

**Library of Congress**  
**Report to the Digital Library Federation**  
**Spring, 2004**

**TABLE OF CONTENTS**

Collections, Services, and Systems  
Projects and Programs  
Specific Digital Library Challenges  
Digital Library publications, policies, working papers, and other documents

**I. Collections, Services, and Systems**

**A. COLLECTIONS**

**National Digital Library Program**

Since its beginning in 1994, the Library of Congress National Digital Library Program (NDLP) has been a premier national resource for noncommercial, high-quality intellectual content on the Internet. More than 8.5 million items from the Library's collections are now available to users anywhere anytime from the Library of Congress Web site at <http://www.loc.gov>. The site, one of the federal government's most popular, handles more than 100 million transactions every month, including 47 million visits each month to the American Memory collections.  
<http://memory.loc.gov>

**1. America's Library**

America's Library is a Web site designed for children and families. The site combines kid-friendly graphics with the incomparable American historical collections of the Library in more than 1,000 stories about our nation's past. America's Library invites users to learn through multimedia materials that have rarely been on view. These materials include cartoons, baseball cards, early "special effects" captured on film, photographs, maps, prints, manuscripts, and audio and video recordings. Interactive elements such as a "Scavenger Hunt" and "Send a Postcard" encourage exploration of the site, and animated "teasers" on the main home page will delight users of all ages. Questions invite children to talk to their family and friends about what they have learned. During fiscal year 2003 America's Library received an average of more than 15 million visits per month. Individual collections include: Meet Amazing Americans; Jump Back in Time; Explore the States; Join America at Play; and See, Hear and Sing.  
<http://www.americaslibrary.gov>

**2. American Memory Historical Collections**

The American Memory Web site continued to receive wide acclaim as a provider of free, high quality, educationally valuable American cultural and historical resources. At the end of 2003, more than 8.5 million items digitized from the collections of the Library of Congress or other institutions were available online or in digital archives.

In fiscal 2003, seven new multimedia historical collections were added to the American Memory Web site, bringing the total to 123. Seven existing collections were expanded with approximately 344,000 digital items.

New American Memory collections from the Library of Congress added in fiscal 2003 (October 2002-September 2003) were:

Chinese in California, 1850-1920 (added through the Ameritech Competition; see below)

The Pearl Harbor Attack Collections

American Notes: Travels in America: 1750-1920

Macao Maps

Wilbur and Orville Wright Papers

American Women's History

Stars and Stripes: The Official Newspaper of the American Expeditionary Forces Printed in France from February 8, 1919, to June 13, 1919

<http://memory.loc.gov/>

In addition, five new Library exhibitions were mounted on the Library's Web site in 2003 and three continuing exhibitions were updated.

The new exhibitions in 2003 were:

Earth as Art: A Landsat Perspective

Ancient Manuscripts: From the Desert Libraries of Timbuktu

Herblock's Gift: Selections from the Herb Block Foundation Collection

Rivers, Edens, Empires: Lewis & Clark and the Revealing of America

Creative Space: Fifty Years of Robert Blackburn's Printmaking Workshop

In early 2004, two more collections were added: The Dream of Flight, marking the centennial of the Wright brothers' epochal first flight; and Churchill and the Great Republic, which celebrates Winston Churchill's life and times and examines his links to the United States. More than 45 exhibitions are now available on the Library's Web site.

<http://www.loc.gov/exhibits/>

### **3. The Learning Page**

Designed for teachers, the Learning Page is a companion Web site to the American Memory collections. It provides search and curricular guidance to teachers and students who use American Memory primary source materials. Throughout 2002 and 2003, the Learning Page was redesigned to enhance usability. Nine new lesson plans, designed by educators who have participated in the Library's American Memory Fellows Program, were made available, bringing the total number of lesson plans on the page to 80.

The Learning Page's Immigration feature presentation continued to add major sections of content that demonstrate how various ethnic groups shaped America. A Community Center was implemented as an interactive section of the site. Eleven discussion themes were introduced. In addition, eight interactive activities were created for teachers' use with their students, and seven Collection Connections were added to provide activities for using the Library's digital materials in the classroom.

The Learning Page also supports the National Digital Library teacher-training initiative and other educational outreach efforts available to all educators with Internet access.

Nineteen workshops from the National Digital Library teacher-training initiatives are now available on the Learning Page, as well as a conference page of national and local presentations by Library staff and teachers working with the Library of Congress initiative and descriptive indexes and reservation forms for workshops given in the National Digital Library Learning Center at the Library of Congress. A page of downloadable handouts for teachers was added. A quarterly electronic newsletter, The Source, written by teachers and school media specialists, is published on the Learning Page through collaboration with the Social Science Education Consortium.

