS.

This pre-conference workshop provided library and museum professionals with a basic introduction to the world of media production, including:

- needs and expectations of film and television producers, who may incorporate archival film footage into productions;
- distribution of video products via new media channels such as high-bandwidth Internet;
- nuts and bolts of video production including pre-production, lighting, cameras, audio, and post-production processes; and
- preservation of digital video.

Marsha Semmel, Deputy Director for Museums and Director for Strategic Partnerships at IMLS, opened the pre-conference workshop with a panel discussion that included Selma Thomas of Watertown Productions, a video production company, and Jennifer Locke Jones, Chair and Curator, National Museum of American History (NMAH, http://americanhistory.si.edu). Semmel emphasized the need for library and museum professionals to consider exploring the new opportunities offered by the media world in light of 1) their institution's mission and vision; 2) their audiences; 3) their resources; 4) core competencies required, and whether the library or museum should have these on staff or hire outside expertise; and 5) the benefits of collaborations, including partnerships with public broadcasters.

Semmel added that the entire concept of broadcasting is changing radically. In today's YouTube world, everyone is a producer, and everyone can be a curator. This is a world of networks and connectivity, of growing interoperability, where fast response and flexibility in creating, combining, and recombining program elements in different media, by curators and programmers, often in concert with various audience segments, is quickly becoming the norm.

Selma Thomas of Watertown Productions suggested that library and museum professionals think of production as collection curation. Producers recognize that museums, libraries, and public broadcasters hold unique and complementary resources and assets. Cultural heritage institutions also have reputations as trusted institutions, have public service missions, and have expertise with different technologies and media such as historical film. She noted the increasing demand for access to information and for lifelong learning opportunities in today's knowledge society, and the increasing convergence of the museum, library, and broadcasting worlds that is being driven by technology. As she stated, "technology is a moving target," and—for that reason—audience needs and institutional goals should be the primary guides in choosing and producing media-based programs. Thomas stressed that finding archival documentary footage can greatly enhance a production but requires a lot of time and research. Producers create a timeline and script and then search for appropriate footage. They must find not only the appropriate item but also the exact segment they need, so an index is critical. For these reasons, producers tend to use what is easily available and accessible. There is a growing concern among producers and archivists that much information is being "lost" because it is not readily available online.

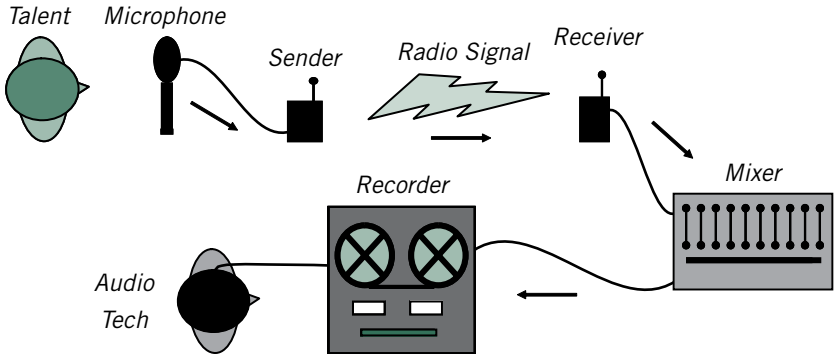
Jennifer Locke Jones of NMAH noted that producers often have very limited time and money to do research in archival sources. Good documentation is essential so that researchers know what is available and can find a particular segment quickly. It is important to have collections in more than one format and to reformat to the current technology for accessibility. The best way to ensure access to film and video collections is to stream video on the institution's Web site, but you must do so according to accepted standards to enable download and use. There is a long-term cost for reformatting that must be considered. However, not all materials are created equal—it may be worth the investment to make some

materials accessible, but not all. Transcripts are expensive, but new tools are being developed that may improve access. Virage (www.virage.com), for example, is a software tool that facilitates the introduction of metadata to indicate changes such as camera angles, but it is expensive and still requires human interpretation. Institutions that cannot or do not want to undertake the costs of converting fragile formats, upgrading to new formats, and creating search mechanisms to help producers find the material they need should make these decisions before they start collecting and perhaps reconsider their acquisitions policies.

Louis Fox, Associate Vice President for Computing and Communications at the University of Washington (www.ischool.washington.edu), spoke about the potential of Internet2 (www.internet2.edu), a nonprofit advanced networking consortium that provides high-speed Internet access to support advanced research and educational applications. The consortium currently comprises more than 200 U.S. universities and partners in 38 states. More than 40,000 public schools, 3,000 libraries, and 125 museums are now connected to Internet2. High-speed Internet can expand distribution of multimedia content beyond broadcasting into webcasting and other emerging distribution channels. Cultural heritage institutions can use this network to engage new audiences that are growing up in the networked world. Internet2 supports multimedia digital programming, real-time discovery-based learning, and emerging learning communities. Libraries and museums can make their collections of film, video, music, and other media accessible to schools and teachers via Internet2. They can also use it to distribute their own productions of one-way or interactive programming and can help a new generation of teachers find and use multimedia resources in the classroom.

Tim Lorang of the ResearchChannel (www.researchchannel.org) provided a whirlwind tour of the video production process, from acquisition or production through editing and postproduction to encoding, storage, and distribution. He emphasized that just because it's digital doesn't mean it's good, and he demonstrated many common mistakes that reduce quality. Lorang noted that quality restricts future use and the value of the message. Quality should be determined by goals, not budget. An institution that doesn't have the resources to meet its goals should change the goals, increase the budget, or not attempt the project. It is easier and more economical to do it right the first time. Some production mistakes can be corrected later, but at a higher cost, while others cannot be corrected at all. In general, it is wise to engage professionals when high quality is important, particularly for

Recording Audio: Multiple Opportunities to Mess Up



Lorang's PowerPoint slide presentation, slide 4.

postproduction and audio. In some cases, in-house filming maymake investing in a camera worthwhile, but frequently this too can be done more efficiently and economically by outsourcing to a company that has a wide range of equipment as well as expertise. Some general production cost estimates are $5,000 for a 30-minute lecture; $20,000–$40,000 for a 15–30-minute documentary; and $5,000–$15,000 for a 30-minute interview. Having a general knowledge of the production process can help libraries and museums work successfully with production companies, whether in partnership with a broadcaster or by hiring them for their own production. If any part of the budget needs to be reduced, it should be at the distribution end rather than the production end, since this can be increased later for a high-quality product.

Nate McQueen of the ResearchChannel closed with a presentation on "The Challenge of Media Management and Content Delivery." He covered issues such as encoding, metadata processing, cataloging, storage, distribution, and content integration, including publishing. The ResearchChannel, a nonprofit organization affiliated with the University of Washington, offers storage and distribution solutions to subscribers through its Digital Well (www.researchchannel.org/tech/digitalwell.asp), which employs open source or open standards tools whenever possible. Encoding and metadata are necessary to identify and retrieve digital video stored in repositories. The Digital Well currently supports only the Dublin Core metadata schema, which requires 10 core fields. In its next version, however, it will support many different schemas, including PB Core, which provides up to 60 fields, and MPEG-7, which is a nested schema providing over 200 fields. It is important to consider scalability in selecting a schema, so that it can continue to serve its purpose effectively when content and use grow—it should be able to expand on demand. Consideration should be given to use of GUIs (globally unique identifiers) to ensure stable location and retrieval over time. It is also desirable to provide metadata that can be harvested via the OAI-PMH in order to reveal institutional assets to search engines that harvest metadata. McQueen discussed some of the common commercial video content providers, such as YouTube, and pointed out that these commercial services have no long-term preservation component. DSpace, developed by the Massachusetts Institute of Technology in cooperation with Hewlett Packard to support institutional repositories, excels at deep preservation but currently does not include digital rights management. Shibboleth (http://shibboleth.internet2.edu) is an open source tool for digital rights management that is becoming widely used. Unfortunately, open source tools are not available for every purpose, and tools have not yet been integrated into comprehensive suites of services available uniformly. The most important thing is to plan for the long term from the outset rather than to opt for short-term solutions that can inhibit preservation and use at a later stage.

Project: Renaissance Society Online Exhibition Archive, 1915–present (see page 81).
Photo: Avery Preesman installation, 2006. The Renaissance Society.

Conference Day One

March 1, 2007

Keynote speaker **Dr. Elizabeth Broun,** Director of the Smithsonian Institution's American Art Museum (SAAM, http://americanart.si.edu), began with a tour de force multimedia presentation showcasing the many resources available on the Museum's Web site. She followed this presentation with an overview of the museum's history and entrée into digital media, beginning in the 1970s with a research database and moving more deeply into the digital realm through the efforts of many talented staff members. When the museum initially ventured onto the Web in the 1990s, the move was undertaken as a way to enhance the museum's brand. Today the Web site is totally integral to the museum. Itno longer separates the "brick-and-mortar" from the "virtual" institution, and it has had to develop new assumptions and business models as a result.

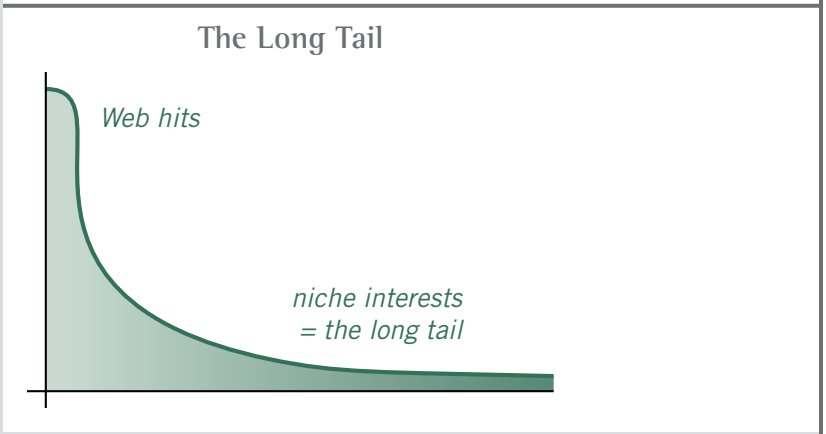
SAAM strives to be *the* place to go for American art. To achieve this goal online, the museum is adopting a "long-tail" strategy for information provision and access. Chris Anderson,[1] editor-in-chief of *Wired* magazine, adopted this phrase to describe an economic phenomenon in the online world where poorly selling products can collectively make up a market share that exceeds the relatively few best-selling products.[2] (The phrase originates in the statistics community, where it is used to describe the sections of a bell curve that "tail off" at the beginning and end of the curve.)

When applied to information access and provision, a long-tail approach is one which identifies the myriad niche interests that exist in the long tail and builds content to meet those interests. The museum's move to a long-tail strategy is supported by analysis of its Web site traffic, which shows that the 10 most frequently used areas of the site generate only 25 percent of the traffic. The remaining 75 percent of traffic comes from less frequently used areas of the site, or the long tail. In other words, the majority of users are not coming to the Web site to see something the museum has created specifically for them, such as a special exhibit, a tour, or some other virtual event. They are coming for assets and content that address an interest of their own.

What interests are driving users to the nonmuseum -mediated portions of the site? SAAM is examining this question closely. The museum has priceless collections, phenomenal staff devoted to research and scholarship, and a reputation as a populist place. It now needs to harness these assets and use them in a way to meet users' needs in the long tail.

Broun cited the museum's new Luce Foundation Center for American Art (http://americanart.si.edu/museum_info/renovation/index.cfm#luce) as a physical manifestation of the long tail. The Luce Center contains more than 3,300 objects arranged in visual storage displays (paintings hung on screens, items in storage drawers, objects arranged on shelves, etc.). The center is extremely popular with visitors, perhaps because it is the antithesis of a traditional museum gallery where curators choose the objects displayed and the method of display. In the Luce Center, visitors choose their own path, not a path created for them. Their interests drive what they choose to view.

To implement a long-tail strategy online, SAAM is changing many of its assumptions about information access and delivery. In the past, for



Broun's PowerPoint slide presentation, slide 95.

1 Anderson, Chris. 2004. "The long tail." *Wired* 12.10 (October 2004). Online at www.wired.com/wired/archive/12.10/tail.html.

2 Amazon.com, for example, may sell three million copies of a Harry Potter book on its release date, but on the same day may collectively sell eight million other books.

example, it published everything online at the *end* of a physical activity: Once an exhibition had been planned and installed, the catalogues written, and everything put through multiple reviews, *then* all the content was placed on the Web site. However, this strategy does not offer what users apparently want. So the museum now is focusing on publishing what Broun termed "microcontent." It publishes information whenever it is available and on many different topics, not just huge swathes of information on a single artist or a single exhibit. For example, last year SAAM published over 110 "small stories" on its blog, which allowed it to address a wealth of different and diverse topics. The goal is to increase the content spread and thus increase the likelihood that it will be offering information that satisfies one of the niche interests hidden in the long tail. This new strategy of publishing more microcontent is scalable and allows SAAM to publish more broadly.

SAAM also wants to be sure that everything it offers online can be easily found. This goal requires better information architecture, search engine technology, and human interface design, but SAAM has an additional problem to address: brand confusion. The museum is associated with a variety of names that reflect its parent organization (the Smithsonian), its museum name and acronym, the name of the complex (the Reynolds Center), the historic name of the building (the Patent Office), its new centers (the Luce Center for American Art and the Lunder Conservation Center), its branch museum at the Renwick, and its close partnership with the National Portrait Gallery. The museum is working to solve this brand confusion, which Broun identified as one of the most difficult roadblocks it faces in communicating with users.

The museum also will strive to put users at the center of everything it does, but this raises large philosophical issues. How do you capture the individual passion of users and harness it? The museum will start by focusing on its core customers, asking them what SAAM can do for them and letting users add their knowledge to the SAAM site (through commentary, tagging, and other yet-to-be discovered means). Broun was emphatic that museums can no longer be "timeless oracles sitting on the sidelines of the future" and must engage their constituency in this manner.

To accomplish all its goals, SAAM must increase capacity. The museum has reached the end of its "build it on a shoestring" strategy when it comes to its Web site. Like the renovation of its building, the renovation of the virtual museum is going to require money, and lots of it. Although it needs to "widen the pipes," fortunately it won't have to start from the beginning. It will build on its strengths, adapt things in a more intelligent manner, build value incrementally, "share its passion," and make more information, insights, and excitement flow between the museum and the public.

In the question and answer session following her keynote, Broun acknowledged that the museum will be giving up some control over its content to engage new audiences in the ways she envisions. Although this initially caused some "teeth gritting" on her part, she finds the situation akin to what occurs with scholarly books. Over time, all scholarly publications and catalogues are found to have errors or outdated perspectives. People accept this fact, and new scholarship and information is posited in response. When this happens, the provenance of the new scholarship and information is transparent. The trick in the online world is to make certain the same level of transparency remains, to be frank about who says what about various content.

In response to a question about how a Web site can be used to generate income, Broun strongly voiced her belief that museum Web sites are not money-making endeavors. She urged museums to fundraise directly for their online needs with as much enthusiasm and focus as they do for their buildings, exhibits, and other activities. Advising the audience that "people give to vision, not to need," she urged museums to develop and articulate a vision for their Web sites so they can successfully fund their development.

Moderator **Günter Waibel** characterized the day's first session, which focused on the challenge of preserving physical collections, as one that "eases us into the conference by being mindful of the physical collections that we care for" and helps lay the foundation for understanding the relationship between caring for physical collections and our digitization activities. Panelists discussed the crisis in cultural heritage preservation and ways to deal with this crisis from a digital preservation perspective.

Kristen Overbeck Laise, Vice President, Collections Care Program, for Heritage Preservation (www.heritagepreservation.org), set the stage by discussing the Heritage Health Index (HHI),[3] the first comprehensive survey of the condition and preservation needs of U.S. collections. HHI was developed in consultation with 35 national agencies and 100 leading collections and preservation professionals. Thousands of small institutions participated, as did over 500 nationally significant institutions (such as the Smithsonian Institution and the National Archives). Together, they represent a full array of collections (from artworks to zoological specimens) stored in every conceivable type of collecting institution (museums, libraries, arboreta, archives, zoos, etc.).

The HHI survey reveals some staggering findings. It conservatively estimates that over 4.8 billion artifacts are held in U.S. repositories, and many of these materials are in an unknown state of preservation. (Laise noted that collections whose preservation status is "unknown" usually *do* need preservation.) A significant percentage of collections (26 percent for collecting institutions; 40 percent for libraries) have no environmental controls whatsoever. Large percentages have had their collection damaged by light (59 percent) and/or improper storage (65 percent). A vast number of institutions (79 percent of historical societies, 74 percent of museums, and 43 percent of archives) have *none* of their collections information online.

Those institutions that do have digital collections report equally sobering statistics. HHI found that approximately nine million digital items (defined as the physical media—i.e., the CDs, floppy drives, magnetic tape, etc.—on which digital information resides) exist in collecting institutions. Thirty-two percent of institutions reported damage to these collections because of equipment obsolescence. Forty-nine percent of these digital media are in an unknown state of preservation. Most surprising is that only 31 percent of institutions felt they had a *responsibility* to preserve digital collections as part of their preservation program or mission.

Based on its findings, HHI strongly recommends the following measures be taken to address the preservation crisis in our cultural heritage collections:

- Institutions must give priority to providing safe conditions for collections they hold in trust.
- Every collecting institution must develop an emergency plan for its collections.
- Every collecting institution must assign responsibility for caring for its collections to members of its staff.
- Individuals at all levels of government and in the private sector must assume responsibility for providing support that will allow collections to survive.

Laise noted that conservation and preservation are ongoing activities. They cannot be considered "deferred maintenance." Yet 68 percent of institutions report conservation/preservation budgets of less than $3,000 per year—a figure that is clearly too low to sustain ongoing preservation work. HHI recommends institutions seek endowed funds to support preservation activities and to ensure that they are continuously applied.

The Heritage Health Index has received an extraordinary amount of national media and press coverage, which helps put the problems in the spotlight and generates support. However,

3 *A Public Trust at Risk: The Heritage Health Index Report on the State of America's Collections.* December 2005 (www.heritagepreservation.org/HHI).

Heritage Preservation hopes the report also will be used by institutions to generate internal and local support for preserving their collections. It urges institutions to share the report with board members, fundraisers, press, and other influential people to highlight their local situation and validate local needs.

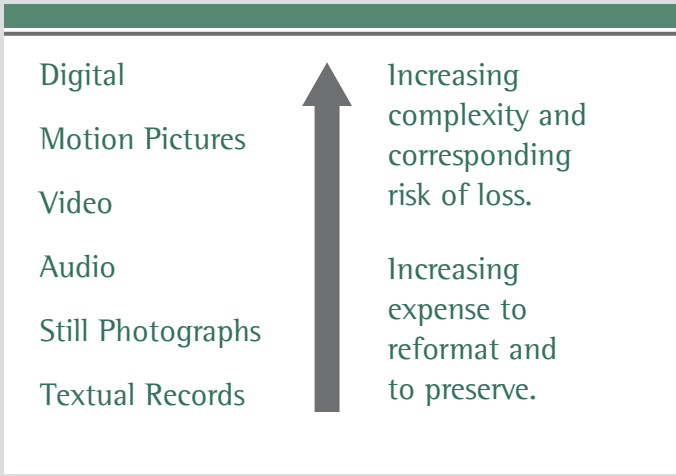
Laise commented that few institutions publicize their preservation needs to their communities. The problems are rarely presented on institutional Web sites, in newsletters, or as the subject of exhibits. Yet preservation has great public appeal, and collecting institutions should take advantage of this fact. She concluded that institutions must bring their local preservation problems to light if we are to have a national impact on the preservation crisis facing our cultural heritage collections.

The next speaker, **Steve Puglia,**[4] Preservation and Imaging Specialist at the National Archives and Records Administration (NARA, www.archives.gov), discussed what can be done from the perspective of digital preservation. In his talk entitled "Overview of Preservation in the Digital Age," Puglia emphasized that technology is just one tool in a cache of tools that can be used for preservation. The key is to select and use the appropriate tool wisely. We need to be critical observers of technology, and to honestly evaluate its strengths and weaknesses.

