



## THE LIBRARY OF CONGRESS

101 INDEPENDENCE AVENUE, S.E.  
WASHINGTON, D.C. 20540-1610  
EMAIL PAO@LOC.GOV

PUBLIC AFFAIRS OFFICE

**Statement of Dr. James H. Billington  
The Librarian of Congress  
before the  
House Subcommittee on Legislative Branch  
U.S. House of Representatives  
March 20, 2007**

Madam Chair Wasserman Schultz, Mr. Wamp and members of the Subcommittee:

It is a pleasure and honor to appear before you today. We are pleased that this important subcommittee has been restored and has given us separate opportunities both to examine the Library's FY2008 budget and to review the Library's broad transformation into the digital age.

I will outline the overall vision and strategy with which the Library has been addressing the challenge of digital technology, and I will then ask my colleagues in the leadership of the Library to describe briefly the goals and challenges ahead in their respective areas.

### **The Library of the 21st Century**

The Congress of the United States has been the greatest patron of a library in the history of the world. Building on its purchase of Thomas Jefferson's large and wide-ranging personal library in 1815, the Congress has created and sustained the largest single repository of recorded knowledge in the widest variety of languages and formats in human history as well as the most exhaustive record anywhere of the private sector creativity of the American people. This body of knowledge helps inspire Americans to create and achieve, providing our country with the intellectual currency that is necessary to maintain our unrivaled position of global competitiveness.

It took two centuries for the Library of Congress to acquire today's analog collection—32 million printed volumes, 12.5 million photographs, 59.5 million manuscripts and other materials – a total of more than 134 million physical items. By contrast, with the explosion of digital information, it now takes only about 15 minutes for the world to produce an *equivalent* amount of information. Researchers at Cal-Berkeley produced estimates of the amount of information produced and circulated on the Internet in 2003 – it was equivalent to 37,000 times the content of one Library of Congress. Most of this information exists *only* in digital form: so-called born-digital items, many of which are already irretrievably lost.

There is a widely-held but false assumption that digital materials accessible today on one's PC or Blackberry will necessarily be available in the future. That is not the case. The average life of a Web site has been estimated to be 44 to 75 days, and information not actively preserved today could literally be gone tomorrow. Other essential digital information—most notably e-journals and data bases—are merely licensed for use in the short term—the information does not belong to the licensee. By contrast, traditional print books and journals collected by the Library for more than two centuries are, and will remain, in the possession of the Library and accessible to researchers. But it is current information that is often most needed by Congress, and current, up-to-date information is increasingly available only in digital form.

To cite some examples: of a sample set of 56 primary sources identified by the Congressional Research Service to support research on Hurricane Katrina in 2005 and 2006, 21% were no longer avail-

able on the Web in 2007. Web sites relating to the national elections of 1994—the first time the Web played a role in such elections—have also vanished. It was not until 2000 that the Library began preserving election Web sites. Political scholars wishing to write the history of how the Web has influenced politics will have to do so without important pieces of the puzzle. Such information is and will remain of critical importance to Congress, researchers in the Congressional Research Service, other government agencies and the public.

How did the Library prepare itself for the explosion of digital information? Since the 1960s, the Library has been utilizing early electronic technology for creating and sharing bibliographic information from the Library's vast catalog. In 1975, CRS made its Issue Briefs available electronically to the Congress for the first time. By 1980, the Library began to track patron use of its electronic material and recorded over 22 million transactions. In 1993, the Library posted its first records and information on the Internet and recorded 7 million transactions; by the following year that number had more than quintupled to 38 million. An overview of the Library's digital milestones is included in the attached charts.

The Library needed to explore new media in order to “get the champagne out of the bottle.” As a historian, I was always especially moved and stimulated by contact with primary documents. As I became aware of the full range of primary materials in the Library's collections, I became convinced of the need to share these unique national and international treasures more broadly with younger people whose knowledge of American history was becoming more and more tenuous.

The requests for access to such material came through a series of 10 national forums with Library leaders in 1988. These meetings convinced me that access to the unique elements in the Library's collection was desirable for teachers and librarians who could not get themselves, let alone their students and readers, to Washington, D.C.