<http://memory.loc.gov/ammem/ndlpedu/index.html>

## **September 11 Digital Archive**

The Library of Congress on Sept. 10, 2002, marked its first major digital acquisition of September 11, 2001, materials with the addition to its collections of the September 11 Digital Archive. The September 11 Digital Archive is a joint project of the City University of New York Graduate Center's American Social History Project and George Mason University's Center for History and New Media, institutions that have explored digital history for more than a decade.  
<http://911digitalarchive.org>

## **Veterans History Project**

With Public Law 106-380, signed by President Bill Clinton in October 2000, Congress directed the American Folklife Center at the Library of Congress to develop and coordinate a program to collect and preserve at the Library the audio and video recorded oral histories and documentary materials such as diaries and letters of America's 19 million war veterans. The Library works closely with interested groups, such as veterans organizations, to coordinate the project and make the collections available to the public, including online presentations. AARP (formerly American Association for Retired Persons) has provided generous partial funding for the project. Audio and visual (moving and still images and text) files of veterans' oral histories are available online, alphabetically under each veteran or grouped under the themes of:

- Courage
- Patriotism
- Community
- Life-Altering Moments
- Hurry Up and Wait
- Buddies
- Sweethearts
- Family Ties
- On a Mission

<http://www.loc.gov/folklife/vets/stories/> (Veterans' stories)

<http://www.loc.gov/folklife/vets/> (Veterans History Project home page)

## **B. SERVICES**

Laura Campbell, director of the National Digital Library Program (NDLP), was appointed as the Library's first Associate Librarian for Strategic Initiatives (ALSI) in October 2000 (start of fiscal 2001) to develop a full range of digital policies and operations for acquiring, describing and preserving content created and distributed in electronic form and for life-cycle management of digital materials as part of the Library's universal collection for the nation. Ms. Campbell retains her title as NDLP director and is also Chief Information Officer for the Library of Congress, with responsibility for oversight of the office of Information Technology Services and leadership of a new internal information technology planning group that works with external partners.

Deanna Marcum was appointed Associate Librarian of Congress for Library Services effective August 11, 2003. The Office of Strategic Initiatives, Law Library of Congress, U.S. Copyright Office, and Library Services all have roles in the delivery of digital products and services to the Nation. In addition, the Congressional Research Service delivers digital content to the Congress.

New developments in the past two years are presented below.

### **1. Digital Reference Team/QuestionPoint**

The Digital Reference Team handles the reference support for the digital collections and the Library's digital reference initiative. After the inauguration of the Ask A Librarian service in June 2002, the team provides both text-based and chat service via direct access from the Library's Web site. The team continues to test and build the knowledge base that is available to QuestionPoint member libraries and a global network. To this end the team has answered over 14,000 inquires and conducted 1600 live chat sessions.

<http://www.loc.gov/rr/askalib/>

Additionally the Digital Reference Team designs and presents demonstrations, on-site workshops, and video conferences for a wide audience. Opportunities for video conferencing and webcasting are continually expanding with 2,800 college and K/12 educators and students participating in 90 video conferences in the past year. Two new video conferences are "The Spy Map and George Washington" and "Gathering Your Community's Stories."

In November 2003 the Library's Public Service Collections Directorate received the 2003 Virtual Reference Desk Director's Award for its role in founding QuestionPoint, which evolved from the Collaborative Digital Reference Service, a collaboration to provide reference service anywhere, anytime using the Internet. QuestionPoint is now a subscription service hosted by OCLC Online Computer Library Center, Inc., with more than 800 member libraries.

<http://www.loc.gov/rr/digiref>

## **2. Global Legal Information Network**

The Law Library of Congress's primary digital initiative, the Global Legal Information Network (GLIN) is an online parliament-to-parliament cooperative exchange that has made the laws and legal materials (often in the vernacular) of forty-eight nations available via the Internet. In fiscal 2003, there were about 1.7 million visits to the GLIN online database.

The Law Library of Congress contracted with the General Services Administration to perform a major technical upgrade of the GLIN system in 2004.

An ongoing project adds retrospective laws to GLIN for the countries of Latin America, covering the period from 1975 through 1995, by scanning the microfilm of official gazettes and producing PDF files. About 28,000 full texts of laws from nineteen Latin American countries have thus been linked in the GLIN database.

The Law Library continued to review foreign legal gazettes, select all appropriate legal instruments, summarize and assign legal subject terms and convert to PDF files. During 2003, 7,200 summaries of laws from 23 jurisdictions were added to the GLIN database by the Law Library. These jurisdictions were primarily in the Americas but also included other Spanish, French and Portuguese speaking countries around the world.

Access to summaries is free to anyone with Internet access, but access to full text of laws and legal materials has been restricted to GLIN members. The GLIN Executive Council adopted in 2003 a new policy of encouraging GLIN members to allow free public access to their legal information in GLIN.

<http://lcweb2.loc.gov/law/GLINv1/>

## **3. Interlibrary Loan**

This year the Library of Congress will process more than 50,000 requests for interlibrary loans. In March 2001 all interlibrary loan requests to the Library began to be funneled through a single computer gateway, ILL Manager, purchased from RLG. The software channels requests from libraries using OCLC, RLIN, and the Library of Congress Web into a single workstream. By October 2003, the Collections Access, Loan and Management Division had digitized and delivered over the Web more than 150 small, fragile pamphlets requested by ILL patrons. The Library has also committed to capturing

the content of scanned items through OCR of the PDF files. Access to the OCR files is currently being explored.