From a preservation perspective, collecting institutions are responsible for preserving three broad categories of materials:

- Static analog originals, such as print publications, photographs, microfilm, etc. With these types of materials, chemical and physical instability cause degradation and limit usable life.
- Dynamic analog media, such as motion pictures, audio and video recordings, and any other analog item that needs a machine to access information from the item. For these types of materials, chemical and physical instability of the media and system obsolescence limit usable life.
- Digital media. System obsolescence is the critical factor for these materials, but chemical and physical instability can also limit usable life.

When assessing preservation needs it is critical to take a systems perspective and to acknowledge that managing and preserving digital data and objects is different from managing and preserving analog ones. As you move along the preservation spectrum from static analog to dynamic analog to digital media, there is increasing complexity, a corresponding risk of loss, and an increasing expense to preserve these materials.



Puglia's PowerPoint slide presentation, slide 7.

When considering preservation strategies, there are a number of tools and methodologies to choose from: proper storage, environmental monitoring, holdings maintenance and conservation, risk condition and archival/curatorial assessment, reformatting, etc. For example, extremely cold storage is the only tool to use to permanently preserve still and motion images, and nitrate and

4 Henry Wilhelm, President of Wilhelm Imaging Research (www.wilhelm-research.com), was originally scheduled to speak in this time slot about "Sub-Zero Cold Storage for the Permanent Preservation of Photographs, Digital Prints, Newspapers, Magazine, Books, Paper Manuscripts and Other Artifacts that are not Born Digital." Circumstances prevented Wilhelm from attending, and Steven Puglia offered to speak in his stead.

other photographic film. The colder the storage, the more longevity you can expect for these materials.[5]

Reformatting is always the most expensive tool in the preservation arsenal, but it is the primary method used to preserve materials in system-dependent formats such as audio or video recordings. It also is essential for allowing access to originals that must be preserved in cold storage or are otherwise unavailable. When considering reformatting as an option, you must take into account archival needs, costs, reproduction quality, stability of imaging materials, and ease of distribution.

The paradox of preservation is that people want an archival medium that can maintain information forever, but no such medium exists. So instead, preservationists talk about "life expectancy,"[6] which Puglia defined as "the length of time that information is predicted to be acceptable in a system at 21 degrees C and 50 percent relative humidity," which are the average conditions of an office environment. Life expectancy is greatly shortened as we move toward digital. For example, in an unmanaged environment, digital media have only 1/100th the life expectancy of the best photographic film available today (5 years vs. 500 years). This sobering fact has everyone working on ways to preserve digital information.

Despite the low life expectancy for digital media, the use of digitization for preservation reformatting is inevitable because analog processes and materials are disappearing from the marketplace. (Audio recordings are now being reformatted solely by digital means. Microfilm will be next, as all predications indicate it will be phased out over the next five years.) In addition, digitization is the only method that facilitates access to objects which cannot be physically handled.

As we rely more on digitization for preservation reformatting, we will need to move away from managing and preserving technology and media to managing and preserving digital objects and data. While much work is being done in this area, we still need to understand more about the characteristics and quality of the digital objects we are creating (e.g., capture device performance, bit depth, resolution, etc.). We also need to improve our risk assessment and work to identify and choose appropriate tools for specific preservation needs.

Puglia suggested promulgating the "lots of copies keep stuff safe" idea. We want to try to keep the originals in their original formats whenever possible, but we need to digitize them for access and for the ability to create multiple copies. All institutions need to use a holistic approach to preservation that is incremental and that is not seen as an "either digital or original" argument. We also should encourage smaller institutions to partner with larger groups to undertake digital preservation activities. Consortial approaches to the problem are the only economically viable option because institutions cannot invest in preservation infrastructure on their own.

Laise's report on the state of cultural heritage collections and Puglia's presentation on preservation strategies provided background for the next discussion—a case study of preservation at Exit Art (www.exitart.org/site/pub/exit_archive/digital_archive.html), an organization that presents conceptual and experimental artwork.

Jodi Hanel, Associate Curator at Exit Art, provided background on the organization, noting that a critical component of its mission is to keep a comprehensive exhibition record and to thoroughly document the organization's programs. Exit Art has presented the work of over 2,500 artists, and provides primary materials about American contemporary art. The organization's founders had the foresight to realize their institution's larger role in the world of contemporary art, so they saved enormous amounts of material during Exit Art's 25-year history.

5 In his discussion of cold storage as a preservation tool, Puglia used Henry Wilhelm's slides as illustration. See Puglia's PowerPoint slide presentation (slides 9 through 42) at www.imls.gov/news/events/webwise07.shtm

6 Exit Art (www.exitart.org/site/pub/main).

Audrey Christensen, Exit Art's archivist, spoke about the Exit Art Digital Archive project, which is creating digital surrogates to preserve and make accessible artwork that can be ephemeral (such as performance art, multimedia, conceptual art) and whose existence is often represented solely by archival materials. The project began two years ago with the chronological organization of materials and media into six distinct series: curatorial files, photographic collections, slides, video/audio files, and born-digital collections. Exit Art is digitizing documents, photographs, drawings, oral history, and other materials that represent the artwork and how it is presented, as well as the organization's institutional history. The goal of the Digital Archive project is threefold: to capture and preserve the trove of information in the Exit Art archives, to make it available to everyone, and to assist colleague organizations that wish to develop digital archives of their contemporary art holdings by sharing the database, their digitization plan, and their metadata schema. The project will also serve as a test bed for the use of descriptive standards being developed for conceptual and intermedia art. When the archive goes online (in the summer of 2007) it will be an open access, interoperable, searchable database of Exit Art's history and the contemporary art it has showcased.

In introducing this session, **Thomas Clareson,** Program Director for New Initiatives, PALINET, Inc. (www.palinet.org), noted that past WebWise conferences often showcased new standards and initiatives which, in subsequent conferences, are reintroduced in the context of actual implementations. The presenters in this panel provide examples of this continuity. They have put theories and ideas expounded only a few years ago into practice today.

The first speaker, **Ann Russell,** Executive Director of the Northeast Document Conservation Center (NEDCC, www.nedcc.org), outlined the results of an IMLS-funded project to survey the digital readiness of cultural organizations. For over 30 years, NEDCC has been surveying traditional preservation needs in small to midsized institutions. As more digital materials enter collections, it has become apparent that a new set of tools and skills is needed for preserving these special materials. Working with other professional organizations,[7] NEDCC brought together digital experts and preservation surveyors to develop a method to assess and address digital preservation issues in cultural organizations.

They began by gathering quantitative data on the status of digital collections in cultural institutions and found that 92 percent of these institutions were digitizing from source materials, but only 29 percent had any written policies or plans for digitization. The disparity in these percentages confirmed NEDCC's suspicions that institutions were creating digital materials without planning for their preservation, and that cultural institutions needed assistance integrating digital activities into their broader strategic planning.

NEDCC next convened a group of experts to review the survey data and develop a strategy that could help organizations conduct assessments of their digital preservation needs. The experts concluded that cultural heritage institutions need a national technical assistance program on digital readiness. This program should include expert-facilitated, on-site surveys that address the goals and resources of an institution. After analyzing several models, it found the Conservation Assessment Program (CAP, www.heritagepreservation.org/CAP) to be the most promising. CAP enables small museums to bring conservation consultants into their institutions to perform general surveys and identify conservation priorities. A similar model for digital preservation assessment was suggested.

NEDCC conducted test bed site visits to develop survey tools and procedures that followed the CAP model. Eight institutions (five museums and three libraries of different sizes and types) were chosen for the test bed. Each institution filled out a survey about its organization and its digital activities. A team of two digital preservation experts ("surveyors") then conducted a site visit at the institutions, where they interviewed key personnel, visited processing and digital storage areas, and gathered detailed information. The surveyors wrote a report documenting existing practices and policies, and made institution-specific recommendations for improvement.

NEDCC staff and consultants now are creating a series of written tools (www.nedcc.org/resources/toolkits.php) for cultural institutions to use to conduct digital preservation assessments. They are refining the institutional questionnaire, writing a surveyor's handbook (which includes a sample survey report), and reporting on the quantitative data they have gathered and the qualitative trends they have observed. NEDCC also will survey large institutions with good digital preservation programs so it can present examples of leading-edge work and identify institutional models for cultural organizations to consider.

For NEDCC, the larger goal of the project is to contribute to a national strategy for digital preservation. Russell concluded by identifying the elements needed in such a strategy:

7 Other partners in the project were the Museum Computer Network, the Center for Research Libraries, the American Institution for Conservation, and Heritage Preservation.

- Statewide solutions for long-term storage that are open to multiple institutions and are backed up by regional centers of digital preservation expertise
- Training in digital readiness and preservation
- A standards-based approach to trusted digital repositories
- A derivative version of the NARA-RLG audit and certification checklist[8] that could serve as a best practice guide for smaller institutions
- A national technical assistance program for digital preservation (based on the CAP model)
- Advocacy efforts to encourage institutions to make digital preservation a part of their mission

Robin Dale, Program Officer for RLG Programs (a program of OCLC's Office of Programs and Research), followed up on one of these elements in her talk on **"**Auditing and Certification of Trusted Digital Repositories: Best Practices and Next Steps." Dale began her presentation by discussing how the concept of trusted digital repositories is applicable to everyone because they are part of a digital readiness plan. All collecting institutions need to understand risk, content, context, capabilities, and solutions for preserving their digital information.

She noted that "confusion reigns" when it comes to the terminology used for digital repositories. Is a digital repository the same as a data archive? Are content management systems repositories? Is an institutional repository the same as a digital repository? We all have notions about some de facto type of repository involving digital preservation and access, but there is no "one size fits all" nor any "out of the box" solution.

For this reason, we need to agree on principles and an objective way to assess, and ultimately certify, digital repositories. The concept of auditing is key: Audits imply checking, examining, making certain something meets set criteria. Audits of digital repositories make us think about the capabilities of the people and systems running a repository. They allow us to assess our risks and, most important, they provide the basis for trust.

We need *trusted* digital repositories to be sure that our digital collections are going to survive and be available for use and reuse in different ways over time. Trust involves verification and clarification about activities and actions ("trust, but verify and clarify"). It's also an iterative process: Repositories will change with new collections, hardware/software changes, staff turnover, etc., so the evaluation process must be ongoing.

While repositories will differ because of local needs, good practice remains universal. For this reason, audit checklists are not prescriptive (i.e., they do not tell you what repository type is best). They *do* provide guidance, allow you to assess the level of risk, and help you build trust with a repository partner. Since most of us will be partnering with repositories, we need to know and trust our partners. Audit checklists are simply a tool to help us in this effort.

Dale summarized some of the high-profile projects that are creating auditing guidelines for digital repositories. The RLG-NARA Digital Repository Certification Task Force (www.rlg.org/en/page.php?Page_ID=367) recently has combined its efforts with the work of the Center for Research Libraries (CRL) Audit and Certification of Digital Archives (www.crl.edu/content.asp?l1=13&l2=58&l3=142) project to create a revised tool known as *TRAC - Trusted Repositories Audit and Certification: Criteria and Checklist* (www.crl.edu/content.asp?l1=13&l2=58&l3=162&l4=91). TRAC is organized around three areas: organizational infrastructure, digital object management, and technical infrastructure and security. It builds on earlier work and identifies additional requirements for specialized repositories. The Center for Research Libraries will maintain this document.

8 Newly revised and now published under the title *TRAC - Trusted Repositories Audit and Certification: Criteria and Checklist* at www.crl.edu/content.asp?l1=13&l2=58&l3=162&l4=91.

On the international scene, efforts include *nestor*,[9] a German-based working group developing guidelines for state, academic, and archival communities to meet certification requirements being proposed by the German government.[10] This program is developing training tools that will be available in English. The Digital Curation Centre (DCC), in association with Digital Preservation Europe, is developing a risk-based self-assessment tool kit (www.repositoryaudit.eu) to help institutions conduct internal audits/assessments of their own digital repositories and measure strengths, weaknesses, and capabilities. And CRL, nestor, and DCC are creating a list of "10 principles" for digital preservation repositories, which Dale characterized as a "Cliff Notes" version of digital repository requirements. Together, these local and international efforts will help institutions evaluate their own digital repositories or the digital repositories run by potential partners.

Dr. Sue Medina, Director of the Network of Alabama Academic Libraries, spoke next about the "Alabama Digital Preservation Network: A Statewide Solution to Preserving Locally Created Digital Collections." This IMLS-funded project is exploring a scalable model of digital preservation archiving that allows institutions of all sizes and types to participate with minimal effort.

The Alabama Digital Preservation Network (ADPNet, http://adpn.org/wiki/Main_Page) began as an adjunct to another IMLS-funded project—AlabamaMosaic (www.alabamamosaic.org)—which created a digital collection showcasing Alabama resources. To participate in AlabamaMosaic, local institutions had to agree to archive their own digital content. This agreement proved difficult for many participants. Small institutions frequently didn't understand the concept of digital archiving and when they did understand it, they often couldn't implement it. Larger institutions also were not very responsive to this requirement. The project leaders soon realized they needed to provide some archiving solution for their partners.

A project known as the MetaArchive of Southern Digital Culture (www.metaarchive.org) offered one possible archiving model. MetaArchive uses LOCKSS ("lots of copies keep stuff safe," www.lockss.org/lockss/About_LOCKSS) software to preserve at-risk collections that document the culture and history of the American South. The AlabamaMosaic team applied to IMLS for funding to create an archival piece to their project based on LOCKSS. The proposal was funded, and ADPNet was born.

The goal of ADPNet is to offer a trusted archival storage service for cultural organizations in Alabama. The only requirement for participation is that organizations must make their digital collections freely available to the public. The archive is "dark" and thus reserved for emergency use only.

Medina gave a brief history of LOCKSS, noting that it was developed as a response to librarians' concerns that their licensed commercial content, particularly journals, was not being archived by the publishers. Worried that this omission might one day lead to the disappearance of this content, librarians wanted to ensure perpetual access to their licensed resources.

LOCKSS is open source software that provides an easy and inexpensive method for preserving and providing access to digital content. Servers communicate with each other in a slow Web crawl to say the machine equivalent of "This is what I have. Does it match what you have?" LOCKSS boxes constantly monitor themselves to make certain the integrity of their content is maintained.

ADPNet is testing LOCKSS with several partners, including the state's three largest academic institutions, the state archives, a small public

---

9 The Network of Expertise in Longterm Storage of Digital Resources (nestor, www.langzeitarchivierung.de). For an English article on nestor, see Dobratz, Susanne, and Astrid Schoger. "Digital Repository Certification: A Report from Germany." RLG *DigiNews,* October 15, 2005. (www.rlg.org/en/page.php?Page_ID=20793&Printable=1&Article_ID=1779).

10 nestor Working Group. *Catalogue of Criteria for Trusted Digital Repositories.* Version1 (draft for public comment) December 2006 (http://edoc.hu-berlin.de/series/nestor-materialien/8en/PDF/8en.pdf.pdf).

university, and a small private college. Each partner brings different challenges to the table, which is important for determining the scalability and usefulness of the LOCKSS model.

The project currently faces some administrative hurdles. A long-term, sustainable governance model is needed, policies and procedures for the network must be developed, and future funding must be procured. Medina and project manager Aaron Trehub acknowledged that they will have to address format migration and equipment obsolescence issues at some point in the near future. However, they feel that it will be easier to do this collectively: Everyone in the archive will have to get their data out and reformat them, so they will need to agree on a single acceptable solution rather than dozens of different solutions.

Despite the challenges, Medina feels the LOCKSS approach to digital preservation archiving is a good alternative for cultural resources in Alabama. It offers low-barrier entry to organizations with varying levels of digital preservation "know-how." A low-cost, low-maintenance, distributed solution is critical for a state like Alabama, which has few resources. (Eventually, ADPNet will move to a fee-based service model, so it is important that the infrastructure costs are minimal to keep fees low.) Although ADPNet is a "proof of concept" project, it is planning a path toward long-term sustainability.

Moderator **Dr. Clifford Lynch,** Executive Director of the Coalition for Networked Information (CNI, www.cni.org), introduced the four speakers on this panel as representatives of some of the most active players in the arena. All of them are working with one another, and with other agencies, to address digital preservation challenges, and their partnerships signal a critical change in the field. "Stovepipe thinking" about digital preservation is rapidly being replaced by collaborative efforts.

Lynch noted another significant change is taking place in the area of strategy. The pursuit of a purely technological solution for digital preservation is a dead end. Rather, the answers lie in designing organizations and incentives, and allocating shared responsibilities among organizations to carry out the job of preserving, curating, and stewarding collections on behalf of society.

The first panel speaker was **Laura E. Campbell,** Associate Librarian for Strategic Initiatives at the Library of Congress. Campbell gave an update on the National Digital Information Infrastructure and Preservation Program (NDIIPP, www.digitalpreservation.gov), a 100-million-dollar initiative funded by Congress in December of 2000 to develop a national strategy for preserving the ever-growing amounts of digital content, particularly materials created solely in digital formats. Using an initial $5 million, the Library of Congress began the initiative by proposing a strategy based on a distributed group of collaborating partners working to build an interoperable preservation network. When this strategy was approved by various congressional committees, the program received another $20 million to spend on developing the model and testing the network. The remaining $75 million was allocated with the condition that it be matched dollar for dollar. Campbell was delighted to announce that this matching requirement had been met.

NDIIPP has two major components: a preservation network consisting of partners with collections preservation responsibilities, and a preservation architecture consisting of tools, services, storage, etc. The former is an "organization of organizations," with different layers of the network representing different communities. The first layer can be termed a "foundation layer" consisting of the collecting institutions themselves. A second layer includes communities of experts (such as professional associations, standards bodies, and government agencies) whose expertise is critically important to the network, but who do not collect materials themselves. The third layer is the providers and suppliers of services needed by the partners, for example, suppliers of metadata services or storage services. Finally, there is the capacity-building layer, consisting of agencies, individuals, and organizations that provide funds for research and development, education, demonstration projects, etc.

Currently, 40 collecting and preserving institutions are involved in the project, and they are saving everything from social science data sets to Web sites to cultural heritage material. They all have mounted digital repositories and are in full operating mode, dealing with issues of interoperability, migration, and copyright—the full range of problems that one would expect. NDIIPP also has 10 partners involved in digital preservation research (funded with the assistance of the National Science Foundation), partners providing fee-based services to the overall network, and international partners in larger communities of practice. In total, NDIIPP has 67 agreements with institutions and groups actively working at different levels in the network.

As NDIIPP moves into its second phase, it will focus on defining the long-term roles and responsibilities of each institution, thus building out the long-term stewardship network. It also will invest in more capacity-building endeavors, such as multistate demonstration projects for repository development that bring several states together to work on one repository that can

serve them all. And it hopes to add technology companies and commercial content players to bring the services of these organizations in to strengthen the preservation network.

Campbell emphasized that material being saved under the NDIIPP initiative would truly be lost if it were not for the work of this project. The partners in this program are collecting at-risk digital content. She praised them as the real "heavy lifters" whose efforts at collaboration are responsible for the successes achieved thus far.

Dr. Kenneth Thibodeau, Director of NARA's Electronic Records Archive (ERA, www.archives.gov/era) spoke about his agency's efforts to develop a preservation framework for preserving electronic records of the federal government. NARA has an enormous responsibility to shoulder. It must be able to preserve any type of electronic media created on any computing platform from anywhere in government. It also must provide discovery and delivery of these records to all those with an interest and a legal right to access them. Furthermore, it must undertake these efforts for "the life of the Republic."

Thibodeau delved into the formidable challenges NARA faces, such as obsolescence issues, the variety and complexity of digital materials, a time frame that stretches forever, the need to ensure authenticity of records, and enormous volume (the scope of the entire federal government). The latter is especially daunting. Thibodeau cited examples of the vast amounts of material anticipated from just a few federal departments: 25 million State Department diplomatic messages, 32 million e-mail messages from the Clinton White House, and military personnel records totaling 1 billion images. Although no one knows the exact volume of digital information created by the federal government, the Congressional Research Service estimated (15 years ago) that more than 96 percent of federal information starts its life in a computer.