Accordingly, in 1990 the Library launched, with a mixture of congressional and private support, a test experiment of primary sources on CD-ROMs in 44 sites throughout the country. We found that the most avid interest and use was not at the higher levels of education but in K-12, with a surprising amount of interest in even the third and fourth grades, where the loss of intellectual curiosity often becomes set for life. In late 1994, we launched a program to digitize 5 million items of American history and culture for educational purposes—the National Digital Library. The budget was \$60 million with a 3-to-1 private match for every dollar of congressional support. By the end of the 1990s, the Library had well over 5 million items of American history on-line. We have continued this process and now have more than 11 million items on our American Memory Web site for educational use by teachers and librarians. The Library has benefited from the support of the Ad Council to promote the Library's educational and literacy programs. Our overall Web usage climbs continuously and now stands at more than 5 billion electronic hits each year.

Paralleling the overnight success and continuous growth of the National Digital Library was the creation and wide-spread use of THOMAS, our Web site launched in 1995 at the request of Congress to provide constantly updated information about the Congress of the United States for the general public. In 1999 we developed an international bilingual Web site with the National Library of Russia, “Meeting of Frontiers”, and more recently, with six other national libraries, the international equivalent of our National Digital Library. These efforts that began with countries that had an impact on or parallel developments with American history are leading now to our current efforts to build a World Digital Library that will include Arabic-Islamic memory and Brazilian memory, just as we did with American Memory.

As our digital presence grew and our adaptation of electronic technologies became prevalent, the Library needed external analysis of what problems might lie ahead, where the information revolution was heading, and whether our structure, staff, and systems would continue to flourish in this rapidly expand-

ing new world of digital information. Accordingly, I asked the National Academy of Sciences to conduct a strategic study, *LC 21: A Digital Strategy for the Library of Congress* (NAS: 2000).

The key challenge posed by the Academy's experts for the Library was to capture, collect, preserve, and provide access to important "born-digital" material and Web-based information. In 2000, the Library requested a special appropriation to meet the digital challenge – the National Digital Information Infrastructure and Preservation Program (NDIIPP) – authorized and funded by Congress to give the Library the mandate and resources to undertake the necessary R&D to capture, preserve, and provide access to essential born-digital information – information that is exploding annually at the rate of 5 exabytes.

The Library has a long tradition of partnerships and collaboration to build and organize its collections. In 2000 the Library convened a major library conference on the future of bibliographic control in the digital age. Similarly, the scope of NDIIPP's purpose was unprecedentedly broad, and the legislation prescribed specific collaboration with both Legislative and Executive Branch agencies as well as other public—and private-sector organizations and institutions. Through NDIIPP, the Library has built a national network of 67 partners to collect, save and provide access to a body of high-quality research and educational content in digital form. We have been working closely with content providers, technology innovators, libraries, archives, and end-users to advance the science and practice of preserving important at risk materials that are perishable and often exist only in digital form. (A full list of collaborating partners is included in the attachments.) The Library now manages a total of 295 terabytes of digital content, including 66 terabytes of digital material preserved by our partners across the nation.

The overwhelming challenge the Library faced a decade ago—and continues to confront in its third century—is how to superimpose the dynamic world of digital knowledge and information onto the still-expanding world of books and other traditional analog materials. The print publishing universe continues to flourish, particularly in the developing world. So how can the Library preserve and seamlessly integrate these two worlds so that we can continue to comprehensively provide Congress and the American people the objective and dependable knowledge that is needed now more than ever before?

The Library's basic mission of acquiring, preserving and making accessible the world's knowledge and the nation's creativity is not changing. But the amount of information and the explosion in the number of creators are driving the greatest revolution in the generation and communication of knowledge since the advent of the printing press.

We are proud that the Library is yielding profoundly valuable information and educational resources for the nation. We are bringing together both the historical digitized materials and the born-digital content that together provide a strategic and unique resource for the nation. No single institution can collect, save and provide access to digital content in the future. Almost all of the Library's digital initiatives involve learning to work in new ways, in a networked environment, where we are working with others to amass critical content and deliver new and improved services. For example, the Library is leading a network of trusted agents that save and deliver "at risk" digital content under NDIIPP. The Law Library of Congress is working with a network of countries all over the world to save official legal documents used by Congress. We are taking our digitized and unique historical materials to a network of schools and teachers across the country. These are just three examples of strategic new directions for the Library, consistent with our core purpose of furthering human understanding and wisdom.