<http://lcweb.loc.gov/rr/loan/>

#### **4. Library of Congress Online Catalog**

The Library of Congress Online Catalog is the public catalog module of the Library of Congress Integrated Library System (LC ILS). In 2003 links were added from the main Online Catalog home page to the Prints and Photographs Online Catalog (PPOC) and the Sound Online Inventory and Catalog (SONIC), which provides catalog access for many of the Library's 2.5 million sound recordings.

<http://catalog.loc.gov>

### **C. SYSTEMS**

#### **1. Digital Audio-Visual Conservation**

In August 1998, the Congress authorized the Architect of the Capitol (AOC) to acquire on behalf of the Library real property and improvements in Culpeper, Virginia for use as the National Audio-Visual Conservation Center. The Center, scheduled to open in phases, June 2005 through March 2006, will implement digital archiving of the Library's collections of film, television, radio and recorded sound and will house the Library's film and audio and video preservation laboratories.

With new digital audio workstations, the Motion Picture, Broadcasting and Recorded Sound Division has begun creating digital preservation files, rather than relying exclusively on analog tape as the reformatting target media. The Recording Lab continued in 2003 to set up a high speed SAN (storage area network) for centralized recording, editing and batch processing of digital audio files. A digital object metadata system for sound recordings was also refined; the system captures METS (Metadata Encoding and Transmission Standard) digital-object metadata in Oracle and Java application hosted in ITS servers. Java tools produce the XML output from the database. A digital workstation for recorded sound researchers was established in the Performing Arts Reading Room of the Library's James Madison Memorial Building on Capitol Hill, including interim access management protection to protect copyrighted content.

<http://www.loc.gov/rr/mopic/avprot/avprhome.html>

#### **2. Integrated Library System (LC ILS)**

The LC ILS includes modules for cataloging, circulation, acquisitions and serials check-in as well as the Library of Congress Online Catalog. The software, procured from Endeavor Information Systems, Inc., is Voyager release 2000.1.3 as of March 31, 2004. The Library made three major improvements to ILS performance and stability in 2003. The first major hardware upgrades to the ILS were implemented, including a migration to new servers; a read-only copy of the production database was implemented, permitting twice the number of simultaneous sessions; and system slowdowns caused by heavy use of keyword indexes and the Z39.50 server were eased.

The Congressional Research Service Voyager application ramped up to full production in fiscal 2003, now utilizing the acquisitions module and the "New Books" add on feature. Also in 2003, a new Voyager database was set up for the National Library Service for the Blind and Physically Handicapped, a division of the Library of Congress. The inventory management system for the offsite collections storage system at Fort Meade, Maryland, was put into operation.

#### **3. Portal Application and Open URL Software**

The Library of Congress Portals Applications Issues Group (LCPAIG) completed and issued two key documents in 2003:

- *List of Portal Application Functionalities for the Library of Congress, First Draft for Public Comment*, issued on July 15, 2003
- *Functional Requirements for an OpenURL Resolver for the Library of Congress*, issued November 4, 2003

Both of these documents are available on the documents page within the LCPAIG Web site, which also links to additional resources, including “Starting Out with Portals and OpenURL: An Introduction”; pages with numerous links to federated search portal and OpenURL resolver products and vendors; portal products used by academic and educational institutions; and a selection of Federal government agency and academic institution Web portals and subject “gateways.” The Web site also includes sections devoted to portal and OpenURL standards, reports and information about portals available online from several national and international library organizations, and links with abstracts of dozens of Web articles on portals and OpenURL resolvers.

Functional requirements for the acquisition of OpenURL software were specified in November 2003. An award was announced in 2004.

<http://www.loc.gov/catdir/lcpaig/documents.html>

## II. Projects and Programs

### A. PROJECTS

#### 1. Bibliographic Enrichment Advisory Team (BEAT)

The Bibliographic Enrichment Advisory Team (BEAT) develops tools to aid catalogers, reference specialists, and searchers in creating and locating information. Major components of the team's work are enriching the content of Library of Congress bibliographic records, improving access to the data the records contain, and conducting research and development in areas that can contribute to furthering these efforts.

<http://lcweb.loc.gov/catdir/beat>

Updates on several of the 18 projects now underway are given below.

#### **Pre-1970 Congressional Hearings**

This recently approved project will improve access to approximately 2,000 pre-1970 Congressional hearings, improving service to the Congress. Other benefits to be derived include central, as opposed to dispersed availability of the hearings, better access to the content of the items, improved cataloging describing these hearings, and the addition of other information concerning alternate data source availability. The methodology to be employed consists of identifying and retrieving approximately 2,000 items, reclassifying them into class KF in the Law schedule and upgrading the cataloging for the items. In addition, the project will arrange to make digital copies available for hearings in poor physical condition.