To fulfill its preservation mandate, NARA is building a system that Thibodeau referred to as a "preservation framework." Developed by Lockheed Martin, this framework is hardware/software independent, extensible, and scalable (so that it can handle a collection of one billion records, as well as a small specialized collection that must be isolated from others). The system allows NARA to identify *how* to preserve any specific body of electronic records, and then identify and bring in different tools appropriate for that line of preservation. The ERA system is being developed in five stages, with the first stage (ingesting and providing managed storage for electronic records) scheduled for testing by NARA staff and select government agencies this September. A public release is anticipated a year later, and the entire system is scheduled for completion in 2012.

The new ERA system will offer innumerable benefits to NARA, the federal government, and US citizens. There will be easier access to archives records and services, one-stop shopping in an "e-government context," automated finding aids, tools for managing the life cycle of government records, and measures that will ensure privacy and confidentiality mandated by law. This system is mission critical for NARA. Without it, NARA will fail to discharge the responsibilities vested in them by law.

As a national archive, NARA must preserve records in their authentic form. It cannot repurpose them for the needs of the many local communities that want access to federal government information. Thibodeau believes that this is where digital libraries will play an important role. He foresees these libraries serving as brokers that access NARA information and provide it to various user communities in the ways they need it.

Dr. Chris Greer, Senior Advisor for Digital Data, Office of Cyberinfrastructure (www.nsf.gov/dir/index.jsp?org=OCI), National Science Foundation (NSF), spoke next about his agency's current efforts. He began by expressing his sense that we are at a historic

juncture that rivals the introduction of the Internet, and he offered a brief history of how we have reached this juncture and where it may lead us.

In the mid-1980s, NSF created NSFNet, a networking backbone that initially provided academic institutions with connectivity to new supercomputer centers. NSF greatly underestimated the demand for this connectivity. Open access and (with the introduction of browser software) ease of use changed the entire dynamic, and demand soared far above predictions. The result was transformative. The Internet offered spaces for innovation, discovery, collaboration, and partnerships. Completely unexpected entities emerged, such as Google and its business model built on information aggregation.

Greer believes we are entering another transformative period, but this time the driving force is our increasing ability to transparently access and interact with divergent and heterogeneous data sets. He predicts that the next "Mosaic"[11] will be a visual, interactive, intuitive information integrator and navigator, and the next Google will be a company that provides easy-to-use information integration services. These and other emerging trends will be built on the foundations of the data capacity, preservation, and access capabilities that we are all working on today.

NSF's effort to help build this foundation is outlined in its new cyberinfrastructure plan.[12] Chapter 3 of this plan describes NSF's objectives for data, data analysis, and visualization, and expresses a vision where information is deposited in well-documented form, is easily discovered and understood by specialists and nonspecialists, and is properly protected and reliably preserved. To move this vision into reality, NSF is working on two fronts: 1) mobilizing national organizations that can provide for this functionality, and 2) supporting the research and development necessary to create the tools and other solutions to make this vision possible. Three new agency programs are being developed to support these efforts: a software development program (currently in place), a community-based interoperability program (to be announced shortly), and funding for digital preservation and access organizations (to be announced over the next several months).

NSF also is working with other federal agencies through its National Science and Technology Council's Committee on Science (www.ostp.gov/nstc). Under this committee's auspices, 22 federal agencies and seven offices and councils are creating a strategic plan for scientific data in the federal (and federally funded research) realm. A draft of this plan will be available in six to nine months, and an implementation plan will follow twelve months thereafter.

In closing his address, Greer noted that when he speaks with colleagues who were active during the early development of the Internet, they often talk about the excitement of those times. Ten years from now, he believes we will look back on today's efforts with a similar sense of excitement as we create what he termed a "dataverse" to go alongside the Internet as driving forces in society.

Dr. Joyce Ray, Associate Deputy Director for Library Services at IMLS, spoke next about her agency's efforts in the digital preservation arena. In 1998, when IMLS began funding the creation of digital content in libraries and museums, the agency immediately recognized that it had a responsibility to fund research and

11 Mosaic was the first popular Web browser, and is largely credited with moving the Internet from a network used by scientists to one used by a broader public. It was created at the National Center for Supercomputing Applications in 1992, but has not be developed nor supported since 1997. See http://en.wikipedia.org/wiki/Mosaic_(web_browser) for more information.

12 National Science Foundation Cyberstructure Council. March 2007. *Cyberstructure Vision for the 21st Century (*www.nsf.gov/pubs/2007/nsf0728/index.jsp).

development efforts that would help preserve this content. It continues to fund preservation research because it remains an important issue for their library and museum constituency.

Ray outlined several significant, high-profile preservation projects that have been funded by IMLS. In 2002, it funded a model preservation program for the California Digital Library's multi-institutional digital assets that resulted in one of the first digital preservation repositories. A study of the repository's operations and policies was an important part of the project, highlighting early on that solutions to digital preservation have an administrative, as well as technological, component. In the same year, IMLS also funded the development of DAITSS ("Dark Archive in the Sunshine State") open source software at the Florida Center for Library Automation. This project was an early instance of a **dark archive** prototype.

In 2005–2006, three significant preservation related projects were funded. The first was an award to the University of Michigan to study institutional repositories in North America and provide case studies that showcase the key elements of a successful repository. Another award was made to Johns Hopkins University, in collaboration with the University of Washington and the University of Edinburgh, to investigate the digital archiving issues involved in providing astronomers long-term, reliable access to data for publishing research. In the third project, the Massachusetts Institute of Technology is studying the role of digital preservation archives in developing strategies for preserving computer-assisted design (CAD) documents (using architect Frank Gehry's designs as a test bed).

Ray believes that digital preservation has risen in prominence over the last few years as more digital content is created, more storage is needed, and perhaps because our communities are feeling more confident that solutions are emerging. IMLS will continue its important activity supporting new efforts in this field. In its 2007 National Leadership Grant program guidelines, IMLS is looking for projects in the Research and Demonstration category (www.imls.gov/applicants/grants/nationalLeadership.shtm) that investigate ways to enhance archiving, preservation, management, discovery, and use of digital assets and resources.

A question and answer period followed the agency updates. One audience member asked the panelists how their agencies were addressing the roadblocks that intellectual property issues pose for digital preservation. Laura Campbell spoke about the Library of Congress's Section 108[13] Working Group (sponsored by the Copyright Office and the NDIIPP program) and its efforts to rewrite the portion of U.S. copyright law that addresses preservation for libraries and archives. She noted that change will be slow and incremental, but efforts must continue in this area or else we will never make progress in moving this outdated portion of the law into a new arena.

Ken Thibodeau said that he initially believed that NARA could skirt this issue entirely, since (with few exceptions) federal government records are in the public domain. He faced a rude awakening with the introduction of digital rights management technologies that allow users to place self-destruct mechanisms inside the digital content they create. These technologies would allow anyone creating a government document to mark it for self-destruction, thereby preventing NARA from preserving the record. NARA is now actively involved in research that will allow users to access content independently of whatever software was used to create that content.[14]

---

13 "Section 108" refers to a clause in U.S. copyright law (U.S.C. Title 17, *Copyright Law of the United States of America*—see www.copyright.gov/title17/circ92.pdf). It outlines the rights that libraries and archives have to reproduce copyrighted works for preservation purposes, and the circumstances in which they may do so. Given the technological transformations of the last decade, Section 108 is widely perceived to be inadequate, particularly by those responsible for digital preservation.

14 See the Multivalent Document Home Page at www.archives.gov/era/related-websites.html?template=print.

Chris Greer mentioned that NSF's cyberinfrastructure report urged the scientific community to address more proactively NSF's assertion that research products funded by the agency should be made openly available. Joyce Ray responded that IMLS could fund research on the impact of current copying restrictions.

Moderator Cliff Lynch noted that the efforts of these agencies suggest there is a new philosophical perspective emerging to address the legal roadblocks and impediments to preservation. This perspective moves the discussion from the narrow realm of "who has legal control over what" into a larger domain that posits reasons why we collectively need to preserve and have access to our cultural heritage. With their efforts, the federal agencies are working to get us to consider a broader notion of stewardship.

This session shifted focus from preserving the digital by-products of cultural heritage to preserving cultural heritage itself, especially the amorphous, intangible aspects of culture —language, oral history, cultural memory—that are difficult to capture by traditional means.

Session moderator **Liz Bishoff,** Head of Sponsored Programs at the University of Colorado at Boulder, noted that documenting cultural heritage is a distinctive type of preservation activity, and the panelists in this session would speak about the issues unique to this endeavor.

Jane Sledge, Associate Director for Museum Assets and Operations at the Smithsonian's National Museum of the American Indian (NMAI, www.nmai.si.edu), spoke about museum recording systems as stewards of knowledge. These systems are not authoritative as is frequently assumed, but instead represent multiple points of view and differing opinions. They are, in a very real sense, a repository for cultural conversations. This perception contrasts markedly with traditional views that interpret these systems as nothing more than sophisticated finding aids.

Sledge illustrated the complexity of information in the NMAI's recording systems with an example of a Potawatami woman's blouse, and showed how its presentation and cataloging in the museum represented both Western terminology and perceptions of use. Insights by Native American staff and other community members added a Native understanding that recontextualized the meaning and use of the blouse. How can this new information and insight, which is often captured and transmitted better through storytelling mechanisms than by data fields, be recorded in an information system?

The answer is that it can't, unless the system has a place to record this information. For this reason, NMAI is working on a robust system architecture that will hold the large amounts of information it expects will be provided by its Native American constituency. Sledge emphasized that most of the Museum's objects do not come with a title and author/creator. The objects must be recontextualized—described and given meaning—by Native Americans themselves. To ensure that opportunities for recontextualization take place, the museum is changing the nature of their documentation practices by fostering collaborations between museum staff and Native groups, and intergenerationally between elders and tribal youth.

Much of the earliest information on NMAI's collections comes from George Gustav Heye, who collected the majority of the museum's holdings. Heye purchased entire rooms full of material from Native Americans, and recorded their origins as the place of purchase. He had no concept of provenance, of how things are traded, inherited, and dispersed through time. In its new collections management system, the museum now has the ability to record these concepts as they apply to a particular object or group of objects. The system can include information about the culture that *created* the object and the culture that *used* it. It also can offer images, show objects in context, and add resources (such as URLs) that point to further sources of information.

Sledge concluded by noting that everyone is demanding more of their information systems. Meeting these new demands will require an integrated approach where everybody—the community, staff, scholars—is responsible for contributing to the system. This idea is new for most museums, and it will require a change in mind-set. As we strive to document and preserve cultural information, we all have to contribute to the stewardship role.

Anne Graham, Senior Computer Specialist in the Libraries Digital Initiative Programs at the University of Washington Library, spoke more about community contributions in her talk entitled "The Olympic Peninsula Community Museum Project: A Window into a Community." (http://content.lib.washington.edu/cmpweb). This IMLS-funded project created an online museum showcasing the rich and diverse histories and cultures of the

Olympic Peninsula region. It did so by enlisting community members to curate, digitize, and describe who they are in their own way and with their own voice. However, the project encountered a number of roadblocks along the way, and Graham characterized her talk as one of "lessons learned."

The Olympic Peninsula Community Museum consists of a Web site with over 12,000 images, multimedia, manuscripts, oral histories, and nine exhibits reflecting life and history in the Olympic Peninsula community, a geographically isolated area of small towns with Native American and Latino populations. The "originals" for much of the cultural material on the site reside with individuals and families, who allowed these materials to be scanned for the project. Graham noted that most of these cultural materials would never have been made public if not for this project.

Although the University of Washington Libraries played a leadership role in the project, they wanted local groups to determine how they would be represented, to speak with their own voice and to tell their own story. This goal presented many logistical challenges. Community mobilization, for example, took much longer than expected, both to spread the word about the project and to develop trust among potential partners. In addition, assumptions about work expectations varied, and these expectations were not well articulated in partner participation agreements. There also were differing concepts of ownership among tribal groups that often led to delays in the signing of agreements. Other problems arose with the recruitment of community volunteers and with a community liaison who could not keep up with the management needs of the project. In addition, the project encountered differing notions of the concept of time, which, in rural communities, meant deadlines were frequently ignored.

Educational issues also surfaced. For example, conveying basic concepts about metadata and Web exhibit design proved difficult. In addition, community and interpersonal dynamics frequently interfered with project activities.

Graham concluded that the Olympic Peninsula Community Museum Project offers useful lessons for collaborative projects that involve community engagement in order to preserve local cultural heritage. The cultural heritage in local communities may be rich, and the communities themselves may be willing, but there are logistical, educational, and community dynamics that can introduce significant roadblocks into the process. Knowing about these challenges in advance can result in smoother project implementations.

Dr. Mark Louden, Professor in the Department of German/Max Kade Institute at the University of Wisconsin, Madison, spoke about "American Languages: Documenting Cultural Heritage Through Language Preservation" (http://csumc.wisc.edu:16080/AmericanLanguages). This IMLS-supported project began in 2003 with the goal of digitizing, interpreting, and making accessible audio collections that capture a variety of American languages and dialects. The majority of project materials come from three distinctive sound collections: the Max Kade Institute North American German Dialect Archive, the Dictionary of American Regional English (DARE) fieldwork collection, and the Mills Music Library ethnic music collection. There are four official project partners,[15] but over time they have developed informal partnerships with tribal colleges, independent scholars, and various community organizations.

The project team is digitizing indigenous varieties of language that are spoken by the *descendants* of first-generation immigrants, which are referred to as heritage languages. They are focusing on English dialects, German dialects, and Native American languages, with the goal of making these language varieties accessible and useful across disciplines and to the general public. The public aspect component is very important: Partners frequently go into local communities to talk about

15 The Max Kade Institute for German American Studies, the Center for the Study of Upper Midwestern Cultures, the Dictionary of American Regional English, and the Mills Music Library.

the project and present excerpts of the material to particular groups. The project also includes an important evaluation component designed to elicit comments from community members.

Efforts to record Wisconsin tribal languages have been particularly challenging. Most of these tribes are small, and very few of their indigenous languages are currently spoken. As Anne Graham mentioned in her case study, cultural dynamics also come into play. Some tribal elders decline to participate because they see language as transferring cultural and spiritual significance and not as an exercise in grammatical analysis. Others feel that recording their language is of no utility, since they are among the last speakers and nothing they do can maintain the language as a "live" (i.e., spoken) tongue.

Project partners have found that working with tribal colleges, as well as with students from various tribes who are studying linguistics and anthropology at the University of Wisconsin, has been very important in building trust. As they continue to develop contacts and build trust with Native communities, staff are forging ahead by highlighting Native American aspects wherever they can. Currently, they are extracting material about Native American culture from the German and English dialect recordings, and are highlighting this material on their Web site to illustrate Native/non-Native interactions in Wisconsin.

Louden noted that when the American Languages project was first proposed, the idea of linking it to a notion of cultural stewardship was never considered. As work got under way, the partners quickly realized they were creating a "loop of interaction" between themselves, scholars, and community members. Individuals and scholars who heard about the project came forward with donations or loans of their personal audio collections for inclusion in the project. In essence, they were developing, on the ground, a notion of cultural stewardship for language preservation.

Before the session came to a close, two questions were posed by members of the audience. The first question asked for clarification on the role of vocabularies as access points in these projects. Jane Sledge responded that language is a priority at NMAI, and it is working on using multiple versions of terms, not in a "preferred term" context but in way that shows all possible terms associated with an object, and having its system encompass the variability. Louden plans to incorporate dictionary projects to provide additional access points. For example, there is a Menominee dictionary project (being developed by a consortium of Native American linguists, tribal community college language teachers and other tribal members) that he would like to make available on the American Languages Web site so users can search the collections by Menominee object names.

A second question asked about the preservation strategies being used for the digital resources created from these projects. The American Languages project is keeping many digital copies in different locations on DVDs and CDs. NMAI is using multiple, redundant storage systems (which hold terabytes of information) that include online access, tape backup, DVDs, and off-site storage. It has a five-year migration plan and, because it holds all images in online format, it can also migrate these images from system to system. The Olympic Peninsula project uses DSpace as its digital repository. It also has multiple copies on portable media and servers.

Project: Florida Folklife Digitization and Education Project (see page 80). Photo: Guitar being played by Tom Walton, White Springs, Florida, 1982.

Conference Day Two

March 2, 2007

Keynote speaker **Dr. Deanna Marcum,** Associate Librarian for Library Services at the Library of Congress (LOC, www.loc.gov), discussed past and present efforts by the Library to preserve original cultural materials and make them accessible using various technologies. As those technologies become increasingly digital, and as their collections incorporate more born-digital items, it is devoting new energy and efforts to the unique preservation needs of these digital materials.

The Library of Congress provides stewardship for over 132 million items, such as manuscripts, books, music, sound recordings, film, and video. The collections grow by approximately 13,000 items a day. Selective digitization of collections began in the early 1990s, and the library quickly realized that the gains brought about by newly enhanced access also brought new challenges in preservation. These challenges were compounded when the library began accepting born-digital materials.

To address the new preservation challenges, LOC identified (in the late 1990s) five strategic methods for digital preservation. The first method focuses on developing better digital storage media, and LOC urged vendors to address this problem. The second strategy is to "refresh digital data" by copying the data from one location to another. A third method is migration, or the transfer of digital materials from one format to a newer, enhanced format. The fourth method, technical emulation, uses technology to imitate the function of the system on which the digital materials were originally created and used. The final method, characterized by Marcum as the "desperation option," is digital archaeology, where efforts are made to reconstruct otherwise unreadable stored bits.

To avoid facing the digital archaeology option, LOC realized that digital preservation would require ongoing management of digital resources from the point of creation. Its goal is now to create a comprehensive system for long-term storage and management of digital materials and metadata for ongoing use. Marcum noted that achieving this goal means committing to long-term processes that have no discernible end.

Over the years, LOC has been involved in a number of activities (in film and audio preservation, and in the NDIIPP, for example) that have spawned preservation efforts in various areas. LOC currently is completing a new national audio-visual conservation center in Culpeper, Virginia. This state-of-the-art complex will house all the collections and facilities of LOC's motion picture, broadcasting, and recorded sound divisions, consolidating staff and materials that are currently dispersed across seven different sites in Washington, D.C., Maryland, Ohio, Pennsylvania, and Virginia. The center will include laboratories for analog and digital preservation of audio-visual materials, a theater, and listening facilities for public programs and exhibits. A major component of the new center is a digital acquisitions and preservation system. The site will also have a training component where scholars, archival students, and other professionals can work and study using the latest preservation methods and technology.

The audiovisual conservation center is just one of several major preservation activities currently under way at LOC. Another project is "Chronicling America" (www.loc.gov/chroniclingamerica), a joint effort with the National Endowment for the Humanities (NEH) to provide digital access to U.S. newspapers of historical value. This project will span approximately two decades and is starting small to allow for iterative evaluation of technical guidelines, selection criteria, and effectiveness.

LOC also has embarked on a project called "Digitizing American Imprints at the Library of Congress."[1] With support from the Alfred P. Sloan Foundation, the library plans to digitize thousands of brittle books from its general collections and make them available on the Web. Marcum described this effort as a demonstration project to help improve the ability to safely scan works that are physically vulnerable. The project will include developing technologies for displaying and turning pages, displaying foldouts, and capturing table of contents and indexes.

---

1 www.loc.gov/today/pr/2007/07-020.html.

As part of this and other projects, LOC has established selection criteria for deciding what materials should be digitally reformatted. First it considers the cultural and societal value of the material. Is it of national interest? If so, can digitization make it more readily accessible while reducing wear and tear on the original? A second factor is the condition of the material. Is the original damaged or fragile to the point where it cannot be used? A third consideration is the usage of the material. Is it a high-demand item? Does it have high retrieval costs? Would digitization decrease those costs? Finally, they considers the material's physical characteristics and format. Do the physical formats lend themselves to digitization at a very high level of reproduction?

In all its preservation projects, LOC strives to preserve both the original material and the digital copy. Preservation of the original is considered a national responsibility, while preservation of the digital copy ensures access and extends the substantial investments made in these digital assets. To ensure the preservation of digital copies, LOC is working on policies for life-cycle management of digital data. These policies include extending the longevity of digital media, enhancing environmental conditions for this media, developing software and hardware requirements to extend the longevity and use of digital data, and creating methods and schedules for checking the integrity of digital files.