By no means, however, will we be able to collect everything that is digital. We will continue to do what we have done for 200 years: identify and select what is critical to a universal collection of knowledge and information.

When Congress created NDIIPP in 2000, it appropriated \$100 million in no-year funds to sustain this enormous effort in a decade-long program. In order to complete funding for other critical priorities in FY2007, NDIIPP's unobligated funds were tapped. Prior to the rescission, we were on the verge of making our next set of investments for the work of our current partners, as well as reaching out to new communities. At risk is not only the work of partners across the nation but essential infrastructure and content for the Library's mission to serve Congress. The fuller extent of the lost investment as a result of the rescission is \$84 million—\$47 million in direct funding plus \$37 million in matching funds already committed to the pending investments.

The Library's NDIIPP plans call for expanding the NDIIPP network to include demonstration projects for preservation of important state government records (legislative data, court records and other state information) that is of vital interest to Congress. We have already made commitments to 35 states to preserve state-based digital geospatial, legislative and agency records identified by the Law Library of Congress and the Congressional Research Service as essential for public policy discussions in Congress.

We are witnessing the transformation to a society where instantly available, reliable and credible information will be as indispensable as electricity, water and transportation, and we are proud to play a large role in this change. We urgently hope that Congress will find a means, working with the Library, to provide funding to allow NDIIPP to complete its essential work on behalf of the nation.

The pending work that will have to be cancelled if funding is not restored includes a project led by North Carolina in partnership with Kentucky, Georgia, Florida and West Virginia to collect and preserve state and local geospatial information; a project led by the Washington State Archives to develop a regional digital archive with Alaska, Louisiana, Georgia, Maine, Idaho, Colorado, Montana and Oregon; and a project led by Minnesota in collaboration with Mississippi, California, Kansas and Utah to collect and preserve state legislative information.

At the time of the rescission, we were also finalizing another new NDIIPP project called Preserving Creative America. This is an initiative to join with commercial producers of creative content—digital film, music, photography, other forms of pictorial art and even video games—in developing strategies for the preservation of American creativity in all its forms. Preserving Creative America will help us identify common problems and solutions that are shared by private industry as well as libraries and archives.

As we begin the second half of the Library's unprecedented effort to preserve at-risk digital content, the Library of Congress and our committed network of partners risk losing the resources that have already been invested – and the benefits of important digital preservation work. It is in our national interest to preserve the born-digital information of today to ensure that Americans will be more enlightened and competitive tomorrow.

## **The Challenges Ahead**

The enclosed overview of the Library's digital milestones makes it clear that every part of the Library has utilized technology in order to remain current in service to our customers, and to ensure that the core of the Library— its collections – continues to reflect our commitment to serve Congress and the American people. I will be joined today by my colleagues who direct each of the Library's service units. They will highlight their key areas of focus and the challenges they face to remain current. Their relationship to the Library's core mission is critical.

The Copyright Office has been housed at the Library since 1870 and has been responsible for the enormous growth of the Library's collections as the mint record of American creativity. The deposit of copyrighted material has brought millions of books, films, sound recordings, prints, maps, photographs,

and – in the digital age – databases and Web sites to the Library’s collections. Because Congress made the wise decision to place the Copyright Office in the Library (after it had been in both the Executive and Judicial Branches), it is Congress that deserves the praise for preserving what its citizenry creates.

For some time now, the Copyright Office, like the rest of the Library, has been faced with evolving digital technologies that are transforming how copyrighted works are created and disseminated. In order to receive born-digital works for registration and deposit in the collections of the Library and to improve operational efficiency, the Office embarked on a major reengineering program seven years ago. I am pleased to report that we are nearing completion of reengineering and are on schedule to achieve full implementation of our new processes and systems later this year. Objectives of the program include providing Copyright Office services online, ensuring prompt availability of new copyright records, providing better tracking of individual items in the workflow, and increasing acquisition of digital works for the Library of Congress collections.