<http://www.loc.gov/catdir/beat/pre1970.hearings.html>

#### **ONIX TOC/Publishers Descriptions**

ONIX (ONline Information eXchange) is a means of representing book industry product information, used by some publishers today to communicate that data electronically. The Library receives these data directly, and with programming developed by BEAT, the project creates Table of Contents (TOC) and Publisher Description records that the Library makes available on the Web. Hyper-links are made from this TOC data to the catalog record, and the reverse, thus allowing researchers to move from or to the Library's Online Catalog where they can make

additional searches for related or other material. To date the project has created about 43,000 ONIX TOC records and 64,000 ONIX Descriptions records.

<http://www.loc.gov/catdir/beat/onix.toc.html>

<http://www.loc.gov/catdir/beat/onix.descriptions.html>

### **Links to Book Jackets**

BEAT staff plan to link images of dust jackets to titles in the ONIX Descriptions and ONIX TOC initiatives. Initially, there will be links to approximately 2,300 dust jacket images, and it is anticipated that the number will grow as publishers can provide that data through existing BEAT channels. This further enriches the data that LC provides to searchers about cataloged materials. This project was to be implemented in 2004.

### **ECIP Tables of Contents (TOC)**

In a new aspect of this project, using programming by three BEAT Team members, a Web-based TOC is created for virtually all ECIP records that contain TOC data. A hot-link in the TOC file is made to and then from the underlying record in the LC Online Catalog. The programs handle most diacritical marks, and also enrich the TOC Web display by adding such LC subject headings as were applied by cataloging staff. In addition to this new development, table of contents data are added in about 30 percent of the bibliographic records by staff in the course of the regular ECIP cataloging process. As of December 2003, approximately 12,000 ECIP TOC records had been added to the Web server.

<http://www.loc.gov/catdir/beat/ecip.html>

### **Web Access to Works in the Public Domain**

This project makes links from the LC Online Catalog to full electronic texts of items represented in the LC collections. Two recent additions are noted. In the most recent development, 379 titles represented in the Library's collections were linked to The California Digital Library ( <http://www.cdl.org> ) to *eScholarship* texts that have been made publicly available through the Library. The texts to which LC was granted permission to link are mostly recent imprints and current scholarship, whereas prior Public Domain projects were for retrospective material primarily of historical value. In another addition to this project, BEAT has added links to bibliographic records for more than 35 titles from Indiana University's *Victorian Women Writers* Project, which includes anthologies, novels, political pamphlets, religious tracts, children's books and volumes of poetry and verse drama.

<http://www.loc.gov/catdir/beat/pub.domain.html>

### **H-NET Reviews**

To enrich access to scholarly materials, BEAT has begun to link catalog records for selected materials in the Library's collections to reviews for them in *H-Net Reviews in the Humanities and Social Sciences*, the online journal of H-NET: Humanities and Social Sciences Online, a resource that contains many scholarly and academic reviews. The reviews cover contemporary publications as well as classic publications in a variety of academic fields. Many of the reviews, in addition to containing relevant Library of Congress subject headings, include footnotes and bibliographies. The initial phase of the cooperative project resulted in 6,173 links from Library of Congress catalog records to corresponding H-Net reviews. An estimated 1,000 to 2,000 additional reviews will be added each year.

<http://www.loc.gov/catdir/beat/hnet.html>

### **BECITES+**

BECITES+ (Bibliographies plus: Enhanced Citations with Indexes, Tables of contents, Electronic resources and Sources cited) enhances staff- produced bibliographies, and the catalog records for the titles included in such

bibliographies, by adding links to their tables of contents, indexes, and sources cited. Another recent initiative has been the scanning and conversion to text of heavily used, but out-of-print guides to Library of Congress collections whose individual items are not easily identified in the Library's Online Catalog. The project uses scanning and OCR to substantially enrich these traditional printed bibliographies. Links in the catalog records are made for each type of data file created for the work in question as well as between all the related files for any work for which a Web file is included. Completed works within this project include guides on business history, African American business, Thomas Jefferson, and materials on Immigrant Arrivals to the United States. A number of additional works are in progress, covering additional business resources, guides to microfilm collections, three guides to Prints and Photographs Division collections, manuscript collections from several Middle Eastern monasteries, and a guide to Ladino publications in the Library of Congress.

<http://www.loc.gov/rr/business/guide/>

## **2. Global Gateway (formerly International Horizons)**

Global Gateway is a major portal that brings together bilingual collections of digitized artifacts, research guides and databases, cybercasts, and information on research opportunities using the Library of Congress collections. The digitized primary research materials are found in the collaborative digitization efforts *Meeting of Frontiers* and *Spain, the United States, and the American Frontier: Historias Paralelas*.

<http://international.loc.gov/intldl/intldlhome.html>

### **Meeting of Frontiers**

The National Library of Russia and the Russian State Library continued to contribute digitized historical materials to Meeting of Frontiers, a Russian-American bilingual Web site.

Recent additions to the Meeting of Frontiers collections are from the Library of Congress; the State and University Library of Lower Saxony of Göttingen, Germany; the National Library of Russia; and the Russian State Library. Project partners since 1999, the Russian State Library and the National Library of Russia contributed rare books, maps and manuscripts to their already extensive collections of digitized materials on the Meeting of Frontiers site. The additions include unpublished memoirs of Russian exiles in Siberia, Russian documentation about the Russo-Japanese War of 1904-1905 and the 1837 translation into Aleut of the Russian catechism by Ioann Veniaminov (St. Innocent). With more than 330,000 digital images, Meeting of Frontiers is the world's largest bilingual collaborative library site.