Marcum also discussed LOC's efforts to develop a new strategic plan for library services. This plan has a strong focus on digitization for access and preservation. It recognizes that access and preservation go hand in hand because access to digital copies reduces use and extends the life of originals. It stresses that LOC must increase its contacts with those creating digital works and expand its collecting skills to the digital world. It also notes that LOC must develop trusted repositories for its digital items and take advantage of digitization to help meet preservation needs.

Marcum concluded by stating that preservation of digital materials is still an evolving area, and there are no guarantees of how long we can preserve these types of materials. Therefore it is vital that we work collaboratively to improve our ability to extend the life of digital materials, and to share our own institutional progress in the digital preservation arena with others. Marcum encouraged the audience not to sit alone "huddled in blankets against the chills of change" but instead to use the technological ability available to us to operate beyond our walls and to help one another develop solutions to digital preservation challenges.

In the question and answer session that followed, Marcum was asked to elaborate on LOC's plans to develop repositories for the long-term preservation of its own digital materials. She responded that LOC has a plan to develop format-specific digital repositories. One such repository for audiovisual material is being created at the new audiovisual conservation center in Virginia. However, Marcum is advocating for a more comprehensive approach to digital repositories at LOC. Funding has been the primary obstacle, but LOC is keenly aware of the need and is working hard to make it a reality.

Moderator **Dr. Joyce Ray** of IMLS opened the session by discussing an infrastructure disparity between the sciences and the humanities. For years, the scientific community has been developing a cyberinfrastructure, funding data and supercomputing centers that store and process vast amounts of information, offering tools for sharing these data, and using this framework to advance scientific research. The humanities and social sciences have no such system in place. The speakers in this session represent organizations that are trying to correct this omission by promoting collaborations, research and development, and funding opportunities needed to create a cyberinfrastructure for the humanities.

The first speaker, **Dr. Steven Wheatley,** Vice President of the American Council of Learned Societies (ACLS, www.acls.org), served as advisor and manager of ACLS's Commission on Cyberinfrastructure for the Humanities and Social Sciences. In 2006, the commission issued a report[2] that proposed a cyberinfrastructure for the humanities. Wheatley summarized the report and its recommendations, and discussed the many influences that are moving the idea of a humanities cyberinfrastructure into reality.

He began by describing cyberinfrastructure as the "middle layer" of material between hardware and users. It is not the bandwidth, the hardware, or even the community of users. Rather, it is the discipline-specific software, the expertise, the best practices, the tools and collections, the training, the policies, and the collaborative environments.

Why do the humanities and social sciences need a cyberinfrastructure? New information technologies are empowering research in traditional areas of study, and most expressions of human creativity are now being "born digital." Scholars need to study these expressions in their original (that is, digital) form. No one, for example, can incisively study political elections today without looking at the blogosphere. The humanities need a platform—a cyberinfrastructure—to successfully conduct research in this new environment and to convey the resulting knowledge to future generations.

The commission's work was influenced by a 2003 report issued by NSF that is informally referred to as the "Atkins Report."[3] This report examined how cyberinfrastructure was revolutionizing science and engineering. It has had a catalyzing effect in the sciences, where it has been used to mobilize scholarly energy and articulate the concept of a cyberinfrastructure for this community. ACLS wants to do the same for the humanities and social sciences communities.

The ACLS report identifies five important characteristics for a humanities cyberinfrastructure. It must be accessible for the public good; financially, technologically, and intellectually sustainable; interoperable; facilitate collaboration; and support experimentation. The report also offers the following eight recommendations:

- Invest in cyberinfrastructure as a strategic priority
- Develop public and institutional policies that foster openness and access
- Promote collaboration between the public and private sectors
- Cultivate leadership that effectively directs investment, encourages choices, and mobilizes energy in the community
- Encourage digital scholarship
- Establish national centers to support scholarship that contributes to and exploits cyberinfrastructure

2 *Our Cultural Commonwealth: Report of the ACLS Commission on Cyberinfrastructure for the Humanities and Social Sciences* (www.acls.org/cyberinfrastructure/OurCulturalCommonwealth.pdf).

3 Named for Dan Atkins, the chair of the committee that issued this report, the report's official title is *Revolutionizing Science and Engineering Through Cyberinfrastructure.* National Science Foundation Report of the Blue-Ribbon Advisory Panel on Cyberinfrastructure. December 2004, (www.nsf.gov/publications/pub_summ.jsp?ods_key=cise051203).

- Develop and maintain open standards and tools
- Create extensive and reusable digital collections

Wheatley is optimistic about the development of a humanities cyberinfrastructure because there are many new investments being made in this area: for example, the NEH's Digital Humanities Initiative, the recent IMLS/NEH partnership to support collaboration among their grantees (See Brett Bobley's talk, below), and NDIIPP's efforts to create extensive collections of digital materials. There also are a number of recent studies showing how the current infrastructure for scholarship is in flux and urging changes in areas of research, publication, tenure requirements, and use of digital resources.[4] These studies influence decision makers and spawn new efforts that help in the creation of a cyberinfrastructure.

Scholars also are changing the equation. Their increasing demand for digital materials, and for the ability to conduct research in digital environments, is having a catalyzing effect. Academic impediments (such as the tendency to weigh the products of digital scholarship as less valuable than the products of traditional forms of scholarship) are eroding under the pressure.

Brett Bobley, Director of the Digital Humanities Initiative at the NEH,[5] spoke about two of the cyberinfrastructure investments referenced by Steve Wheatley—the Digital Humanities Initiative and the NEH/IMLS Advancing Knowledge Partnership. Bobley credited the ACLS cyberinfrastructure report with inspiring a series of NEH programs. He wholeheartedly agreed with the report's recommendation that federal funding agencies work together to promote collaboration, and NEH is acting on this recommendation in its partnership with IMLS and with discussions currently under way with NSF, the Department of Energy, and other grant-making agencies to create more joint, interdisciplinary programs. Bobley believes that the era of "boxed-in" federal programs is breaking down as everyone realizes collaboration across disciplines is critical.

Bobley elaborated on four new programs implemented by NEH over the last year under the umbrella of its new Digital Initiatives Program. The first program, titled "Digital Humanities Start-up Grants," is designed to help small projects develop prototypes or plans. This program funds institutions or individuals who have an innovative idea and are in need of start-up funds to develop the idea more fully. Unlike other federal programs that try to mitigate risk and ensure that federal dollars are well spent, this program is encouraging risk as a necessary part of innovation. The program will offer what Bobley called "small dollar grants" to give innovative ideas a "push" and see if they prove fruitful.

A second program is a joint collaboration between NEH and IMLS called "Advancing Knowledge." This program fosters large-scale digital humanities collaborations among museums, libraries, archives, and universities. It encourages collaborative "blue sky" thinking that will advance the field. It also requires (at the end of the grant period) a white paper that explains the collaborative work and identifies both the effective and ineffective aspects of the project.

The "Digital Humanities Fellowships" award is NEH's third new program in this area. These fellowships are designed for scholars conducting digital humanities work, ideally in a digital humanities center. A secondary goal of this program is to support long-term relationships. NEH hopes that scholars will maintain their

---

4 *MLA Report on Evaluating Scholarship for Tenure and Promotion* (www.mla.org/tenure_promotion); *Understanding the Use of Digital Resources in Humanities and Social Science Undergraduate Education* (http://cshe.berkeley.edu/research/digitalresourcestudy); Ballon, Hillary and Mariet Westermann, 2006. *Art History and its Publications in the Electronic Age* (http://cnx.org/content/col10376/latest).

5 See Digital Humanities Initiative at www.neh.gov/grants/digitalhumanities.html. Information on grant programs discussed in this section can be found at this site, unless otherwise noted.

relationships with the digital humanities centers once they return to their home institutions, thus extending the reach and impact of these centers and of digital humanities scholarship.

The last program is the "Digital Humanities Challenge Grant." Designed to support and strengthen infrastructure for digital technology, this award can be used, for example, to cover costs of construction and renovation of facilities, for equipment purchases, or to augment or establish endowments that provide income for staff, fellowships, ongoing maintenance of archives, etc. Bobley characterizes this program as an excellent opportunity for those who wish to create a digital humanities center.

NEH also has several programs that are *not* under the Digital Humanities Initiative umbrella, but still support digital programs. Its "Preservation and Access: Humanities Collections and Resources" program [6] funds the digitization of collections and preservation reformatting. An "Education and Training" program [7] supports the creation of training and materials used for preservation and access activities. In addition, NEH has a "Research and Development"[8] program that funds the development of technical standards, best practices, tools for preservation and access, and use of scientific and technical methods to preserve humanities collections. Bobley encouraged the audience to take advantage of these funding opportunities and to speak with him or any other NEH program officer about their ideas and how they may match particular agency programs.

Roy Rosenzweig, Director of the Center for History and New Media at George Mason University (http://chnm.gmu.edu), spoke about "Collaboration and the Cyberinfrastructure: Academic Collaboration with Museums and Libraries in the Digital Era." He began by noting that the urgent sense of collaboration fostered in the ACLS report and supported by NEH's programs runs counter to the tradition in his own scholarly discipline of history. To drive home this point, he cited a study of authorship of 32,000 articles indexed by a leading history journal. Only 6 percent of the indexed articles were written by more than one author.

The solitary endeavor of traditional historical research differs markedly from what occurs in digital humanities research. Citing the experience of the Center for History and New Media, Rosenzweig displayed a list of the nearly 400 collaborators —both individuals and institutions—that the center has worked with over the past dozen years. Rosenzweig is particularly pleased about the collaborations with libraries, museums, and archives. Academic historians have grown distant from these institutions, a reversal of the close association that used to exist between these parties from the 1880s to the 1930s, when groups like the American Historical Association led the fight to establish the National Archives, and when historians could be found in the Society for American Archives or heading divisions in the Library of Congress.

Rosenzweig believes that the digital era will force a "course correction" of sorts, because it brings a renewed imperative for collaboration between academic historians and museums, libraries, and archives. He cited four reasons this collaboration makes sense. First, museums, libraries, and archives have the materials that historians need. Rosenzweig demonstrated several projects at the center that integrate the study of history with material objects used as historical evidence. One project with the Smithsonian's National Museum of American History provides students and teachers of U.S. history with museum collections and curatorial expertise. Entitled "The Object of History" (www.objectofhistory.org), the project Web site helps students improve both their knowledge of history and their ability to understand material culture in the context of history.

---

6 NEH has merged two older programs ("Preserving and Creating Access to Collections" and "Reference Materials") under this new program name. See www.neh.gov/grants/guidelines/pcahc.htm and www.neh.gov/grants/guidelines/referencematerials.html.

7 www.neh.gov/grants/guidelines/pet.html.

8 www.neh.gov/grants/guidelines/researchdevelopment.html.

A second reason for collaboration is that museums and libraries have a permanence that is lacking in the more transitory world of individual scholars. The center is concerned about permanence because it is accumulating a large archive of digital assets and it has turned to librarians for assistance. In one of these collaborations, the Center is working with the Library of Congress to preserve over 150,000 digital objects in its "September 11 Digital Archive" (http://911digitalarchive.org).

A third reason to collaborate is for the benefit of our audiences. The Web has opened all of us to more heterogeneous groups, and museums and libraries have more experience dealing with diverse audiences than do scholars. The center's collaboration with the Smithsonian on the "September 11 Digital Archive" was critical in helping it build the audience it needed for collecting the online objects that are central to the archive.

Finally, collaboration is essential because the experience of museums and libraries complements that of scholars. Museums have materials and the curatorial expertise associated with those materials. Librarians have expertise in the development of tools and digital projects. The center has drawn heavily on the latter. For example, when the Center developed its research tool "Zotero" (www.zotero.org), it relied on the accumulated expertise of librarians in the area of reference management.

Scholars also bring a great deal to these collaborations. Their scholarly knowledge and subject expertise is used in museum and library exhibits, and they offer insights on the kinds of historical questions scholars will want to ask about collections that are being assembled by these institutions. Often they bring technical expertise to a collaboration, especially with smaller museums that lack a budget to carry out a digital project. Scholars also bring their teaching experience to bear on the educational objectives of libraries and museums by helping identify how students learn to analyze historical evidence.

For all their merits, collaborations do come with significant challenges. Different professional and work cultures, distance, and concerns about all sides doing their "fair share" are important considerations because they can engender distrust and conflict. The process of collaboration is hard work, and often more work than one expects. But digital collaborations can foster change in the humanities, not only by building more effective cyberinfrastructure but also by enriching the interpretation, preservation, and presentation of our cultural heritage.

In the ensuing question and answer period, an audience member expressed concern that the amount of money federal agencies awarded to humanities cyberinfrastructure projects was not enough to cover the larger costs of these projects. After a parent organization takes its overhead amount from a grant, there is very little money available for any particular project. Bobley agreed and pointed out that one of the less frequently cited reasons that NEH collaborates with other federal agencies is precisely because of the money issue. Cross-agency collaboration is one way to provide access to larger pools of money. In the meantime, NEH is working with the Office of Management and Budget and with Congress to try to get more money for its own programs. Joyce Ray reminded the audience of Elizabeth Broun's comment from the previous day that funders "give to vision, not need." The cultural community is starting to articulate this vision.

A second question asked about project sustainability. The federal agencies are funding digital programs, but when funding runs out, programs often disappear just when others have come to rely on them and their products. Bobley acknowledged that sustainability is critical. and NEH encourages grant applicants to provide a sustainability plan with their applications. He also emphasized that collaboration by its nature, often fosters sustainability. Scholars who create a repository or other Web resource will find they can sustain it over the long term if they partner with their university campus library, or with another organization.

A final questioner asked, "What can librarians, archivists, and museums do to assist humanities scholars in their research needs?" Rosenzweig felt librarians could help scholars address computational issues with tools that aid in complicated forms of analysis. Wheatley added that libraries, archives, and museums could help by providing institutional repositories for digital projects and materials that scholars create and use.

At this year's eighth annual WebWise Conference, the theme of "Stewardship in the Digital Age" urged us to consider how we can maximize our preservation successes in the digital era. Over the course of two and a half days, through workshops, panels, updates, keynotes, and Q&A"s, nearly 40 professionals discussed efforts, both large and small, to enhance preservation and access to our cultural heritage.

The overarching concerns of these discussions were threefold: preserving the physical (objects) and intangible (language, ritual, etc.) forms of cultural heritage; preserving the digital surrogates that we are using to encapsulate that heritage for purposes of access and/or preservation; and preserving the cultural materials that are increasingly "born digital," that is, the art, writing, research, and other creative expressions that are digitally created.

A number of underlying themes crosscut the many presentations. The first can be characterized as "everything old is new again." The concepts of curation, preservation, and access discussed throughout this conference are actually age-old traditions rooted in our institutions' missions and core values. Putting the word "digital" in front of these concepts means new challenges, *not* new changes, in our mission. As our collections and activities grow increasingly more "digital," we will adopt new tool sets and methodologies, but our core values remain intact.

A second major conference theme was collaboration. No one can work in isolation on digital preservation and access issues, because the needs and requirements are too great. We all benefit from (and generate) economies of scale, pooled expertise, larger funding, and more robust infrastructure when we collaborate. And collaboration means not just crossing over our museum/library/archives divisions, but entering into whole new communities such as science and engineering, and the commercial sector.

Additional themes emerged from the discussions of digital preservation activities. One of these themes is that digital preservation is a holistic endeavor. We cannot preserve just a digital object or a digital collection: We must preserve the entire digital ecosystem. Following closely on this idea is that digital preservation requires life-cycle management strategies. Planning for digital preservation should begin the moment a digital resource is created, using methodologies such as data curation and archiving, and must continue for the life of the resource. There is, as Deanna Marcum noted, no discernible end. Another theme is that digital preservation is a transformative process. Unlike the preservation of physical materials, the only way to preserve something digitally is to change it. For this reason, we must consider the "essence" of what we wish to preserve, not solely the form.

Another frequently stated theme was also one of the most succinct: "There is no simple solution. There also is no *one* solution." The conference presenters offered many strategies and methodologies for preservation and access. We must be cognizant of best practices, current trends, and new developments, and choose solutions that work within our own local context.

Conferences such as WebWise routinely point out the challenges or roadblocks ahead. For digital preservation and access, those challenges include the need for file format standards, vocabulary services (such as query expansion, vocabulary-assisted searching), greater clarity about the characteristics and quality of the digital objects we are creating, distributed networks of trusted repositories, large-scale storage technologies, and engagement of more communities such as commercial industries and the private sector. We also need training at all levels, from our professional training programs to less formal training for cultural institutions that are small and poorly staffed.

Most immediately, we need to raise awareness of preservation issues and needs within and outside of our communities. The staggering results of the preservation surveys presented at this conference make it clear that cultural heritage collections are at great risk, a large portion of our communities are ill prepared to deal with the problems, and we have inadequately

conveyed the issues to our local constituencies. We also need to raise awareness in our daily activities. When we create a digital resource, we should be asking ourselves, "How are we going to preserve this?" and start implementing good preservation practices in our everyday routines.

In addition, we need to borrow more from others. The OAIS reference model, for example, was developed for the sciences but has been embraced by the humanities, and the many instances of its use by conference participants demonstrate how fruitful this borrowing has been. Similarly, the ACLS report on cyberinfrastructure was modeled on efforts undertaken by NSF for the sciences and engineering and is now being used to motivate action in the humanities community. Borrowing and building on the works of others clearly yields significant returns and is another reason for breaking out of our institutional barriers, collaborating, and staying aware of developments in all sectors.

We also must broaden our constituency in our digital preservation and access efforts. The admirable and effective efforts of large-scale collaborations among the federal agencies, and the work undertaken by some of the major cultural institutions of our nation, are critical for seeking solutions for high-profile, at-risk collections and for keeping the topic on the national (and congressional) agenda. We now must bring aboard smaller institutions and their vast, but largely hidden, cultural materials. Small, poorly staffed institutions are thought to comprise more than half of all collecting institutions in our nation. If we fail to bring them along, we are committing a form of benign neglect, leaving behind a significant portion of the community and untold cultural collections.

In examining preservation and access in the digital age, we need to reconsider our notions of stewardship. What defines successful stewardship in this arena? For physical collections, it is maintaining objects in a more or less homeostatic state so they can be accessible for research, reflection, and enjoyment. For digital art, it is encapsulating the behavior of a work rather than its form so it can be "performed" in the future for the same purposes. Stewardship, it seems, can take different forms to reach a common end.

When thinking about stewardship, particularly preservation and access of physical collections and their digital expressions, loss is a looming specter. So when the tangible and intangible products of culture *can* be preserved, we have scored a victory over the vicissitudes of time. But when something no longer physically exists, that does not mean it ceases to be. If cultural institutions can collect the essence and importance of a work by digital means, and transmit these qualities to successive generations, that is a victory as well. WebWise 2007, with its rich and diverse program, illustrated that successful stewardship for cultural heritage collections is a multifaceted endeavor.

Project: Cylinder Preservation and Digitization Project (see page 78). Photo by Rod Belle. University of California, Santa Barbara, Library.

Resources

Many thanks to Anne for that very beautiful introduction. Little did I know that she shares my passion for Albert Pinkham Ryder. I do encourage all of you to come and see our museum. It is just beyond dazzling. We worked there for many years without realizing what a beautiful building it was, and it never looked better than today.

I especially want to acknowledge the three organizations that put this conference together because over the years, not just the WebWise conference, but so many other occasions to bring us together to talk about these issues have been provided by these three—by IMLS, OCLC, and the Getty Trust. I just cannot say how important it has been for me personally to be able to network and learn from these events.

[PowerPoint video presentation shown, available at www.imls.gov/news/events/webwise07.shtm.]

That was the fun part of the presentation. As I stand up here I have to give special thanks to Michael Edson on the front row for creating that for us.