The overall responsibility for the Library’s collections, not including the Law Library, rests with Library Services. The Library revolutionized access to information with the creation of the MARC record in 1960. Now nearly half a century later, the Library is leading a nationwide group on the future of bibliographic access; the results of the group’s work will be made available this fall.

The Library’s goal is to provide the researcher with the optimum amount of information regardless of format. Over the next five years, the Library will need to be able to move high-quality, authentic, valid content to the Web where it can easily and conveniently be identified and retrieved. We will build on the deep knowledge of the Library’s curators and subject specialists who have traditionally served as mediators between the collections and the user. Curatorial expertise will also be moved to the Web to accompany content.

### **The Challenge of Globalization: World Digital Library and GLIN**

The Library of Congress is making good progress in its initiative to build a World Digital Library (WDL) for use by other libraries around the globe. The project is supported through funds from nonexclusive public and private partnerships. Our first partnership is with Google, which has provided a \$3 million grant to plan and begin pilots for the WDL.

The WDL will draw upon the experience of the Library of Congress and other national libraries and cultural institutions from around the world to create an unprecedented collection of significant primary materials in digital form that document the achievements of many different cultures. Content will come through digitization of unique and rare materials, including manuscripts, maps, rare books, musical scores, sound recordings, films, photographs, drawings, and other materials. Most of the material will be older and thus free of copyright restrictions.

As the Global Legal Information Network (GLIN) expands its reach beyond its current 46 jurisdictions and continues to focus on emerging democracies, it will provide Congress with up-to-date information about new laws and legal trends in key nations worldwide.

Domestically, the Congressional Research Service (CRS) will continue to utilize the emerging technologies of “Web 2.0” such as podcasts to link the work of CRS analysts directly with their customers in the United States Congress. CRS issued its first electronic issue briefs in the early 1970s and launched the Legislative Information System (LIS) in 1997. The XML version of LIS will be launched in 2010. Meanwhile, CRS has continued to refine the ease with which Members and staff are able to utilize the capacity of Congress’ policy staff in CRS.

## **The Library of Congress: America's Information Reserve**

The dynamism of America's democracy depends on knowledge and our access to it. If we take steps now to collect, preserve and make accessible the exponentially growing body of knowledge, we will leave to our descendants an invaluable legacy. We continue to stress the principles of free and equitable access, as well as the absolute necessity of long-term preservation.

Just as the Library has acquired, preserved and made accessible more than 134 million traditional analog items (books, manuscripts, maps, music and movies), we are now applying the skills and values of traditional librarianship to the digital world. I have been told by members of Congress and their staff that if they want information, they simply find it on Google, and you can indeed find a flood of information on Google – sometimes hundreds, and even thousands, of sources for a single query. Our goal is to integrate the best available electronic information into the knowledge, judgment and wisdom contained in books and in the minds of our curators so that Congress and the American people continue getting the same authentic, reliable information and knowledge that have been the hallmark of the Library since its inception in 1800.

We must transform much of our workforce into a new kind of “knowledge navigator” able to draw equally on new digital materials and traditional artifactual items. And we are helping develop standards and protocols for the electronic sharing of bibliographic records just as the Library did for the print world by making its cataloging records available to others at the dawn of the 20th century.

We are not just creating endless digital data files; we are giving our collections context and making them increasingly accessible to the world. Congress' foresight in encouraging the creation of THOMAS in 1995 gave the public a new, simple way to access legislative information. THOMAS remains the predominant free-access tool for the public to find out what is going on in Congress, and we continue to work with the Government Printing Office to add information from the pre-digital era, back to the 100th Congress. Similarly, the Legislative Information System (LIS), designed to serve Congress' needs, has become the cornerstone of legislative work being done by both Members and Committees.