<http://frontiers.loc.gov/>

### **Kiowa Stories**

Library of Congress collections added to Meeting of Frontiers in fiscal 2003 include the Kiowa Stories from the papers of Hugh Lenox Scott and the Eleanor L. Pray Album. Scott was a West Point graduate and career military officer who served at various Western posts between 1876 and 1897. In 1892, he was assigned to Fort Sill, Okla., and given command of Troop L of the 7th Cavalry, an all-Indian unit comprising Kiowa, Comanche and Apache.

### **Eleanor L. Pray Album**

Images from Vladivostok in 1899-1901 and the life of an American merchant family living in the city at that time.

### **Asch Collection**



The latest Meeting of Frontiers update also includes, from the State and University Library of Göttingen, a large portion of this institution's extraordinary Asch Collection. The collection, which is a creation of Georg Thomas von Asch (1729-1807), a German who studied medicine at Göttingen and then entered the Russian National Service, is a comprehensive record of Russian expeditions to Siberia in the second half of the 18th century.

<http://frontiers.loc.gov/intldl/mtfhtml/mfdigcol/subcoll.html>

#### **Spain, the United States, and the American Frontier: Historias Paralelas**

This site, created in 2001, brings together unique collections of the Library of Congress with those of partner institutions in Spain, most notably the National Library of Spain and the Biblioteca Colombina y Capitular of Seville. The collection draws upon rich Library of Congress collections relating to Spain and Spanish America digitized from the Law Library, the Library's general collections and collections of the Geography and Map, Manuscript and Rare Book and Special Collections divisions.

<http://international.loc.gov/intldl/eshtml/>

#### **Portals to the World**

This Global Gateway project is designed to provide access to online resources for all international regions of the world. Portals are created using OCLC Connexion Pathfinders. To date, more than 700 Pathfinders have been created.

<http://www.loc.gov/rr/international/portals.html>

### **3. Elektronische Zeitschriftenbibliothek (EZB)**

In 2003 the Library of Congress joined the Elektronische Zeitschriftenbibliothek (EZB), a consortium of more than 200 libraries in Europe and the United States. It offers a database of member holdings for more than 15,000 scientific and academic journals available in full text on the Internet. More than 5,000 of these e-serials are available free of charge, and they are currently being evaluated individually as possible additions to the Library's digital collections.

### **4. International Children's Digital Library**

The Library participated in the development of the International Children's Digital Library (ICDL), a cooperative project with the Internet Archive and the University of Maryland to build a Web site containing international literature for children. The Library selected and digitized approximately 60 children's books from its general, rare, and special collections as its initial contribution to the site. The ICDL Web site was launched Nov. 20, 2002, at the Library of Congress in conjunction with the celebration of National Children's Book Week.

[www.icdlbooks.org](http://www.icdlbooks.org)

### **5. MINERVA**

The Library of Congress continues to develop a Web preservation project, Mapping the Internet Electronic Resources Virtual Archive (MINERVA), initiated in fiscal year 2000 to capture, save, and preserve collections from select Web sites for use by future generations of researchers. During phase I of the project, the Library designed a prototype Web site, selected pilot sites to archive and developed general guidelines. Bibliographic records were created and open access was provided for the captured Web sites. In the second phase, the LC MINERVA Team created the Election 2002 Web Archive, in collaboration with WebArchivist.org of the State University of New York Institute of Technology and the Internet Archive, with additional funding from The Pew Charitable Trusts through the University of Washington Center for Communication and Civic Engagement. The Election 2002 Web Archive is a selective collection of nearly 3,000 sites archived between July 1, 2002 and November 30, 2002.

The initial March 4th, 2003 release of the Election 2002 Web Archive included Web sites produced by congressional and gubernatorial candidates, and featured a unique “Drill search” interface developed with WebArchivist.org, allowing access to Web Archive Records and archived sites. In July 2003, the Library of Congress began hosting the archive directly from LC servers (previous collections had been temporarily hosted by the Internet Archive). Additionally, the Library of Congress, in collaboration with the Internet Archive, webArchivists.org and the Pew Internet & American Life Project, launched a September 11 Web Archive. This new archive preserves the Web expressions of individuals, groups, the press and institutions in the United States and from around the world in the aftermath of the terrorist attacks on the United States.

<http://september11.archive.org/>

In fiscal 2003 the MINERVA project also collected Web sites relating to the 107<sup>th</sup> Congress and the War in Iraq. Collection-level AACR2/MARC catalog records were created for each thematic Web site collected in order to represent these items in the LC ILS. Building upon traditional methods, we are in the process of supplementing the collection level metadata by experimenting with the creation of title-level descriptive metadata for each Web site within the collection using the Metadata Object Description Schema (MODS).