This is probably the moment for me to confess the truth, which is everything I know about this subject I have learned from my talented staff over many years. Our very first research database dated back to the early 1970s and was prepared as a celebration of America's bicentennial in 1976. One of the participants in doing it was Eleanor Fink. She was one of the first to introduce me to the potential for searchable databases. Then she was succeeded by Rachel Allen, who is now the deputy director of our museum and integral to all of our programs, one of the wisest content managers I have ever known. We were fortunate, in the early nineties, to hire a publications chief to help us produce books. His name was Steve Dietz, and as soon as the Web became a factor he would prove to be one of the most intuitive people about the potential of the Internet in museums. He was the one who helped me over some of those crucial humps like, can we really afford to put our images into digital format online? Won't someone steal them? He was the one who helped me through what I think of as early questions.

We went on to have a great many talented people, too numerous to mention. I might just cite Thorny Staples, who was particularly gifted in figuring out applications that are richly meaningful for the humanities. We currently have a terrific internal team. I might just mention Christine Hennessey, who heads our Research and Scholars Center, a long-timer with superb database management skills, and Theresa Slowik, our current publications chief. They have adopted and integrated technology into all the ways we work every day—not just for a Web program, but for tagging and filters and everything that makes our content more useful. And now we have Mike Edson to head our technology programs, and I feel very privileged to be able to work with him and a great team.

We want to be the place to go for American art. We want to be the crossroads for everyone who has a question about American art. But we know that, although we have been able to achieve some great things in the past and we are very proud of our accomplishments, this field is changing faster and faster every day and it is harder and harder for us to keep on top of the way we use the Internet and all of the other new media programs. It is all changing too quickly.

When we started, back in the mid-nineties, a Web site was a nifty extra. You could get some recognition for putting some images online, or doing something related to one of the exhibitions, and it would enhance your brand. But today it is totally integral. There is really no way any longer for public museums and libraries to separate their bricks-and-mortar business from their virtual business. They have to integrate with each other, and it is no longer optional to do these things. It is now required.

We know that one of the impacts of all of these new media programs and Internet programs is to cause some really serious disruptions in traditional business models. It is a big threat right now to the print and broadcast media businesses, and, in a way, if we don't get on top of it, we could find our very own nonprofit business empires threatened as well. So new media and

Internet strategies are not just an option, not just an enhancement, but a requirement.

As we think about strategy, we have decided that we should adopt a long-tail strategy. Most of you know this is Chris Anderson's idea that the Internet has transformed how business cultures work. It is no longer necessarily the handful of top hits that will drive most of your traffic or business. If you aggregate all of those little niche interests and individualized hits that come in the long tail, you can actually create a business model around that.

Imagine that Ryder is somewhere there on the long tail. He is not O'Keeffe, he is not Hopper, he is more of a cult figure in American art with a passionate following, but it is small. And after Ryder we have something like 7,000 artists represented in our collection, and about 6,900 are somewhere there on the long tail.

We think this is a strategy that will work for us. Ken Hamma talked about it earlier. It is not that we have to do something dramatically new and different. What we need to do is figure out the strengths of what we have already been doing and find a way to adapt those to this new culture. We have a priceless collection. We have a wonderful and talented staff. We have always been devoted to research and scholarship, so our files are bursting with information and content. We want to share knowledge. We have a role as a federal museum with an entire national purview, and the museum has always been a populist place that believes in giving back to the public. So what we need to do is harness all those assets and use them for the benefit of American Art 2.0.

We have analyzed our Web site to try to figure out who is coming, for what reasons. It is interesting to find out that the 10 most visited areas of our web site, if you add them up, only generate 25 percent of our traffic. The other 75 percent comes from the long tail. At least two-thirds of our visitors do not come to see something that we have shaped and created specifically for them, such as our online exhibitions. Instead they are coming for our assets, our data, our content; they have a question about American art and they are hoping to find the answer somewhere in the assets that we provide. They are not coming to hear what we want to tell them. They are coming to find what they want to know.

Anne talked a little about the Luce Foundation Center for American Art in our museum; this is a publicly visible art storage and study facility. On the left you see the extremely beautiful historic wing where it is housed. It has a floor level and two mezzanines. On the upper right you get a glimpse of the 64 glass cases and below you see one of the nifty drawers pulled open. We have drawers for portrait miniatures, for medals and medallions, and for the craft jewelry. This facility has allowed us to go from traditionally showing 1,000 artworks in our public galleries, to being able to show four times that many artworks, because 3,400 artworks are on view in this center.

It is immensely popular with the public. We think one reason is that you have the feeling of getting behind the scenes. Visitors think, "Okay, the curator has decided what I should be looking at in the galleries," and that's a very kind of top-down authoritarian kind of structure. But when you come to the Luce center, you get to decide for yourself and see all of the things that the curators didn't choose.

While we were preparing this physical place, we had a team of people working to generate an artist's biography for every one of the artists represented in this center. We have data and information and research on every one of the 3,400 objects there. All of this is posted online. We have a number of media assets as well, including 100 artist interviews and videos. So we have tried to make it a very rich place for knowledge, and I think that is driving a lot of traffic to our Web site. This may be our poster child for bringing the bricks and mortar museum and the virtual museum in synch with each other.

Now where do we go from here? We think we have been using some of these long-tail ideas in an intuitive way over a while, but we really need now to establish it as a firm strategy. We need to figure out what resources and

requirements are needed. And this slide is our diagram of what we think is involved.

We want to determine the scope of what we intend to put online and publish. We think the answer to that is to publish everything. Then we want to figure out how we are going to make it easy to find, because without that users will become frustrated or will never know it is there. Third, we have to find a way to put our users or our customers at the center of everything we do. Finally, we have to recognize the harsh reality that this is going to take a lot more financial and staff resources than we have allocated so far. We can't continue to do it on a shoestring.

Let's start with "publish everything." Our museum has objects that have been collected since 1829, so we have assets and information dating back a long way. We started our research programs in a major significant way in the 1970s. We have enormous files, enormous records, huge databases, and fabulous ephemera that have been accumulating now for at least 37 years. Not all of it is accurate but it is in there, and we think little by little we should find a way to create content on-line that spreads across the entire long tail, the whole scope of the collections.

In the past we have tended to put our online projects at the end of the museum food chain. We would prepare a big show—like George Catlin and His Indian Gallery. We would spend a couple of years generating research, writing catalogue essays, preparing label text, doing public programs, gathering together all the assets, sending them through multiple reviews, sending them to the editors, and then finally, at the end of all that, we would post an enormous site on George Catlin and it would be a wonderful rich resource for scholars.

We are moving away from that strategy now. We want to do more with what we think of as "micro" content. We would like to be leaner and faster, we would like to be posting things more frequently. We're moving away from an "all or nothing" risk, where if something doesn't work you have lost an enormous amount of effort.

We would like to make it a habit that content is constantly being posted. If it needs further editorial review or if there is some error found, we know we can always go back and do that. Getting away from a monolithic structure and looking at something that is a little faster and leaner and more about micro-content allows us to sprinkle information all across the collection assets and not just focus for a year or two on a single artist.

We have been doing something similar with our blog. We posted 110 small stories last year that gave us a chance to address a wealth of topics and to engage people in conversation about them.

Another way we have used micro-content is to ransack all of our past publications and also the work we are doing for the Luce Foundation Center for biographical information on artists. Recently we posted 1,300 artists' biographies, which turn out to be a very rich draw for users.

So we are looking for ways to adopt a strategy that is not so labor intensive to get information online. It is challenging, however, because to use this micro-content and make it meaningful, you have to have the structure to support it. We know this means getting more robust and seamless art information databases.

Right now we are drawing on our Web site from 23 data sources, created in many formats over many years. We have over one million records. It is a jumble. We would like to rationalize with a better, stronger infrastructure. We need better data design. We need content management to be made more robust. We need to emphasize a lightweight technological framework.

The second thing we want to do is make all the content more easily findable. We are going now from "the book" with a table of contents and an index to the Web site. We know that "real estate" on our homepage is limited. We have exactly 12 topics listed on this very first screen. It doesn't begin to scratch the surface of what is in the 23 databases and all the other assets we have. I am constantly hearing from staff that "we have this fabulous asset but nobody can find it online."

How are we going to make all this findable?

We can create all the content in the world but if it is not easy for users to find, it won't do us very much good. Somehow we need better human interface design. We need better graphic design; we need better information architecture; we need much better search engine technology. That comment earlier today about how everybody is coming to Google rather than the library home pages is absolutely true of museums as well.

There are other challenges also to accomplishing our goals. For instance, we are the poster child for brand confusion. These nine names are just some of the names that are associated with our program. Our parent organization, our museum name, our new centers, the Reynolds Center name that is applied to the entire complex, the historic name of the building, our branch museum at the Renwick Gallery, and of course our close partnership with the National Portrait Gallery—all have individual brand names.

I don't know if we will ever completely resolve this dilemma, but we can do a lot to improve the situation. How I would love to just have the name "MOMA"!

We are making this a top priority as we figure out how to communicate with our users.

The third element is how to put our customers (or viewers or users) at the center of everything we do. They are extraordinary. We find that people who are interested in American art come to us from an enormous range of backgrounds and interests. They might love American art; they might be interested in American history or some aspect of American culture; or it might be something completely different that brings them to us. It might be about how artworks are made or artists' materials—just an infinitude of ways to access what we do.

I know Mike Edson says that in his neighborhood, demographically and economically, everybody looks pretty similar, but one of his neighbors races Chinese dragon boats, another one is an amateur astronomer, and a third is a backcountry kayaker. These niche interests are out there and people come to us for peculiar reasons that we can't always know. What we do know is, just from looking at eBay or Wikipedia or Flickr, is that if we could harness that we would really be onto something great. In this long-tail strategy, finding a way to capture the passion that our fabulous customers or users are bringing to us is really key. If we can find a way to make that the centerpiece of our strategy, we will be okay.

We are going to start by focusing on our core customers, the people who know and love what we do. They are probably best positioned to tell us why they come and where they want us to go. Then there are others who come to "sample us." They come, have that first flush of enthusiasm, and then what? How do we capture them and make ours an important site for them to come back to? One way to do that is to allow them to add their knowledge to our site. I don't know if that means tagging or adding reviews or some other way of posting what they know. We really haven't solved that yet.

Of course there is always some conceptual "stone in the road" that you have to deal with. There's one that I am stuck on right now. I know that our museum, like all of your institutions, represents something very special to the public. We are regarded as experts who have authenticated knowledge. People believe they can trust what they see on our site because we will have gotten it right. We don't want to dilute that sense of authority by having a free-for-all with everybody's information completely jumbled up with our own expert knowledge.

But others have confronted this and we think we can do it too. Frankly, the relationship between institutions like ours and the people they serve has been turned upside down. Not for nothing did *Time* magazine say "you" are the Man of the Year. Things are different now, and we must confront this. Otherwise we are going to become "timeless oracles" who are sitting on the sidelines of the future.

Again, we can start by going back to what we are already doing well. We use a lot of volunteers

in a lot of ways, and they already contribute a lot. How do we adapt some of those volunteer programs so they can "volunteer" online as well? There's one classic example, which I really love: For about a decade we did an inventory of all public sculpture across America. This was a wonderful program that we did in partnership with Heritage Preservation. We involved thousands of volunteers over all the 50 states to find, catalogue, and assess the condition of public monuments. We put all the data into a searchable database online. We did not verify the information, though we corrected anything we knew was wrong, but for the most part this was user-created or volunteer-created information. It is an absolutely fabulous, searchable inventory of 32,000 public monuments across America, a truly great resource. So we are going to try to build on ideas like that.

These notions of trust are changing. We need to be a site that is alive and listening. We need to inspire devotion from our users that makes them want to participate and contribute. After all, we are an art museum and we are here to honor the creativity in each and every one of us. We need to keep that idea foremost as we address these issues about how to capture the value of our users.

One way that we are doing it now is through our national education program, where we are engaging a lot of teachers and students to use online resources and to contribute their help creating curriculum and content.

The fourth element naturally follows from the other three. We have achieved what we have accomplished so far just by raiding the museum budget (i.e. snatching a few thousand dollars here and there, or maybe moving a position from one office to the IT office). We have worked on a shoestring. We have never had enormous grants or new federal allocations for technology or media. We have had to be smart.

I think we have reached the end of that strategy. We just opened our great historic landmark building after a long and expensive renovation. We didn't try to do it on a shoestring or out of the existing budget. We confronted head-on the fact that doing this renovation properly was going to require hundreds of millions of dollars. We took the case to the Congress; we took the case to private funders. We accomplished something magnificent that everyone is very proud of, but we acknowledged up front that it was going to cost a bundle—and it did.

So I think we have to do the same with American Art 2.0. In retrospect I wish I could go back seven years and, as we were planning the museum renovation, embed within the plan all of these ideas for the future of American Art 2.0. Instead I now say, look, we have completed renovation of the bricks-and-mortar museum. Now we are going to turn to renovating the virtual museum. And I think it is going to take a lot of money. I don't see any other way. We need a strong strategy and we must persuade people that it is worth doing.

To use another buzzword, we want to widen the pipes. We have the right history and the right collections and the right mission and the right attitude and the right audience to succeed in this new marketplace. We don't have to strike out in some radically different direction. We have to build on what we have done, get smart about how to adapt it, and try to use the strengths that we already have.

This highly complicated technology architecture slide symbolizes some of my more apprehensive moments, when I'm concerned about how we are going to manage to do all this. But we can simplify the message. We want to build value incrementally. We want to enjoy and learn from interacting with our users and our audiences. We want to share our passion and we want to make a lot more information, a lot more insight, and a lot more excitement flow between us and them.

Abstract

Our time's digital information revolution makes being a librarian exciting. The Library of Congress, like others, is exploring new ways of using digital technology for both access and preservation. This work, and the excitement, will grow as the library completes moving its audiovisual resources into its new National Audiovisual Collection Center. The library hopes to share new developments and work with others in meeting the challenges of the digital information era.

Introduction

Amid the daily challenges of dealing with personnel problems, budget questions, and other administrative headaches, I sometimes forget the most important thing. In earlier days, American librarians were happy if they could finance shelving and keep a wood stove going. I recently read about a nineteenth-century librarian who "wrapped herself in a blanket, with a soapstone at her feet, during the coldest of Saturday afternoons, the only time the library was open" (Heidinger, 2006, p. C-2). I easily forget how far libraries have advanced in a relatively short time.

We have far outrun what that early librarian, even if she got warm, could possibly have imagined. Since the 1990s, we have developed the electronic magic of the Internet to make material in our institutions quickly available to anyone, anywhere, with computer access. Our digitized materials include sound recordings and film as well as texts.

We are working today in the midst of the digital information revolution. We are part of an exciting time in the cultural history of humankind. This paper reports on what we are doing and thinking at the Library of Congress about the stewardship of library resources in this era of digital information.

Discussion

I call this paper "Digitization for Access *and* Preservation." The thought that the two could be linked is important to me because we at the Library of Congress have so many resources to preserve. We must provide stewardship for more than 132 million items. These include more than 58 million manuscripts, 30 million books, 13 million prints, 5 million maps, 5 million pieces of music, 3 million sound recordings, and 1 million films and videos. Moreover, our collections grow by approximately 13,000 items every day (Library Services, 2006a, p. 2).

Now we have the opportunity to make much of this material accessible far beyond our buildings' walls in Washington, D.C. How will we make use of that opportunity while also safely preserving so many things? I will answer with a story that began a decade and a half ago.

1 How exciting it was back in 1990 when the Library of Congress launched an experiment with the new technique called digitization. For the next four years, we identified audiences for digital collections, established technical procedures, wrestled with intellectual property issues, and explored distribution formats such as CD-ROM. We sent CD-ROMs containing some of our materials to 44 schools and libraries across the country. We received enthusiastic responses, but the format proved inefficient and costly (Library of Congress, 2007a).

Then came the Internet. In October 1994, with $13 million from private donors, we announced our plan to go online with a National Digital Library Program. The program's flagship became the American Memory historical collection that we had begun digitizing in the experimental project. The collection contained historical documents in multiple media. The Web became the means of making access efficient and affordable. Private donations to the program soon tripled. And the Congress gave us another $15 million for five years (Library of Congress, 2007a).

In 1996, we expanded partnerships in the program. With $2 million from the Ameritech Corporation, we opened a competition. We invited nonfederal libraries, museums, archives, and historical societies to submit proposals for digitizing material to add to our American Memory collection. We placed 23 prize-winning digital collections on our American Memory Web site (Library of Congress, 2007a).

By the year 2000, we exceeded our goal of putting five million items on the site. Today it contains nine million items in more than 100 thematic collections. Included are digital copies of books, manuscripts, pamphlets, prints, photographs, maps, sheet music, sound recordings, and films. Each collection appears with explanatory features. And all collections can be searched electronically (Library of Congress, 2007b).

In the 1990s, many other libraries in the United States and abroad also built digital collections for access online. This seemed wonderful. But at the same time, most of us realized that we did not know how long we could *preserve* the resources we were creating. Digital media lack the durability of paper. Digital documents depend for readability on computer systems that quickly become obsolete. And digital preservation has to deal with formats of many, changing kinds (Library of Congress, 2007c).

We did not digitize with the intent of replacing original materials. But we wanted to preserve also the digital copies in which we made such substantial investments (Arms, 2000). Thus our gains in access brought new challenges in preservation. The challenges increased when we started accepting materials *created* digitally. These included the selection of Web sites that we and the Internet Archive began trying to preserve.

Toward the end of the twentieth century, we drew upon experts outside as well as inside the library to identify five major methods for digital preservation (Arms, 2000). The first essentially called on vendors to develop better digital storage media. The second called for "refreshing" digital data, which basically meant copying streams of digital "bits" from one location to another.

The third, more complex strategy came to be called "migration." This meant transferring digital material from one format to a newer, and hopefully enhanced, format.

Our fourth strategy, called "technical emulation," required programming new computer systems to mimic systems on which digital material had originally been generated.

Our fifth strategy became a desperation option that we called "digital archaeology." This meant trying to reconstruct the meaning of digital material that had become otherwise unreadable. Of course, we also paid attention to storage conditions, data replication, data validation, and other measures for basic security.

We knew that waiting to act until digital material had deteriorated would leave us only the digital archaeology option. To avoid that, we began to try managing digital data from their creation. That included trying to retain metadata—information that helps us manage and retrieve digital data. We envisioned creating a long-term, comprehensive system for storing and managing multiple kinds of digital materials and the metadata needed for their ongoing use (Arms, 2000).

2 With those thoughts in mind, we began in the twenty-first century to do three things for preservation. One, we continued our experimentation. Two, we began building a National Audiovisual Conservation Center to house audiovisual resources in digital as well as traditional forms. Three, we crossed our fingers and prayed that solutions to the digital preservation challenges would emerge.

A preservation expert on our staff expressed our hope more confidently. In concluding an excellent report in 2000 on our digital preservation program, she wrote "Technology advances, while sure to present new challenges, will also provide new solutions for preserving digital content" (Arms, 2000).

We were not so confident, however, as to wait around for that to happen. We recognized that it would not happen without a major effort—an effort larger than the Library of Congress could undertake alone.

Therefore, in 2001, we joined others in persuading the U. S. Congress to appropriate nearly $100 million for a national program to "ensure the long-term storage, preservation, and authenticity" of digital collections (Library of Congress, 2004, p. 175). This became the National Digital

Information Infrastructure and Preservation Program—known by the acronym NDIIPP.

Many libraries and others helped with NDIIPP's development. And institutional partners have received NDIIPP grants for projects to improve digital preservation. The Library of Congress joined with the U.S. National Science Foundation to administer the NDIIPP grant program (Library of Congress, n.d.a.).

Some of the grant recipients work on preserving nontextual resources. For example, the University of North Carolina at Chapel Hill is using an NDIIPP grant to develop a framework for preserving digital video collections. The San Diego Supercomputer Center at the University of California, San Diego, is using an NDIIPP grant to develop a process for managing the preservation of videos from creation through ultimate use (Library of Congress, n.d.a.).

In addition, NDIIPP is helping SCOLA (Satellite Communications for Learning Associations) to archive high-interest television programs. SCOLA is a nonprofit educational corporation that receives and retransmits television programs of potential research value from around the world (Library of Congress, 2006). The Library of Congress and many others expect to learn from such NDIIPP-assisted projects.