Building on our early digital experiences, we are now, with Congress' support, preparing to provide added service to Congress and the public in new ways:

- Providing digital analytic support on over 150 current legislative issues, available 24/7 from the Congressional Research Service (CRS);
- Providing links from the Library's catalog records to detailed descriptions of publications and tables of contents;
- Planning and designing the upcoming release of a new program of Digital Talking Books for the blind and physically handicapped;
- Reengineering our public delivery of copyright services to accommodate future growth in electronic registrations;
- Collecting and making accessible first-person stories under the Veterans History Project and other documentary efforts to capture and preserve histories of ordinary citizens;
- Working with teachers and university faculty in nine states to integrate our primary-source digital collections into K-12 curricula;

- Expanding our international capacity and outreach through our Law Library's Global Legal Information Network and the World Digital Library;
- And creating LC Net, a Congress-only Web site providing online service for book loans, tours, and event planning at the Library.

The scope of our digital strategy encompasses every aspect of the Library and envisions our playing a central role for the nation in three ways: (1) **digitizing and distributing** online for educational purposes primary materials from the Library of Congress and other repositories, (2) **gathering and preserving** in the Library and other cooperating institutions important digital material produced elsewhere and in danger of disappearing for use by Congress and the nation, and (3) **converting** as many of the Library's processes and products into electronic and digital forms as possible.

**Sustaining our collections.** There are certain overarching goals for the future. Determining how to select, acquire, and store the digital and online works that are required to keep our collections complete enough to meet the information needs of the Congress of the United States is our first priority. Collection strategies must be strategic, current, and agile. We will need to collect and evaluate electronic databases, multimedia creations, digitally linked resources, and digital material in formats yet to be invented. We will develop processes for both physical and digital items received through copyright registration and mandatory deposit that will protect intellectual property rights. The Library's work through NDIIPP has taught us that the traditional library notions of fair use and requesting "best editions" need to be rethought in the digital age. Under the aegis of NDIIPP, the Library has convened the Section 108 Study Group (referencing the section of current copyright law that provides limited exemptions for libraries and archives) to prepare findings and recommendations that can bring copyright law into alignment with current technologies.

**Preservation.** As digital works are added to the collections, we need a technology infrastructure that will ensure that their content will be available for future generations. We are taking the lead on developing national solutions and making R&D investments to ensure long-term storage, preservation and authenticity of digital content that recognizes the certainty of continued technological change.

**Access.** Digital technology has helped us provide ways to deliver content to a vastly broader range of library users. In the last ten years, we have made millions of items of all kinds from our collections widely available through the Internet. We need to understand how users want to access and navigate through our collections and, to the maximum extent possible, meet the users' requirements, rather than simply imposing a system on them. We will think differently about how we describe material, working closely with other libraries that rely on records we create. We will encourage the use of bibliographic records created outside the Library when appropriate. Our own cataloging efforts will focus increasingly on creating metadata for Internet searchers, to keep online material "readable" from one generation of hardware and software to the next, to authenticate the accuracy and reliability of electronic copies, and to secure them against tampering and unauthorized users and adjust any violations of U.S. copyright law and international agreements.

**Our workforce.** The Library's fulfillment of its mission will always depend on the foresight of our staff. Library staff need to use and, to varying degrees, become expert in how changing technologies apply to our work.

To achieve these goals, we have begun an intensive strategic-planning process that will ultimately transform our collection policies, institutional infrastructure, buildings, and workforce. The Library has developed an agency-wide framework for program assessment of every division and support office. Con-

gressional support has already enabled us to reengineer copyright functions and to create a state of the art National Audiovisual Conservation Center. We are developing new roles for key staff to become objective “knowledge navigators” who can make knowledge useful from both the artifactual and the digital world.

We have shared with Congress some of our ongoing efforts to ensure the professional development of our staff, training, mentoring, and performance planning and evaluation. We have a large number of staff who are retirement-eligible, and we will have to hire many new staff with specialized skills that are often hard to find.

### **Preservation Infrastructure in the Digital Age: The National Audiovisual Conservation Center**

A significant new asset for the Library of the 21st century is the National Audiovisual Conservation Center (NAVCC) located in Culpeper, Virginia. It is a center to develop, preserve and provide broader access to the Library’s unique, comprehensive and highly valued collection of the world’s audiovisual heritage. Because the NAVCC matured during a period of rapid development in digital preservation, its final plans now include significant digital technology that involve state-of-the-art approaches and components.