## **B. PROGRAMS**

### **1.National Digital Information and Infrastructure Preservation Program**

The National Digital Information Infrastructure and Preservation Program (NDIIPP) is funded by a fiscal 2001 appropriation of \$99.8 million (originally \$100 million, reduced by the government-wide rescission of 0.22 percent in December 2002) from the U.S. Congress. Congress has asked the Library to lead this nationwide effort “in collaboration with other federal and nonfederal entities.” The goal of NDIIPP is to:

- encourage shared responsibility among many institutions and organizations for the collection, storage and preservation of digital content,
- seek national solutions for the continuing collection, selection and organization of historically significant cultural materials regardless of evolving formats,
- ensure the long-term storage, preservation and authenticity of those collections and
- work toward persistent, rights-protected access for the public to the digital heritage of the American people.

In December 2002, Congress accepted “Preserving Our Digital Heritage: Plan for the National Digital Information Infrastructure and Preservation Program.” The plan was released to the public on February 14, 2003. The result of extensive fact-finding, planning and consultation with a broad range of stakeholders, the plan outlines the steps the Library will take as it leads the national effort to develop a digital preservation infrastructure consisting of:

- a network of committed partners and
- the technical architecture to support long-term digital collection, storage and preservation.

Concurrent with the plan’s approval was the release of \$35 million to begin the program’s next phase.

Oversight of the plan rests with the 26-member NDIIPP Advisory Council, which includes the Librarian of Congress, the Chief Executive of the British Library, the Archivist of the United States, the Secretary of Commerce, the director of the National Library of Medicine, and representatives from the White Office of Science and Technology Policy, private industry, and academia.

In August 2003, the National Digital Information Infrastructure and Preservation Program issued an announcement seeking applications for projects that will advance this nationwide program.

This first set of projects will focus on two major NDIIPP goals:

- The selection and collection of at-risk and historically significant digital materials for which no analog equivalent exists and
- the development of a network of committed NDIIPP partners with defined roles and responsibilities to support the long-term collection and preservation of digital content.

These project applications will develop and test models for the collection of digital materials, specifically those that are historically significant and at risk of disappearing if they are not captured. Awards are to be announced in late spring 2004.

[www.digitalpreservation.gov](http://www.digitalpreservation.gov)

### **Ameritech Competition Concludes**

One of the seven new American Memory collections this year was the result of a Library of Congress-Ameritech award. The new presentation brings the total number of collections made available through this program to 23 and concludes the project. Beginning in 1997, the Library of Congress sponsored this three-year competition with a gift from the Ameritech Corporation to enable public, research and academic libraries, museums, historical societies and archival institutions (with the exception of federal institutions) to digitize American history collections and to make them available on the Library's American Memory Web site. Those digital collections complement and enhance the Library's online resources. Thirty-three institutions received \$1.75 million of support to digitize 23 projects.

<http://memory.loc.gov/ammem/award/index.html>

## **2. Adventure of the American Mind**

In fiscal 2003, the Adventure of the American Mind (AAM) program grew to include a total of 17 partners in five states. This program was created by Congress, and implemented by the Library of Congress with the Educational and Research Consortium of the Western Carolinas. Its purpose is to bring the riches of the Library's online collections to students in the classroom by providing educators with the tools and training to integrate electronic primary sources into their teachings.

Several partners joined the AAM program in fiscal 2003, including Loyola University, DePaul University, Governor's State University, the Federation of Independent Illinois Colleges and Universities and Southern Illinois University, Carbondale (all in Illinois); the University of South Carolina in Spartanburg; Brevard College in North Carolina; and the Arlington, Alexandria and Falls Church school system in Virginia.

## **3. National Library Service for the Blind & Physically Handicapped (NLS/BPH)**

The National Library Service for the Blind and Physically Handicapped (NLS), Library of Congress, completed the five-year development phase of digital talking book technology during fiscal year 2003 and began the conversion phase to implement the new digital talking books and playback machines by 2008. To ensure that NLS/BPH has an adequate store of digital materials available when it begins its transition from analog to digital distribution, all contractors producing audio books for NLS/BPH have been given a schedule for converting their studios, culminating in 100 percent digital mastering in fiscal 2004.

NLS has set a target date of 2008 to have 20,000 DTB titles and 50,000 players ready for initial distribution. The first request for proposals (RFP), including detailed specifications on

procurement and quality assurance, has been issued. Contractors will begin producing DTBs according to those specifications in 2004. Conversion of 10,000 of the 40,000 analog recorded catalog titles to DTB format is planned, and an RFP for the pilot conversion of 200 titles has been issued. NLS also sponsored the development of a system whereby network libraries can duplicate analog cassettes from digital masters and convert analog recordings to digital format. It is currently defining the hardware and software needed to manage book files during recording and quality assurance.

In the early 1990s, NLS instructed braille contractors to provide electronic copies of the books they produced. Those books and all new braille books and magazines are now available to users online through the Web-Braille system. Web-Braille allows access to more than 5,900 digital braille book files, twenty-six national magazines, and six national sports schedules. At the close of fiscal 2003, more than 3,000 users were registered for this Internet service. The NLS *International Union Catalog* for braille and audio materials also linked to Web-Braille. As a result, Web-Braille books could be accessed directly from the catalog by using author, title, subject, language, keyword, and other search parameters.