At the same time, other preservation efforts have emerged from other national programs.

For example, in 1992, the U. S. Congress passed a National Film Preservation Act. It financed a fact-finding study by the Librarian of Congress. The completed study, entitled *Film Preservation 1993*, reported that only half of films made before 1950 survive. Only 20 percent of feature films made in the 1920s survive. And only 10 percent of those made in the decade beginning in 1910 survive (Library of Congress, n.d.b).

The study also found that even the surviving films suffered from preservation problems. These included color fading, film-base decay, soundtrack deterioration, and flammable film stock (Library of Congress, n.d.b).

Out of the study came a plan for action, entitled *Redefining Film Preservation*. The Library of Congress worked on this plan with the National Film Preservation Board, and with archivists, educators, filmmakers, and film-industry executives. The plan, released in 1994, called on all concerned to take such actions as the following:

- Make wider use of low-temperature, low-humidity storage to retard film deterioration and buy time for restoration projects.
- Increase the availability of films for education and public exhibition.
- Develop a public-private partnership to share preservation information, restore important films, and search foreign archives for "lost" American films.
- Create a foundation to raise money to preserve newsreels, documentaries, independent films, avant-garde films, and socially significant amateur film footage. (Library of Congress, n.d.c.).

Similarly, we have joined with others to focus attention on needs for preserving sound recordings. Eight years after the Film Preservation Act, the Congress passed the National Recording Preservation Act. It called for establishing a National Recording Registry in the Library of Congress.

Also to help preserve significant sound recordings, the act called for a National Recording Preservation Board. The act charged the board to study current preservation needs and practices and to plan a national audio preservation program.

Participants in this planning have declared that "audio preservation today is not simply a matter of collecting and storing, or transferring endangered records to the digital domain." Long-term preservation requires commitment to long-term processes, which may have, as one expert put it, "no discernible end" (Library of Congress and National Recording Preservation Board, n.d.).

In preparing the audio preservation plan, we held public hearings and solicited comments from representatives of sound-recording

archives, recording companies, audio engineers, and interested organizations of scholars. Also we consulted specialists in intellectual property law and individuals with collections of recorded sound (Library of Congress and National Recording Preservation Board, n.d.).

The study coincided with our work to build a new National Audiovisual Collection Center. We basically completed the center in March 2007.

Since 1999, our Motion Picture, Broadcasting, and Recorded Sound Division and our American Folklife Center have worked on ways to make digital copies of moving image and recorded sound collections. Among other things, we have explored means of scanning motion picture film, of transferring video recordings from tapes to digital files, and of packaging digital materials (Library of Congress, n.d.d.). We carry on such work in our new National Audiovisual Collection Center.

3 What exactly is this center? It is a complex of four structures. They cover 45 acres near Culpeper, Virginia, in the United States, 60 miles south of Washington, D.C. The complex occupies 415,000 square feet. We built much of it into the west face of a mountain, covering the buildings with earth, grass, and trees to keep the site as natural as possible (Dalrymple, 2006, pp. 167–168).

In March 2007 we accepted the complex officially. The Packard Humanities Institute transferred it to the Architect of the U.S. Capitol, who oversees the operation of our buildings.

Packard managed and financed the construction at a cost that we estimate will reach $150 million, the largest gift in the library's history. The U. S. Congress provided an additional $52 million for buying shelving and equipment, relocating collections and staff, and hiring new staff (Dalrymple, 2006, p. 171).

Once we have fully moved in, the complex will house all collections and facilities of our Motion Picture, Broadcasting, and Recorded Sound Division. Also it will provide space for a staff that we anticipate will grow to 150.

A Collections Building will store all our audiovisual collections, except for those on nitrate film. This flammable film will go into specially constructed vaults in a second building. A third structure contains a central plant for heating and air-conditioning the complex. Our fourth structure is a three-tiered Conservation Building. It houses administrative, curatorial, and processing staffs. It also contains a theater and two laboratories for the preservation of all kinds of films, videos, and sound recordings (Dalrymple, 2006, pp. 168, 170).

Because the Collections Building is underground, it efficiently provides ideal conditions for audiovisual storage: low temperature and low humidity. The building contains large vaults with compact shelving for all of our media formats (Dalrymple, 2006, p. 168).

The 175,000 square feet of the Conservation Building contain a state-of-the-art facility for listening to sound recordings. This area and exhibit spaces are open to the public. Also, the building has naturally lighted work spaces. And it has a 200-seat theater with an organ console for music that used to be heard with silent movies (Dalrymple, 2006, p. 168).

The complex enables our library to consolidate collections previously stored in three Washington buildings and five others in Maryland, Ohio, Pennsylvania, and Virginia. Audiovisual reference services remain in our recorded sound and moving image research rooms in Washington. Sound and videotape collections may be accessed there electronically. Film collections, at least for now, are brought to researchers in Washington from Culpeper on a regular schedule (Dalrymple, 2006, pp. 169, 171).

Audiovisual materials comprise a rising proportion of the world's historical record. We expect our complex to have room for additions to our audiovisual collection for at least the next 25 years. This estimate includes storage for materials created digitally (Dalrymple, 2006, p. 169).

Our new audiovisual center includes a system for acquiring and preserving digital materials. There we intend also to preserve digitally the analog

materials that we previously would have transferred to analog formats that are growing obsolete.

As Gregory Lukow, chief of the division in charge of our audiovisual collections, has explained, "The change will be evolutionary and sequenced." Already we have begun to preserve sound recordings digitally. Now we are working on the digital preservation of videotape. Eventually we hope also to use digital technology to preserve and manage film. This is more difficult to do. We will need improvements in technology that can lower costs. (Dalrymple, 2006, p. 170).

Our plans include sharing with other cultural institutions the innovations we expect to develop in the new center. It contains meeting places for visiting scholars, archival professionals, and students from graduate courses in moving image and recorded sound archiving. There they will be able to discuss curatorial and technical challenges and examine improvements in audiovisual preservation and access (Dalrymple, 2006, p. 170).

For example, the new center has an experimental image workstation. In it we will use newly developed technology to speed the digitizing and preserving of 78-rpm shellac and acetate recordings. The new technology comes from the Lawrence Berkeley National Laboratory in Berkeley, California, U. S. A.

The laboratory calls the technology IRENE, which stands for "Image, Reconstruct, Erase, Noise, Etc." IRENE is a kind of restoration software. It enables us to create high-resolution digital maps of the grooved surfaces of deteriorated recordings. From these images, technicians can remove debris and extraneous sounds, and repair damaged portions (Sternstein, 2006).

The laboratory used a grant from the National Endowment for the Humanities to create the image machine and demonstrate that it works. The laboratory is preparing software that will enable library technicians to use IRENE with just basic training (Sternstein, 2006).

Though we are excited about getting into our new audio visual preservation center, we are also excited about other preservation projects. For example, in 2005, we joined with the National Endowment for the Humanities to provide digital access to preserved newspapers of historical value.

Since 1982, the endowment has spent $54 million on grants to preserve some 70 million pages of newsprint on microfilm. These grants have gone to repositories in every state and three territories of the United States. The grants support microfilming of newspapers published since the eighteenth century. The Library of Congress has provided technical assistance since the project's start. The endowment expects to conclude the project in 2007 (National Endowment for the Humanities, n.d.).

Next we will work with the endowment to make historical newspapers accessible via the Internet. Over the next two decades, we will develop an online National Digital Newspaper Program. We will digitize historically significant newspapers published in all the U. S. states and territories between 1836 and 1922. These will be available in a free, searchable database (National Endowment for the Humanities, n.d.).

We have begun by supporting projects to digitize 100,000 pages from newspapers published in California, Florida, Kentucky, New York, Utah, and Virginia between 1900 and 1910. This first batch will help us evaluate technical guidelines and selection criteria. Also we will evaluate whether the program effectively enables users to browse and search newspaper pages (National Endowment for the Humanities, n.d.).

If we receive continued funding, we will make grants in every state and territory. In each, one organization will coordinate newspaper digitizing by several partners. If this project succeeds, the microfilm copies will ensure long-term preservation while we use the digital copies to provide access (National Endowment for the Humanities, n.d.).

We feel excited also about a project that we announced in January 2007. We call it "Digitizing American Imprints at the Library of Congress." The Alfred P. Sloan Foundation has given us $2

million to digitize thousands of books. These include "brittle books" that we are in danger of losing. We hope to make this a demonstration project from which many libraries can learn how to scan their physically vulnerable works safely (Library of Congress, 2007d).

Our digitization project uses book-scanning technology called "Scribe" from the nonprofit Open Content Alliance. In addition, the project will develop technology for electronically turning pages, displaying foldouts, and capturing tables of content and indexes (Library of Congress, 2007d).

Besides brittle books, we will digitize other works, all in the public domain: American history books, genealogies, regimental histories, other U.S. Civil War material, and six collections of rare books. We also will digitize works about photography, particularly artistic publications, biographies of photographers, and works on technical aspects of photography (Library of Congress, 2007d).

We have established formal selection criteria for deciding which works to reformat digitally:

- We consider the *value* of materials—are they of national interest? If so, digitization can make them easier to access while reducing wear and tear on the originals.
- We consider the *condition* of materials—are digital copies needed because the originals are damaged, fragile, or on unstable media?
- We consider the *use* of the materials—are they often in demand? Or do they have high retrieval costs that could be reduced by providing digital copies for use?
- We look at materials' *characteristics*—can their physical formats be digitized at an acceptably high level of reproduction? Do we need to test their reformatting possibilities?
- We also give digitizing priority to *access aids*—guides, indexes, and databases that help users identify and locate useful materials (Library of Congress, n.d.e) .

In the Sloan project, as in others, we want to preserve the original copies. But we want also to ensure that digital copies will be accessible for a long time. Consequently, we are working throughout the Library of Congress on policies to manage digital data over time. We seek more durable media, better storage conditions, and improved technologies for managing digital data. We also work on methods and schedules for checking and maintaining the integrity of digital files (Library of Congress, n.d.f). Additionally, we have begun work on a new, overall strategic plan for the Library Services unit of the Library of Congress. The plan will cover fiscal years 2008 through 2013. It will give major attention to digitization for access *and* preservation. The plan is a work in progress, but here is a summary of items stressed in the current draft:

> *. . . we need to better understand what is being created digitally and increase our contact with those creating these works. Our skills at collecting traditional works need to expand to the digital world. We need to identify digital resources as they are created and apply our collections specialists' knowledge to determine which items should be collected . . . .*
>
> *[Also we must] work closely with the Library's Office of Strategic Initiatives to advance the science and practice of preserving digital works, and to develop trusted repositories for digital items in the Library's collections"* (Library Services, 2006b, p. 8).

Conclusion

I think our new strategic plan will emphasize the following general points:

Like most libraries, we will continue to digitize as much material as we can. And we will take advantage of the Internet for making our resources available worldwide. We must do so to enable people far from our physical libraries to use and enjoy our holdings. Because we now have the means to extend the reach of our libraries, I think we also have a moral imperative to do so.

Additionally we will take advantage of digitization to help us meet preservation needs. We will recognize that providing access to digital copies

enables us to reduce use of, and thus preserve longer, our fragile originals. In that sense, digitization and preservation go together.

However, our plan also will recognize that we do not yet know how long we can fully preserve the material we are digitizing. Nor are we confident of our long-term ability to preserve material digitally created. Both kinds of digital material increase daily, and their preservation needs accordingly grow. The current wisdom seems to be that no "silver bullet"—no universal solution for digital preservation problems—will emerge. But we can progress by developing and refining different techniques for preserving different digital formats.

The need for multiple approaches makes it important for librarians to work together, as we do in the national programs I described. And libraries need to share with others the digital preservation advances they individually make. Thinking again of the early librarian whom I described, I hope we will not sit alone in our institutions, huddled in blankets against the chills of change, warming our feet on the soapstones of tradition. We now have the technological ability to operate far beyond our walls. Let us also cross over our walls to help each other do it.

References

Caroline R. Arms, 2000. "Keeping Memory Alive: Practices for Preserving Digital Content at the National Digital Library Program of the Library of Congress," RLG.

DigiNews, volume 4, number 3 (June), at www.rlg.org/legacy/preserv/diginews/diginews4-3.html, accessed 9 February 2007.

Helen Dalrymple, 2006. "Film and Sound Treasures in the Mountain Lair, Audiovisual Conservation Center Takes Shape in Culpeper," *The Library of Congress Information Bulletin,* volume 65, numbers 7–8 (July–August), pp. 167–171, also at www.loc.gov/loc/lcib/06078/navcc.html, accessed 13 February 2007.

Katherine Heidinger, 2006. "The Loneliest Buildings in Town," *Bangor Daily News* (12 October), p. C-2.

Library of Congress, 2004. "Office of Strategic Initiatives," *Annual Report of the Librarian of Congress for the Fiscal Year Ending September 30, 2003* (Washington, D.C.: Library of Congress), pp. 175–189.

—, 2006. "Library Partnership Supports Preservation of Foreign News Broadcasts," news release, 21 July, at www.loc.gov/today/pr/2006/06-147.html, accessed 9 February 2007.

—, 2007a. "Mission and History," American Memory program, at http://memory.loc.gov/ammem/about/index.html, accessed 9 February 2007.

—, 2007b. "About the [American Memory] Collections," at http://memory.loc.gov/ammem/about/about.html, accessed 9 February 2007.

—, 2007c. "Technical Information [on American Memory]: Preservation," at http://memory.loc.gov/ammem/about/techIn.html, accessed 9 February 2007.

—, 2007d. "$2 Million Sloan Foundation Grant to Help Digitize Thousands of Books," news release, 31 January, at www.loc.gov/today/pr/2007/07-020.html, accessed 9 February 2007.

—, n.d.a. "Library of Congress-National Science Foundation Digital Preservation Project Descriptions," at www.digitalpreservation.gov/partners/researchproj.html, accessed 9 February 2007.

—, n.d.b. "National Film Preservation Plan, Overview and Background," at www.loc.gov/film/implemen.html, accessed 9 February 2007.

—, n.d.c. "Redefining Film Preservation: A National Plan," at www.loc.gov/film/planpr.html, accessed 9 February 2007.

—, n.d.d. "Digital Audio-Visual Preservation Prototyping Projects," Web page at www.loc.gov/rr/mopic/avprot/avprhome.html, accessed 9 February 2007.

—, n.d.e. "Selection Criteria for Preservation Digital Reformatting," at www.loc.gov/preserv/prd/presdig/presselection.html, accessed 12 February 2007.

—, n.d.f. "Life-Cycle Management of Digital Data," at www.loc.gov/preserv/prd/presdig/preslifecycle.html, accessed 12 February 2007.

Library of Congress and National Recording Preservation Board, n.d. "Study on the Current State of Recorded Sound Preservation," at www.loc.gov/rr/record/nrpb/FedRegrev1024web.pdf, accessed 13 February 2007.

Library Services, 2006a. "Overview of Library Services," Strategic Plan, FY 2008–2013, Library Services, the Library of Congress, unpublished paper, draft 3.0, June, pp. 2–3.

—, 2006b. "Strategic Goal One," Strategic Plan, FY 2008-2013, Library Services, the Library of Congress, unpublished paper, draft 3.0, June, pp. 8–9.

National Endowment for the Humanities, n.d. "National Digital Newspaper Program," at www.neh.gov/projects/ndnp.html, accessed 9 February 2007.

Aliya Sternstein, 2006. "Hello, IRENE, Library of Congress to Use Image Tool for Sound Restoration," FCW.com News (13 February), at www.fcw.com/article92272-02-13-06-Print, accessed 10 February 2007.

Murtha Baca holds a PhD in art history and Italian language and literature from UCLA. She is Head of the Getty Vocabulary Program and the Digital Resource Management Department at the Getty Research Institute in Los Angeles. Her publications include *Introduction to Art Image Access* and *Introduction to Metadata*, and she is a member of the editorial team that produced *Cataloging Cultural Objects: a Guide to Describing Cultural Works and Their Images* (American Library Association Editions, 2006). Baca has taught workshops and seminars on metadata, visual resources cataloging, and thesaurus construction at museums, universities, and other organizations in North and South America and in Europe; she teaches a graduate seminar on metadata in the Department of Information Studies at UCLA.

Priscilla Caplan is Assistant Director for Digital Library Services at the Florida Center for Library Automation, where she is responsible for overseeing the Florida Digital Archive. She is the author of Metadata Fundamentals for All Librarians (ALA Editions, 2003) and numerous articles on metadata, digital libraries and digital preservation. She co-chaired the OCLC/RLG Working Group on Preservation Metadata: Implementation Strategies (PREMIS) and currently serves on the PREMIS Editorial Committee.

Erin Coburn is Manager of Collections Information at the J. Paul Getty Museum. She oversees the creation, management, and dissemination of data pertaining to the collection of the Getty Museum, and is responsible for ensuring access to information on the collection and related interpretative material through digital and multimedia initiatives. Cpburn is a member of the board of directors of the Museum Computer Network (MCN), a member of the Combined Committee of the American Library Association, Society of American Archivists, and American Association of Museums (CALM), and a member of the advisory committee for *Cataloging Cultural Objects: A Guide to Describing Cultural Objects and their Images* (Chicago: ALA Editions, 2006). Her publications include "Descriptive Metadata," co-authored with Murtha Baca and Sally Hubbard, in *Museum Informatics*, eds. Paul Marty and Kathy Jones (Taylor & Francis Group: forthcoming 2007) and "Beyond the Gallery Walls: Tools and Methods for Leading End-Users to Collections Information" co-authored with Murtha Baca in *ASIST Bulletin* online (June/July 2004).

Louis Fox is Associate Vice President of Computing & Communications at the University of Washington and a research professor in the Information School, where he has been for the last twenty-five years and has held numerous academic and administrative posts, all with obscure titles. Lacking hobbies, Fox also leads the *National Internet2 K20 Initiative*, which brings together Internet2 members (180 research institutions) with primary and secondary schools, colleges and universities, libraries, and museums to get new technologies—advanced networking tools, content, and applications—into the hands of innovators, across all educational sectors in the United States, as quickly and as "connectedly" as possible, and to connect these innovators to similar communities around the globe. Casting aside any last shreds of a normal life, Fox also leads technology initiatives for the Western Interstate Commission for Higher Education, based in Boulder, CO.

Valerie Glenn has been actively working to preserve born-digital government information for several years, with collections such as the University of North Texas' CyberCemetery and Congressional Research Service Reports Archive. Currently living in Alabama, she is active in the American Library Association, serving as the chair of the Government Documents Round Table's Government Information Technology Committee and as a member of the Committee on Legislation's Government Information Subcommittee.

Cathryn Goodwin manages Collections Information at the Princeton University Art Museum. A participant of the RLG Museum Collections Sharing Working Group, Cathryn has been involved in museum data standards and collections sharing initiatives throughout her 16 year career as a Museum Information Professional.

Mary Ide has been Director of the WGBH Archives since 1995. She has presented at conferences, and written about developing appraisal criteria for the selection of media for acquisition and preservation. Ide is a team member working on the National Digital Information Infrastructure Preservation Program (NDIIPP), which is investigating solutions for the long term preservation of digital media content. Ide is a member of the Society of American Archivists, past president of the New England Archivists and former member of the Board of Directors of the Association of Moving Image Archivists. She holds a MA in history from the University of Vermont and an MLIS from Simmons Graduate School of Library and Information Science.

Michael Jenkins is Manager of Met Images at The Metropolitan Museum of Art. The Met Images project seeks to preserve and protect the Metropolitan Museum's archive of images through the implementation of a secure centralized repository for the storage, management, and distribution of images. Prior to working on the Met Images initiative, Michael spent several years in the Met's Collections Management

department working on projects to improve access to information contained in various digital resources related to the Met's encyclopedic collection. Michael is a member of the steering committee of *steve*, an open source project investigating the usefulness of social tagging in the museum space (www.steve.museum) and is a participant in the RLG Collections Sharing Working Group.

Jennifer Locke Jones has worked in the Military History collections at the National Museum of American History since 1983. Having multiple positions as a researcher, collections manager and specialist, she became the Assistant Chair for the History of Technology from 1995 to 2004, and in 2004 was selected as the Chair for the Division of Military History and Diplomacy. She has worked as a curator on many exhibitions including most recently "The Price of Freedom: Americans at War", and "A More Perfect Union: Japanese Americans and the US Constitution" (including its traveling panel version, and an award winning web site).

Timothy J. Lorang is the Director for National Production Services and Participant Relations for ResearchChannel and the past Manager of Production for UWTV Productions at the University of Washington. Lorang is an accredited member of the Communication Media Managers Association and has been working in television production for over 30 years.

Nate McQueen is a systems architect working in the area of digital asset management for the ResearchChannel. His projects include DigitalWell, ResearchChannel.org and building infrastructure for streaming services at the University of Washington. His previous experience includes ground breaking work in Webcasting and Live streaming for the 96' Democratic National Convention and infrastructure management during the .com era.

Richard Rinehart is Digital Media Director and Adjunct Curator at the UC Berkeley Art Museum/Pacific Film Archive and a digital media artist. Rinehart has taught digital art studio and theory at UC Berkeley for six years and has also served as visiting faculty at the San Francisco Art Institute, UC Santa Cruz, San Francisco State University, Sonoma State University, and JFK University. Rinehart sits on the Executive Committee of the UC Berkeley Center for New Media and on the San Jose Airport Art Project Oversight Committee. Rinehart manages research projects in the area of digital culture, including 'Archiving the Avant Garde', a national consortium of museums and artists distilling the essence of digital art in order to preserve it. He also manages 'Museums and the Online Archive of California', a state-wide project bringing together museums with the archives and libraries across the state of California to provide standards-based access to cultural collections.

Marsha L. Semmel is the Director for Strategic Partnerships at IMLS, where she maintains oversight of federal-state partnership activities, initiates and implements collaborations with other federal agencies and organizations, and manages special projects and initiatives. As Deputy Director for Museums she manages the agency's museum grantmaking programs and is a key member of the executive team contributing to overall agency policy development. From 1998 to 2002, Semmel was President and CEO of the Women of the West Museum, in Denver, Colorado, where she was the first director of a new, multi-disciplinary museum with a mission to discover, explore and communicate the continuing role of women in shaping the American West. Previously, Semmel was President and CEO of Conner Prairie, a living history museum in Indianapolis that interprets the lives, attitudes and values of the early settlers in the Old Northwest Territory. From 1984 to 1996, Semmel worked at the National Endowment for the Humanities, in Washington, DC, serving as program officer; Assistant Director for Humanities Projects in Museums and Historical Organizations; and Director, Division of Public Programs. She began her museum career as curator and educator at the Taft Museum in Cincinnati, was deputy director of the B'nai B'rith National Jewish Museum in Washington, DC, and Program Coordinator for the Resident Associates Program at the Smithsonian Institution. In 1979, Semmel was a Fellow in the Museums Program of the National Endowment for the Arts.

Sarah Shreeves is currently the Coordinator for the Illinois Digital Environment for Access to Learning and Scholarship (IDEALS), UIUC's institutional repository. Her previous position was the Coordinator of the IMLS Digital Collections and Content project. Her past work and research has focused on the use of the OAI protocol and the need for shareable metadata. She is a co-editor of the *Best Practices for OAI Data Provider Implementations and Shareable Metadata* and was the past chair of the Metadata Working Group of the Digital Library Federation's Aquifer Initiative. Sarah has published and presented numerous times about the importance and implementation of shareable metadata.

Selma Thomas, is the founder and principal of Watertown Productions, Inc., a media design and production firm based in Washington, DC. A filmmaker with a background in history, Thomas produced several award-winning public television documentaries before

beginning her work with museums and libraries. She has designed and produced electronic programs, both site- and web-based, for a variety of cultural institutions. A partial client list includes: the Smithsonian Institution, the National Museum of American History, the National Gallery of Art, the Children's Discovery Museum of San Jose (CA), the Chicago History Museum, the Exploratorium, the Franklin Institute Science Museum and the Library of Congress. A frequent author and speaker on the strategic and interpretive uses of media, Thomas is Media Editor of *Curator: The Museum Journal* and co-editor of, and contributor to, *The Virtual and the Real* (an exploration of the interpretive role of media in museums). She is the author of "Private Memory in Public Spaces: Oral History in the Museum," in the upcoming *Oral History and Public Memories* (Temple University Press, 2007).

Barbara Thompson graduated from the Courtauld Institute, after which she worked for the *Witt Computer Index*, serving her apprenticeship in the emerging discipline of cataloguing works of art, using a complex, custom-built relational database. She worked with the *Index,* and related cataloguing projects, for seven years, moving along the corridor, in 1993, to become part of the Witt Library team; She became Witt Librarian in 2002. Collections' care and preservation issues, concerning core material, special collections and negatives are central to my present role in tandem with making the resources available to all. She is presently researching the histories of commercial art galleries in London, cataloguing the Witt Library archive papers and have a special interest in the history and conservation of photographic materials.

Liz Bishoff is Special Assistant to the Dean of Libraries and Head of the Office of Sponsored Programs, University of Colorado-Boulder Library. Previously she was Vice President for Digital Collection Services at OCLC, and former Executive Director of the Colorado Digitization Program. Bishoff has worked with libraries and museums in several states including Alabama, Kansas, South and North Carolina, Missouri, Minnesota, New Mexico, New York, and Tennessee on various aspects of their collaborative digitization initiatives. Liz led the development of collaborative best practices in metadata, including the Western States Metadata Dublin Core Best Practices.

Bishoff's current research interest involves library and cultural heritage institution's preparation for digital preservation. She has worked with the Northeast Document Conservation Center on the development of Digital Preservation Readiness Assessment, visiting museums, historical societies, and libraries of various sizes gathering information on the current status and trends in digital preservation trends and approaches.

Bishoff has been the program coordinator for WebWise 2006 and 2007, a speaker at several WebWise conferences, and a faculty member for the NEDCC School for Scanning program, Off the Wall program, and other NEDCC programs. She is a frequent speaker at digital library conferences and general library programs. She holds an MLS from Dominican University (formerly Rosary College), and has post-graduate work in public administration at Roosevelt University.

Brett Bobley serves as the Chief Information Officer of NEH. Reporting directly to the NEH Chairman, Bobley is the agency's senior advisor on all issues pertaining to technology. Bobley developed the agency's current enterprise architecture, which has made NEH a leader in e-government. Four years ago, all of NEH's grant applications arrived on paper. Today, nearly 100% of them arrive electronically via the Web. Bobley has been recognized numerous times by the Grants.Gov Project Management Office for the work his agency has done to promote the adoption of electronic grants government-wide.

Bobley is also the Director of NEH's Digital Humanities Initiative (DHI). Under DHI, Bobley has helped to launch five new grant programs designed to spur innovation in the area of digital humanities. Bobley is also the co-chairman of the Small Agency CIO Council which is a forum for the nearly 100 small agency CIOs to meet, share best practices, and discuss technology policy issues. He also sits on the federal CIO Council representing small agency concerns.

Prior to his current assignment, Bobley served as the Chief of Systems Operations for U.S. Coast Guard Headquarters. In that capacity, he played a critical role in the design, testing, and deployment of the Coast Guard's Standard Workstation III. As part of that project, Bobley helped put into place innovative deployment and maintenance practices, which greatly reduced the total cost of ownership for the Coast Guard. Many of those innovative practices are now used government-wide. In 1989, Bobley worked as a Systems Analyst for Unisys Corporation, supporting IT systems for the government.

Bobley has a master's degree in computer engineering from the Johns Hopkins University. He has a bachelor's degree in philosophy from the University of Chicago.

Elizabeth Broun is responsible for the nation's premier collection of American art, as well as major exhibition, research, publication, education, and new media programs. During Broun's tenure, the museum has become a leader in providing electronic resources to schools and the public, and in developing a national education program. In addition, Broun conceived and secured funding for many of the museum's core programs and new public spaces—a conservation center, an art storage and study center, and an auditorium—in the museum's main building, a magnificently renovated National Historic Landmark located in the heart of a revitalized downtown cultural district. The innovative Lunder Conservation Center is the first art conservation facility that allows the public permanent behind-the-scenes views of the preservation work of museums. The Luce Foundation Center for American Art is the only visible art storage and study center in Washington. The Nan Tucker McEvoy Auditorium, a 346-seat space equipped with a state-of-the-art sound system, is the first such facility for the museum.

Broun lectures extensively across the United States. Her research interests include contemporary art, nineteenth century art, and prints and drawings. Her 1989 exhibition catalogue on Albert Pinkham Ryder won the prestigious Alfred H. Barr Award for Distinguished Scholarship. She has also curated exhibitions and published on the art of Thomas Hart Benton, Stuart Davis, Childe Hassam, Patrick Ireland, Pat Steir and James McNeill Whistler.

Broun came to Washington in 1983 as chief curator and assistant director of the museum, following seven years as a curator and interim director at the Spencer Museum of Art, University of Kansas in Lawrence. She has served as director of the Smithsonian American Art Museum and its branch museum, the Renwick Gallery, since August 1989.

Broun earned a doctorate (1976) in art history at the University of Kansas for her work on American

art exhibited at the 1893 Chicago World's Fair. She also holds a Certificate of Advanced Study from the University of Bordeaux, France.

Laura Elizabeth Campbell was appointed to the position of associate librarian for Strategic Initiatives in 2000 by the Librarian of Congress, James H. Billington. He noted that "Laura has already made many significant contributions to the realization of the Library's digital future. Her demonstrated leadership, technical knowledge and network of expert colleagues in the public and private sectors will bring added strength to our decision-making process."

Creation of the position of associate librarian for Strategic Initiatives responds to a recommendation contained within the 2000 National Academy of Sciences report *LC21: A Digital Strategy for the Library of Congress.* Responsibilities of the position include overall strategic planning for the Library, oversight of the Information Technology Services directorate, and leadership of the $100 million National Digital Information Infrastructure and Preservation Program (www.digitalpreservation.gov), which was established to build a nationwide network of partners to preserve important digital content that is at risk of being lost. Campbell is also director of the National Digital Library Program, a cooperative national effort to digitize and make available online primary source materials of American history and culture. The program's award-winning Web site, American Memory (memory.loc.gov), offers more than 10.5 million items from the collections of the Library of Congress and those of its partners.

Campbell is a graduate of Pennsylvania State University (BA, 1973), the University of Maine (MA in management, 1979) and Georgetown University (MS in accounting, 1983).

Audrey Christensen joined the Exit Art staff in October 2006 as Archives Manager. From 2004 to 2006 she was the Digital Archivist at Pentagram Design, and spent three years at the Museum of Modern Art Library previous to that. Christensen was closely involved in the production of the first Exit Art exhibition of 2007, *Renegades: 25 Years of Performance at Exit Art, a selection from the Archives.* She has a BA in art from University of Northern Iowa and an MS in information science from Pratt Institute of Art.

Thomas F.R. Clareson joined PALINET as Program Director for New Initiatives in October 2005. Leading PALINET's digital collections creation and management services, preservation services, and consulting activities, he is responsible for establishing new services and funding sources, grant writing, and outreach to the museum and historical society communities. With over 15 years' experience in preservation and digitization services, Clareson was previously Global Product Manager at OCLC Online Computer Library Center, Inc.; he also served in various capacities at Amigos Library Services, Inc. He holds an MLS from Kent State University, an MA from Ohio State University, and a BA from Ohio Wesleyan University. Currently a representative from the Society of American Archivists to the Joint Committee on Archives, Libraries, and Museums, he also serves on the board of trustees of Heritage Preservation.

Robin L. Dale is a program officer in RLG Programs, a part of OCLC's Office of Programs & Research. Previously, she was a program officer at RLG for over nine years. Until February, Dale was also the Project Director of the Center for Research Libraries Auditing and Certification of Digital Archives project, a Mellon-funded activity to develop processes to audit and certify digital archives and repositories. Her current work focuses on data curation, mass digitization, scholarly communications, and cooperative storage. She co-chaired the RLG-National Archives and Records Administration Digital Repository Certification task force, which produced the recently released *Trusted Repositories, Audit and Certification: A Checklist (TRAC)*. A regular speaker on digital preservation initiatives, she is active in digital preservation standards and best-practice building activities, including the development of the Open Archival Information System (OAIS) international standard and various preservation metadata best practices.

Anne Graham is a Senior Computer Specialist in the Digital Initiatives unit of the University of Washington Libraries. Graham has managed several digitization grants and projects from a variety of federal, community, and university sources. With a background in IT and databases, she also maintains the university's installation of CONTENTdm, the digital collection management software, which holds over 160,000 images and digital objects (http://content.lib.washington.edu).

Chris Greer received his PhD in biochemistry from the University of California, Berkeley, did postdoctoral work at CalTech, and was a tenured faculty member at the University of California, Irvine, before joining the National Science Foundation. He is currently Program Director with responsibility for digital data activities in the Office of Cyberinfrastructure. Dr. Greer recently served as Executive Secretary for the Long-lived Digital Data Collections Activities of the National Science Board and is currently Co-Chair of the Digital Data Interagency Working Group of the National Science and Technology Council's Committee on Science.

Kenneth Hamma is Executive Director for Digital Policy and Initiatives at the J. Paul Getty Trust. From 1996 to 2004 he was Assistant Director and from 1987 to 1996 Associate Curator of Antiquities at the Getty Museum. He currently serves as a member of the RLG Programs board at OCLC; member of the Steering Committee of the Coalition for Networked Information; and director of the Museum Domain Management Association, the sponsor of the museum Internet TLD. He has served as a board member for the Art Museum Image Consortium, the Consortium for the Interchange of Museum Information, and the National Initiative for Networked Cultural Heritage.

Jodi Hanel has been Exit Art's Associate Curator since 1997. She has coordinated and co-curated more than 40 exhibition projects, specializing in the work of young and emerging artists and graphic design. Most recently she coordinated The Drop, an exhibition and public program project that explored the contentious role of water and the environment; Terrorvision, an interdisciplinary exhibition that examined definitions of terror in today's society, and L Factor, an exhibition and expansive series of screenings, music, public forums, and performances that explored the work of a young generation of Latino artists.

Jay Jordan became the fourth president in OCLC's 38-year history in May 1998. He came to OCLC after a 24-year career with Information Handling Services, an international publisher of databases, where he held a series of key positions in top management, including president of IHS Engineering.

Jordan graduated from Colgate University in 1965 with a BA in English literature and served as a U.S. Army officer in Germany. He has spent more than seven years living and working outside the United States. He is active in professional organizations, including the American Library Association and the Special Libraries Association. He is a Fellow of the Standards Engineering Society.

He is a member of the Board of Visitors of the School of Information and Library Science, the University of North Carolina at Chapel Hill, and a member of the editorial board of the *Journal of Library Administration*. He is also a member of the Louis Round Wilson Academy and Knowledge Trust at the University of North Carolina at Chapel Hill.

Under Jordan's leadership, OCLC has built a new technological platform, introduced new services, created a library advocacy program, and introduced new initiatives to make library holdings and libraries more visible on the open Web. He has overseen a period of remarkable growth for OCLC. Since 1998, the number of libraries participating in the OCLC cooperative has grown from 30,000 to 57,000. The number of participating institutions outside the U.S. has increased from 3,200 in 64 countries to more than 11,000 in 110 countries.

Mark L. Louden received his undergraduate and graduate training in linguistics, with a focus on Germanic languages, at Cornell University. From 1988 to 2000 he was on the faculty of the University of Texas at Austin. In 2000 he accepted an appointment at the University of Wisconsin–Madison, where he is currently a professor of German. From 2002 to 2006 he co-directed the Max Kade Institute (MKI) for German-American Studies at Madison. He continues to direct MKI's North American German Dialect Archive. A fluent speaker of Pennsylvania Dutch, Prof. Louden's research has dealt heavily with structural and sociolinguistic aspects of German dialects spoken in the United States, which, along with regional varieties of English, comprise an important part of the IMLS-funded "American Languages" project.

Clifford Lynch has been the Director of the Coalition for Networked Information (CNI) since July 1997. CNI, jointly sponsored by the Association of Research Libraries and Educause, includes about 200 member organizations concerned with the use of information technology and networked information to enhance scholarship and intellectual productivity. Prior to joining CNI, Lynch spent 18 years at the University of California Office of the President, the last 10 as Director of Library Automation. Lynch, who holds a PhD in computer science from the University of California, Berkeley, is an adjunct professor at Berkeley's School of Information. He is a past president of the American Society for Information Science and a fellow of the American Association for the Advancement of Science and the National Information Standards Organization. Lynch serves on the National Digital Preservation Strategy Advisory Board of the Library of Congress; he was a member of the National Research Council committees that published *The Digital Dilemma: Intellectual Property in the Information Infrastructure* and *Broadband: Bringing Home the Bits*, and now serves on NRC's committee on digital archiving and the National Archives and Records Administration.

Deanna Marcum was appointed Associate Librarian for Library Services on August 11, 2003. In this capacity she manages 53 divisions and offices whose 2,400 employees are responsible for acquisitions, cataloging, public service, and preservation activities, services to the blind and physically handicapped, and network and bibliographic standards for America's national library. She is also responsible for integrating the emerging digital resources into the traditional

artifactual library—the first step toward building a national digital library for the 21st century.

In 1995, Dr. Marcum was appointed president of the Council on Library Resources and president of the Commission on Preservation and Access. She oversaw the merger of these two organizations into the Council on Library and Information Resources (CLIR) in 1997 and served as president until August 2003. CLIR's mission is to identify the critical issues that affect the welfare and prospects of libraries and archives and the constituencies they serve, convene individuals and organizations in the best position to engage these issues and respond to them, and encourage institutions to work collaboratively to achieve and manage change.

Dr. Marcum served as Director of Public Service and Collection Management at the Library of Congress from 1993–95. Before that she was the Dean of the School of Library and Information Science at the Catholic University of America. From 1980 to 1989, she was first a program officer and then vice president of the Council on Library Resources. Dr. Marcum holds a PhD in American studies, a master's degree in library science, and a bachelor's degree in English.

Sue O. Medina, Director of the Network of Alabama Academic Libraries (NAAL), works with outstanding academic librarians to ensure that Alabama students and faculty members have the highest-quality library resources and services they may need for their instruction and research. NAAL participants include all of Alabama's public and private nonprofit four-year colleges and universities. While Alabama libraries traditionally share poverty better than anything else, NAAL cooperative programs have significantly improved the delivery of library services and resources statewide. With funding assistance from IMLS, NAAL is encouraging development of a portal and digital collection for Alabama history at www.alabamamosaic.org.

Medina earned her PhD in library science from Florida State University. Her advanced studies focused on organizational role and change. She also has an MS in library science and a BA in history.

Kristen Overbeck Laise is the Vice President for Collections Care Programs at Heritage Preservation, a national, nonprofit organization that advocates for collections. She directed the Heritage Health Index, the first comprehensive survey of the condition and preservation needs of U.S. collections. The survey, which released its results in December 2005, was coordinated by Heritage Preservation in partnership with IMLS with funding from the Getty Foundation and other private foundations. Laise is currently directing another national initiative, Rescue Public Murals. Previously, she coordinated the Conservation Assessment Program, a technical assistance program for small museums administered by Heritage Preservation in cooperation with IMLS. She holds a BA in history from Earlham College and an MA in art history from the University of Wisconsin–Madison, where she worked with the History of Cartography Project.

Steven Puglia works as a Preservation and Imaging Specialist, managing the Digital Imaging Lab at the National Archives and Records Administration (NARA). Puglia has worked at NARA for 18 years, he began working in the field of preservation 22 years ago as a technical photographer duplicating historic negative collections, and he has worked with digital imaging for over 15 years. He is involved with the ISO standards committee working on the stability of imaging materials, and is chair of the task group dealing with the stability of color pictorial images. Puglia lectures frequently at conferences on digital imaging, traditional and digital preservation, and related topics.