[www.loc.gov/nls/braille](http://www.loc.gov/nls/braille)

#### **4. Digital Standards**

The Office of Strategic Initiatives and the Network Development & MARC Standards Office (NDMSO), Library Services, are the focal point for technical library, network and digital standards and related planning at the Library of Congress.

##### **METS**

NDMSO is the development and maintenance agency for the Metadata Encoding and Transmission Standard (METS), an important standard for repository management of digital objects. METS is a schema for creating XML documents that express the hierarchical structure of digital library objects, the names and locations of the digital files that comprise those objects, and the associated metadata. In 2003 NDMSO worked with the international METS Editorial Board in the release of Version 1.3.

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

##### **MODS**

The METS Editorial Board endorsed MODS (Metadata Object Description Schema), which is a simpler XML companion to MARCXML, and the Metadata for Images in XML schema (MIX) as extension schema. The MODS schema is intended to carry selected data from existing MARC 21 records as well as to enable creation of original resource description records. Version 3.0 was made available in December 2003.

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

##### **QATP (Question/Answer Transfer Protocol)**

The chief of NDMSO chairs the NISO (National Information Standards Organization) Standards Committee AZ (Networked Reference Services). In April 2004, the Committee announced that a draft Question/Answer Transaction Protocol that supports Q&A between library patrons and reference sources has been released for a one-year trial use period April 5, 2004-April 5, 2005. The draft standard and guidelines for registration and use are available on the NetRef Committee Web page.

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

##### **ZING/SRW**

In 2003 NDMSO finalized version 1.1 of SRW (Search/Retrieve Web Service), a protocol of the Z39.50-International Next Generation (ZING) initiative that evolves the Z39.50 information retrieval protocol to a new Web service. With the release of SRW Version 1.1 on February 13, 2004, Version 1.0 was officially deprecated. The companion service SRU (Search and Retrieve URL Service), also released in Version 1.1 on February 13, encodes the request in the URL.

### **5. New Books Program and Electronic Cataloging in Publication**

The ECIP program now includes more than 2,800 publishers, and nearly half of all Cataloging in Publication requests are received electronically.

New Books, a Library of Congress Cataloging Directorate initiative, is designed to link highly desirable auxiliary information such as author biographies, images of book jackets, and summaries or additional subject terms to the catalog records for forthcoming books. In 2003, close consultation with the library and publishing communities enabled the project developers to decide that the New Books system would use the unabridged Book Industry Standards and Communications subject terms for subject access to New Books records. In addition, New Books would be integrated into the ECIP home page, effectively converting the ECIP home page to a New Books/ECIP home page.

<http://lcweb2.loc.gov:8081/ecip/celli/welc01.html>

### **6. Open Archives and Cultural Materials Initiatives**

The National Digital Library Program engages in efforts to broaden access to the unique materials digitized for American Memory and to encourage interoperability among cultural repositories. As a beginning, the descriptive records for some American Memory collections have been made available for sharing with other institutions using the Open Archives Initiative standard. The records now available are either in MARC format or presented as Dublin core records. American Memory collections are contributed to the Cultural Materials Initiative, a pooled resource being assembled by RLG, from digital resources of participating members.

### **7. Preservation Digital Reformatting Program**

The Preservation Digital Reformatting Program has put its Interlibrary Loan Brittle Book digitization project into production. More than 100 items in the public domain that were too brittle for traditional interlibrary loan delivery have been digitally reformatted into approximately 9,000 page images. Selection criteria and principles of digital reformatting have been issued on the Library of Congress Preservation Directorate Web site.

<http://lcweb.loc.gov/preserv/prd/presdig/presintro.html>

## **III. Specific Digital Library Challenges**

### **Collections Access, Preservation, and Security**

The digital explosion has imposed on the Library of Congress a new mission-critical workload and the need to expand high-quality, free online services to the Congress, K-12 educators, and the American public. This work must be carried out in conjunction with the Library's equally critical traditional services of acquiring, cataloging, preserving, serving and storing artifactual materials. The Library of Congress must continue to construct digital-repository architecture and basic technology infrastructure to preserve current and future digital assets. As the Library's vast multiformat collections grow, it must continue to invest in securing and preserving these cultural records. The Library's ability to launch, deliver, preserve and secure digital content and services for the Congress and public depends on congressional support.

The internationally recognized need for a distributed, coordinated approach to digital preservation is one of the Library of Congress's primary challenges, shared with many other institutions. The Plan for the National Digital Information Infrastructure and Preservation Program addresses this challenge. In the course of carrying out the plan, the following challenges will have to be addressed:

There is no ready commercial answer for migrating content from the World Wide Web to accessible archives.

Technology may be less important than organizational and cultural factors and issues of responsibility.

Intellectual property rights are critical. There is confusion about copyright and what partner institutions can and cannot do under the law.

There is a need to define what digital material will be collected, who will collect it, who will store and maintain it, and how access to digital material will be provided.