Anne-Imelda M. Radice has a strong record of public service. She was most recently Acting Assistant Chairman for Programs at the National Endowment for the Humanities. Before joining the National Endowment for the Humanities, Dr. Radice served as Chief of Staff to the Secretary of the United States Department of Education. In the early 1990s she served as the Acting Chairman and Senior Deputy Chairman of the National Endowment for the Arts. From 1989 to 1991 Dr. Radice was Chief of the Creative Arts Division of the United States Information Agency. She was the first Director of the National Museum of Women in the Arts (1983–1989). Dr. Radice has also been Curator and Architectural Historian for the Architect of the Capitol and an Assistant Curator at the National Gallery of Art.

Joyce Ray is Associate Deputy Director for Library Services at the Institute of Museum and Library Services. She is responsible for agency-wide digital initiatives—such as the annual WebWise Conference—as well as for discretionary library grant programs, including the National Leadership, Laura Bush 21st Century Librarians, Native American Library Services, and Native Hawaiian Library Services programs. These programs collectively distribute approximately $40 million per year to libraries, institutions of higher education, and other organizations throughout the U.S. Before joining IMLS in 1997, she was head of special collections at the University of Texas Health Science Center at San Antonio and also held positions at the National Archives and Records Administration and the National Historical

Publications and Records Commission. An archivist by training, she has a PhD in American history and a master's degree in library science, both from the University of Texas at Austin. She currently serves on the program committee of the Joint Conference on Digital Libraries.

Roy Rosenzweig is Mark and Barbara Fried Professor of History and New Media at George Mason University, where he also heads the Center on History and New Media. He is the author, co-author, or co-editor of several books, including *The Park and the People: A History of Central Park; The Presence of the Past: Popular Uses of History in American Life; Eight Hours for What We Will: Workers and Leisure in an Industrial City, 1870-1920; History Museums in the United States: A Critical Assessment; Presenting the Past: Essays on History and the Public, A Companion to Post-1945 America;* and *Digital History: A Guide to Gathering, Presenting, and Preserving the Past on the Web.* In 2003, he won the Richard W. Lyman Award (awarded by the National Humanities Center and the Rockefeller Foundation) for "outstanding achievement in the use of information technology to advance scholarship and teaching in the humanities."

Ann Russell has served as Executive Director of the Northeast Document Conservation Center since 1978. She received her undergraduate degree from Harvard University and holds a PhD in English literature from Brandeis University. She has served on the boards of directors of Heritage Preservation and the Intermuseum Conservation Association in Cleveland, Ohio, as well as on preservation committees of the American Library Association and the Society of American Archivists. She currently serves as Chair of the Association of Regional Conservation Centers and as Treasurer of the Society of American Archivists. She has organized conservator exchanges and training programs in Russia, Mongolia, Cuba, South Africa, Latin America, and Central Europe. She has written two books and numerous articles on preservation.

Jane Sledge is a thoughtful visionary in the area of museum information resources. She has worked with the Canadian Heritage Information Network, the Smithsonian Institution, the UNESCO-ICOM Museum Information Center in Paris, and the Getty Information Institute, and is now with the National Museum of the American Indian (NMAI). As Associate Director for Museum Assets and Operations, her areas of responsibility cover the overall management of NMAI's physical and intellectual assets and include (1) stewardship for the care and management of NMAI's collection of more than one million artifacts, photographs, media, and archival materials; (2) technology systems and support of technology-based operations at NMAI, including exhibitions, Web, Intranet, and application systems documenting the cultural experience and life of Native peoples of the Western Hemisphere and providing automated support for a variety of museum activities; (3) buildings and facilities management including safety and emergency preparedness. Sledge has management expertise in project management, program administration, content engineering, and large-scale collaborative projects. She enjoys working with staff in museums and cultural not-for-profit organizations to apply innovative solutions to solve challenging problems.

Kenneth Thibodeau is Director of the Electronic Records Archives (ERA) Program at NARA. ERA is NARA's strategic initiative to preserve all types of electronic records and deliver them to future generations of users on future generations of technology. Dr. Thibodeau has over 30 years' experience in archives and records management, and is an internationally recognized expert in electronic records. He has served as Chief of the Records Management Branch of the National Institutes of Health, Director of the Center for Electronic Records at NARA, and Director of the Department of Defense Records Management Task Force. He earned a PhD in the history and sociology of science from the University of Pennsylvania, and held several post-doctoral fellowships in computer science. He has been a visiting professor and lecturer at universities in the U.S., Canada, France, Germany, Italy, and the U.K. A Fellow of the Society of American Archivists, he has published over 30 papers and spoken at more than 150 conferences around the world.

Günter Waibel is a Program Officer in the OCLC Programs and Research division. He specializes in standards for describing and representing cultural materials in a networked environment, as well as the intersection of museums, libraries, and archives in providing access to primary materials. Further areas of interests are digital asset management and digital preservation. Waibel serves as the RLG Programs liaison to the museum and art library community. Before joining RLG in 2003, he worked at the UC Berkeley Art Museum & Pacific Film Archive as well as the Oakland Museum of California.

Waibel is a board member of the Museum Computer Network and the Association of American Museum's Media & Technology Committee. He recently guest-edited a special issue of *RLG DigiNews* entitled "Managing digital assets in U.S. museums" (December 2006), and blogs at www.hangingtogether.org. Since 2004, he has been teaching as adjunct faculty in the School of Information Studies at Syracuse University, New York.

Steven C. Wheatley is the Vice President of the American Council of Learned Societies (ACLS). He holds a BA from Columbia University and MA and PhD degrees in history from the University of Chicago. He is the author of, among other works, *The Politics of Philanthropy: Abraham Flexner and Medical Education* (University of Wisconsin Press, 1988) and a new introduction to Raymond Fosdick's *The Story of the Rockefeller Foundation* (Transaction Books, 1988), and the editor (with Katz, Greenberg and Oliviero) of *Constitutionalism and Democracy: Transitions in the Contemporary World* (Oxford University Press, 1993). He has served as a consultant to the Ford Foundation, the Carnegie Corporation of New York and the Lilly Endowment, Inc., and as a member of the Task Force on the Artifact of the Council on Library and Information Resources. He was for eight years a member of the Academic Advisory Council of the Rockefeller Archive Center of Rockefeller University, and has taught at the University of Chicago and at New York University, where he was appointed an Adjunct Professor. In 2005-06, he was staff to and an adviser of the ACLS Commission on Cyberinfrastructure for the Humanities and Social Sciences.

Diane M. Zorich consults for cultural organizations on information management issues. Her clients include the J. Paul Getty Trust, the American Association of Museums, the Council on Library and Information Resources, and numerous cultural heritage institutions. Before establishing her consultancy, she was data manager at the former Association of Systematics Collections in Washington, D.C., and documentation manager at the Peabody Museum of Archaeology and Ethnology at Harvard University. She was past president and board member of the Museum Computer Network, and chaired that organization's Intellectual Property Special Interest Group.

Zorich is the author of *Introduction to Managing Digital Assets: Options for Cultural and Educational Organizations* (1999, The J. Paul Getty Trust), *Developing Intellectual Property Policies: A "How-To" Guide for Museums* (2003, Canadian Heritage Information Network), and *A Survey of Digital Cultural Heritage Initiatives and Their Sustainability Concerns* (2003, Council on Library and Information Resources). Her latest publication on information policies in museums will appear in *Museum Informatics* (Routledge, 2007). She presently serves as project manager for *Cataloging Cultural Objects: A Guide to Describing Cultural Works and their Images*, a project of the Visual Resources Association.

University of North Carolina at Chapel Hill's School of Information and Library Science—Chapel Hill, NC

Folkstreams

Building on a previous IMLS grant, UNC's ibiblio.org digital library will archive, digitize, and video stream and will add additional documentary films to the Folkstreams.net Web site by transferring them from 16 mm format to digital betacam. Partnering with Folkstreams, Inc., and the University's Southern Folklore Collection, the project will produce a guide to best practices in video digitization and expand its Video Aids for Film Preservation. The project has also produced a multimedia Web site demonstration highlighting the skills required for transfers from 16 mm film to digital formats.

Contact:
Paul Jones
Director, ibiblio.org
Clinical Associate Professor
School of Journalism and Mass Communication and
School of Information and Library Science
University of North Carolina at Chapel Hill
Manning Hall, Room 213
Chapel Hill, NC 27599-3360
Phone: 919-962-7600
Fax: 919-962-8071
E-mail: paul_jones@unc.edu
Project Web site: www.folkstreams.net

North Carolina Aquarium/Roanoke Island—Manteo, NC

Turtle Trails

This unusual project used digital technology to help preserve living species and to increase the public's understanding and appreciation of species preservation. The project team used satellite telemetry to track rehabilitated cold-stunned juvenile loggerhead sea turtles (Caretta caretta) to assess post release survival and behavior. It disseminated project results to the public via the Internet.

Contact:
Joanne E. Harcke
Conservation and Research Coordinator
North Carolina Aquarium at Fort Fisher
900 Loggerhead Road
Kure Beach, NC 28449
Phone: 910-458-8257, ext. 237
Fax: 910-458-6812
E-mail: joanne.harcke@ncmail.net
Project Web site: www.ncaquariums.com/turtletrails

Northeast Document Conservation Center—Andover, MA

dPlan

The Northeast Document Conservation Center (NEDCC), in partnership with the Massachusetts Board of Library Commissioners, has created a free, online disaster plan template for cultural heritage institutions. dPlan generates a customized, updatable plan that contains contact information for staff and key personnel, preventive maintenance checklists, salvage techniques, and more. NEDCC is currently working on a tool for statewide disaster planning for cultural heritage institutions, including libraries, archives, and museums.

Contact:
Lori Foley
Northeast Document Conservation Center
100 Brickstone Square
Andover, MA 01810-1494
Phone: 978-470-1010
Fax: 978-475-6021
E-mail: lfoley@nedcc.org
Project Web site: www.dplan.org

Nebraska State Historical Society—Lincoln, NE

Saving Nebraska's Treasures

Building on widespread interest in the public television show *Antiques Roadshow,* the Nebraska State Historical Society, Nebraska Library Commission, and statewide public television network Nebraska Educational Telecommunications (NET) are showing families and museum and library staff how to preserve their treasures. Professional conservators offered information and advice about family heirlooms at three community-based conservation clinics. NET produced and broadcast a television program featuring the clinics and visits to the Gerald R. Ford Conservation Center in Omaha to illustrate more detailed conservation techniques. The conservators presented four in-depth videoconferences for staff and volunteers of the more than 600 not-for-profit museums and libraries in the state. Through these activities and a complementary Web site, the project will help Nebraskans save personal and public historical materials, educate staff in cultural heritage institutions in conservation techniques, and increase community knowledge and appreciation of their heritage.

Contact:
Lynne Ireland, Chief Education and Research Officer
Nebraska State Historical Society
1500 R Street
PO Box 82554
Lincoln, NE 68501
Phone: 402-471-4758
Fax: 402-471-3100
E-mail: lireland@nebraskahistory.org
Project Web site: www.nebraskahistory.org/oversite/whatsnew/save_treasures.htm

Indiana University—Bloomington, IN

Variations3

Information technology has become an essential part of how music libraries deliver services and collections to music students and faculty. Over the past four years Indiana University (IU) has developed an experimental digital music library system known as Variations2. Building on IU's past experience in creating the original Variations, one of the world's first digital music library systems, Variations2 provides a complete environment at IU in which students and faculty can discover, listen to, view, annotate, and interact with music in both sound and score form. This project will create Variations3, a turnkey digital music library and learning system that can be easily deployed at a wide range of college and university libraries with minimal technical support and at minimal cost to the institutions.

Contact:
Jon Dunn, Associate Director for Technology
Digital Library Program
Wells Library E170
1320 E. 10th Street
Bloomington, IN 47405
Phone: 812-855-0953
Fax: 812-856-2062
E-mail: jwd@indiana.edu
Project Web site: www.dlib.indiana.edu/projects/variations3

University of California, Santa Barbara, Library—Santa Barbara, CA

Cylinder Preservation and Digitization Project

The University of California, Santa Barbara, Library has digitized 6,000 wax cylinder recordings from its collection and made these resources available online. The project has also developed a model for digitizing historic recordings on older formats such as cylinders, 78 rpm recordings, and unique acetate recordings, for delivery via the Internet.

Contact:
David Seubert, Curator of the Performing Arts Collection
Davidson Library
University of California, Santa Barbara
Building 525, 3rd Floor, MC 9010
Santa Barbara, CA 93106-9010
Phone: 805-893-5444
Fax: 805-893-5749
E-mail: seubert@library.ucsb.edu
Project Web site: http://cylinders.library.ucsb.edu

Autry National Center—Los Angeles, CA

Electronic Cataloging Initiative

The Electronic Cataloging Initiative (ECI) of the Autry National Center seeks to create, update, or enhance records related to artifacts and archival materials in the permanent collections; to create digital images of these materials; and to make this data available to scholars, students and teachers, and the general public over the Internet and through kiosk terminals on-site at the center. Advancement of the ECI is also enhancing the ability of Autry curatorial staff to research and organize future exhibitions. This project specifically targets materials held in the Southwest Museum and Braun Research Library collections, combining information about these collections with information about the collections at the Museum of the American West and the Autry Library. The ECI is an essential component in the center's comprehensive efforts to conserve, protect, interpret, and create broad public access to the Southwest Museum's important collection of Native American and American Southwest material.

Contact:

Rebecca Menendez
Project Manager, Electronic Cataloging Initiative
Collections Management Department
Autry National Center
4700 Western Heritage Way
Los Angeles, CA 90027
Phone: 323-667-2000 x 201
Fax: 323-663-4435
E-mail: rmenendez@autrynationalcenter.org
Project Web site: www.autrynationalcenter.org/collections

Michigan State University Museum—East Lansing, MI

The Quilt Index

This project is developing a Quilt Index as an innovative national model for distributed online management and presentation of thematic collections for museums and libraries. The three main goals are to (1) create a critical mass of quilt objects and information online; (2) enhance online access to the U.S. quilt and quilt information collections in museums, libraries, and archives through improved content management and interoperability; and (3) enhance the value, usefulness, and relevance of the index's thematic presentation. The model for distributed collections development around a theme can be applied to many different cultural heritage and natural resource areas, from fossil types to historic toy objects. The index's innovative design pilots a distributed system for entering customized local data that can be replicated locally and shared globally. For libraries, museums, exhibitors, and collections with object-specific foci, the project will offer a model and a road map for creating an online preservation, management, and presentation system.

Contacts:

Marsha MacDowell, PhD
Curator, Michigan State University Museum and Professor, Michigan State University
The Quilt Index, Co-Principal Investigator
West Circle Drive
Michigan State University Museum
East Lansing, MI 48824-1045
Phone: 517-355-2370
E-mail: macdowel@msu.edu

Mark Kornbluh, PhD
Professor and Chairperson, Department of History, Michigan State University
Director, MATRIX: Center for Humane Arts, Letters, and Social Sciences OnLine
The Quilt Index, Co-Principal Investigator
310 Auditorium Bldg.
Michigan State University
East Lansing, MI 48824-1120
Phone: 517-355-9300
E-mail: kornbluh@msu.edu

Mary Worrall
Assistant Curator, Michigan State University Museum
The Quilt Index, Project Manager
West Circle Drive
Michigan State University Museum
East Lansing, MI 48824-1045
Phone: 517-355-2370
E-mail: worrall@msu.edu

Florida, Division of Library and Information Science, Bureau of Archives and Records Management—Tallahassee, FL

The Florida Folklife Digitization and Education Project

This project has digitized 12,000 images and created an index to 52,000 images and over 6,000 sound recordings from the Florida Folklife Collection documenting performances by, interviews with, and fieldwork surveys of folk musicians, craftspersons, storytellers, folklife interpreters, and cultural tradition-bearers in such areas as children's lore, foodways, religious traditions, Native American culture, maritime traditions, ethnic folk culture, material culture, and occupational lore.

Contact:
Joanna Norman
500 South Bronough Street
Mail Station 9A
Tallahassee, FL 32399-0250
Phone: 850-245-6700
Fax: 850-488-4894
E-mail: jnorman@mail.dos.state.fl.us
Project Web site: www.floridamemory.com/Collections/folklife

Monticello/Thomas Jefferson Foundation—Charlottesville, VA

Thomas Jefferson's Libraries

In this project, all available information on the books of Thomas Jefferson (much of it in Jefferson's own hand) is being analyzed to create a complete inventory and annotated bibliographic database of his book collections including 5,000 titles that he sold to Congress in 1815 and 4,000 additional titles. The contents of 12 different lists will be available worldwide for searching through the Thomas Jefferson portal online catalog, enabling correlations between Jefferson's thinking and writing and the vast array of published material that was part of his known world. The principal activities include editing and transcribing manuscript sources, compiling full bibliographic records, and enhancing the title-by-title information with transcription, commentary, classifications, and references to current holdings.

Contact:
Jack Robertson
Foundation Librarian
Jefferson Library, Thomas Jefferson Foundation
P.O. Box 316
Charlottesville, VA 22902
Phone: 434-984-7545
E-mail: jrobertson@monticello.org
Project Web site: www.monticello.org/library

University of Southern California—Los Angeles, CA

The West Semitic Research and InscriptiFact Projects

The University of Southern California has formed a partnership with Cultural Heritage Imaging to develop a tool for three-dimensional, multiview representation of cultural objects that will be downloadable and available over the Internet. Improving on an earlier version that visualized only one surface of an object, the new three-dimensional tool will be easily used by almost any museum. The target audience includes museums of all sizes, scholars and students of material culture, cultural heritage professionals, and the interested public. The project will also produce the complete process history for each digital object, enabling replication by scholars. It has the potential to set a new standard of best practice for digital representations of cultural heritage objects.

Contact:
Dr. Bruce E. Zuckerman
Director, West Semitic Research and InscriptiFact Projects Director, Archaeological Research Collection
University of Southern California Ahmanson Center
130 Los Angeles, CA 90089-1481
Phone: 213-740-0271
E-mail: bzuckerm@usc.edu

Renaissance Society—Chicago, IL

The Renaissance Society Online Exhibition Archive: 1915–present

The Renaissance Society, founded in 1915, is one of the nation's longest-running museums devoted to art of the current moment. As a noncollecting institution, the society offers a lasting contribution to art history through the documentation of its exhibitions, which trace the development of contemporary art through every major movement in the last century. The society's current digital archive project uses Internet technologies to expand the museum's Web site into a vivid public archive, providing unprecedented access to archival images, essays, and other descriptive materials from the entire programming history. Project activities accomplished so far include the development of a searchable database platform on the Web site, which allows visitors to research information by artist name, exhibition title, date, or artistic media. Photographic and textual archives for the last 174 exhibitions dating from 1971 to the present have been digitized and are now available online, including more than 2,800 images, 109 essays, and four audio-visual clips of video-taped artist talks. In the next phases of the project, The Society will focus on adding more audio-visual components and integrating archival material from 1915 to 1970.

Contact:
Lori Bartman
Director of Development
The Renaissance Society
5811 South Ellis Avenue 4th Floor
Chicago, IL 60637
Phone: 773-702-8670
Fax: 773-702-9669
E-mail: lbartman@renaissancesociety.org
Project Web site: www.renaissancesociety.org

aggregators—service providers who gather metadata records and make them available for others to gather.

checksums—values generated by an algorithm and used to ensure data are stored or transmitted without error.

dark archive—in digital preservation lingo, refers to data storage that cannot be accessed by users. Its purpose is to function as a fail-safe repository that can be used for data recovery in the event of a disaster.

emulation—the use of a program or device to imitate the behavior of another program or device.

format standardization—transforming a variety of data types to a single, standard type in order to preserve a resource. A frequently cited example is preserving a word processed file in ASCII format.

harvesting—gathering metadata from distributed resources into a combined data store.

media refreshment—the periodic transfer of data from an existing tape/disk to a new tape/disk.

migration—in a digital preservation context, the transfer of data from older media to newer media.

software engineering—in a preservation context, creating or using various software methods to provide a simpler way of keeping an obsolete format accessible. If, for example, the only functionality that needs to be preserved for a particular digital resource is that its content be viewable, then the only software needed is one that renders the content with its original "look and feel."

INSTITUTE OF MUSEUM AND LIBRARY SERVICES
1800 M Street NW, 9th Floor
Washington, DC 20036
202-653-IMLS (4657)
www.imls.gov