## **Digital Assets Security**

The Library continued to enhance the security of its digital collections and other digital assets in 2003. A new information technology security policy was drafted, and the LC Computer Security Coordination Group involved all service and support units. Computer security awareness training was made part of mandatory staff training. The National Security Agency performed a risk assessment of the Library of Congress network and digital assets held by Information Technology Services, and audits of Library of Congress servers were initiated.

Threat surveillance and selective dissemination of alerts and advisories have been extended to 24 hours a day, seven days a week. The enterprise firewall has been upgraded, including stronger encryption standards. Library of Congress systems engineering staff designed and constructed an Alternate Computing Facility that mirrors the Library's high priority applications. All enterprise-level digital data are remotely stored in backup tapes, and all primary servers, storage, and backup systems are replicated at the Alternate Computing Facility to ensure continuity of service and data security in the event of a disaster in the Library's main computer facility in the James Madison Memorial Building on Capitol Hill.

## **Digital Life Cycle Planning**

The Office of Strategic Initiatives and the Library Services service unit worked together to develop a Digital Life Cycle Framework, which collects information on plans and costs of each digital initiative proposed within the Library, to enable the Library to project costs and manage its staff and other resources.

A significant cost of the digital life cycle is associated with the provision of bibliographic access, whether through the library catalog or via a Web portal or general search engine. The Library's cataloging divisions have made a commitment to increase the degree of bibliographic access to digital content dramatically in 2004 and 2005. New modes of cataloging and new workflows have been proposed.

<http://lcweb.loc.gov/catdir/stratplan/goal4wg2report.pdf> (Modes of Cataloging)

<http://lcweb.loc.gov/catdir/stratplan/goal4wg4report.pdf> (Workflow)

## **IV. Digital Library publications, policies, working papers, and other documents**

Bibliographic Control of Web Resources: A Library of Congress Action Plan. Rev. Feb. 5, 2004.  
<http://lcweb.loc.gov/catdir/bibcontrol/actionplan.pdf>

It's About Time: Research Challenges in Digital Archiving and Long-term Preservation. Proceedings of a workshop sponsored by the National Science Foundation and the Library of Congress, April 12-13, 2002. 15.9 MB.  
[http://www.digitalpreservation.gov/repor/NSF\\_LC\\_Final\\_Report.pdf](http://www.digitalpreservation.gov/repor/NSF_LC_Final_Report.pdf)  
(Available from NDIPP Web site.)

Preserving Our Digital Heritage: Plan for the National Digital Information Infrastructure and Preservation Program.

Part I: Executive Summary (74 pages, 3.1 MB)

[http://www.digitalpreservation.gov/repor/ndiipp\\_plan.pdf](http://www.digitalpreservation.gov/repor/ndiipp_plan.pdf)

Part II: Appendices (266 pages, 16.9 MB)

[http://www.digitalpreservation.gov/repor/ndiipp\\_appendix.pdf](http://www.digitalpreservation.gov/repor/ndiipp_appendix.pdf)

Brylawski, Samuel. Preservation of Digitally Recorded Sound. Environmental scan commissioned by the Council on Library and Information Resources for the Library of Congress, 2001.  
<http://www.digitalpreservation.gov/index.php?nav=3&subnav=7>

Flecker, Dale. Preserving Digital Periodicals. Environmental scan commissioned by the Council on Library and Information Resources for the Library of Congress, 2001.  
<http://www.digitalpreservation.gov/index.php?nav=3&subnav=6>

Friedlander, Amy. Background Summary of Results from Interviews and Essays. Summary of interviews with leaders in digital librarianship and information and six environmental scans, commissioned by the Council on Library and Information Resources for the Library of Congress in 2001.  
[http://www.digitalpreservation.gov/repor/interviews\\_summary.pdf](http://www.digitalpreservation.gov/repor/interviews_summary.pdf)

Ide, Mary et al. Understanding the Preservation Challenge of Digital Television. Environmental scan commissioned by the Council on Library and Information Resources for the Library of Congress, 2001.  
<http://www.digitalpreservation.gov/index.php?nav=3&subnav=9>

Lyman, Peter. Archiving the World Wide Web. Environmental scan commissioned by the Council on Library and Information Resources for the Library of Congress, 2001.  
<http://www.digitalpreservation.gov/index.php?nav=3&subnav=5>

Romano, Frank. E-Books and the Challenge of Preservation. Environmental scan commissioned by the Council on Library and Information Resources for the Library of Congress, 2001.  
<http://www.digitalpreservation.gov/index.php?nav=3&subnav=8>

Wachtlar, Howard D., and Michael G. Christel. Digital Video Archives: Managing Through Metadata. Environmental scan commissioned by the Council on Library and Information Resources for the Library of Congress, 2001.  
<http://www.digitalpreservation.gov/index.php?nav=3&subnav=10>

National Digital Information Infrastructure and Preservation Program Convening Sessions, November 5-6, 7-8, 15-16, 2001: Summary Report.  
[http://www.digitalpreservation.gov/repor/cs\\_summary\\_012802.pdf](http://www.digitalpreservation.gov/repor/cs_summary_012802.pdf)