Unprecedented in size, scope and funding for the Library of Congress, construction of the NAVCC has been made possible by a three-way partnership among the Library of Congress, the Packard Humanities Institute (PHI) and the Architect of the Capitol. Authorized by Congress (P.L. 105-44), the NAVCC has been built and funded by the Packard Humanities Institute at a projected cost to PHI of more than \$150 million. This will be the largest single private gift in the history of the Library – and one of the largest ever made by a private foundation to a government entity. Congress has generously appropriated funds to support the facility and purchase preservation and storage equipment at Culpeper; and this has been a splendid private-public collaboration.

For the first time, the Library will consolidate its more than 5 million item audiovisual collections, currently held in less-than-ideal conditions in three states and the District of Columbia, at one state-of-the-art facility where they can be stored and preserved in an environment with the most appropriate temperature and humidity and made more easily available to scholars from around the world. The current design provides for 25 years of collections growth.

The Library’s audiovisual collection consists of more than 1 million moving images of theatrical films, newsreels, television programs, educational, industrial and advertising material; nearly 3 million audio collection items including commercial sound recordings, radio broadcasts, and early voice recordings of historical figures, as well as more than 1.7 million supporting documents, screenplays, manuscripts, photographs, and press kits. To date, more than 2 million items have already been transferred to the Culpeper facility.

NAVCC incorporates the best of proven digital technology in systems that are being developed in a highly modular fashion, allowing nimble and cost-effective responses to future preservation and access needs. Cutting-edge policies and procedures developed at the Center will be adopted elsewhere, both internally and by the broader library and archival community throughout the country and the world. NAVCC will fully integrate the acquisition of born-digital and converted material into a single processing flow.

The construction of the Packard Campus began in August 2003. The 415,000-square foot complex will include four buildings. Under Phase 1, the Collections Storage and Central Plant were turned over to AOC and the Library in 2005. Phase 2, the remainder of the site (including the Conservation Building with staff offices, preservation labs, and a 200-seat theater), is scheduled for delivery in the spring of 2007; a separate building will house two large storage pods containing 124 specially constructed



vaults for the delicate and combustible nitrate film collection.

The NAVCC campus is largely underground, except for the west front of the Conservation Building, which will curve out from the side of the mountain in a half circle. In a novel and complex landscaping feat, the top and side of Mount Pony were scraped off the building site and set aside during construction. The earth is being replaced over the tops of the completed buildings, and the mountain slope and surrounding landscape are being replanted with 7,700 trees, shrubs and plants showcasing 75 different species—making the site the largest reforestation project on the East Coast. David Packard has not only generously funded this campus, but he has also added many features to the buildings and its surroundings that should make the campus more efficient and elegant.

### **Jefferson Building New Visitors Experience: Lifelong Learning**

Currently, about 1.4 million visitors each year tour the Library's magnificent Thomas Jefferson Building, which Congress restored magnificently in the 1990s. In 2008, the Library's New Visitors Experience (NVE) will open in the Jefferson Building and will connect millions of new visitors to Capitol Hill with the architecture, creativity and history represented at the Library. This experience will complement exhibits in the Capitol Visitor Center (CVC), to which it is being connected by a new passageway.

The NVE will create an interactive, innovative and evolving experience with the Library of Congress to inspire and sustain lifelong learning and creativity for a greatly increased number of visitors (estimated to be as many as 3.5 million annually). The NVE will be unique in its focus on Web-based learning with the development of MyLOC.gov as a means for individual visitors to capture their experience at the Library and to later explore the Library's vast collections and other resources on-line.

Visitors will be able to enter the Jefferson Building through either the new passageway from the Capitol Visitors Center or directly through the grand bronze doors above the Neptune Fountain. They will experience the art and architecture of the building, entering into the adventure of learning by means of a "Passport to Knowledge." Visitors will be able to personalize their visits and arrange for subsequent electronic learning experiences on their own computers.

The New Visitors Experience will include an orientation gallery; an interactive tour of the Great Hall and view of the Main Reading Room; exhibits showcasing the "Creation of the United States"; Thomas Jefferson's own Library; and a gallery of materials spanning the history and cultures of the indigenous peoples of Mexico, Central America and the Caribbean, including the exhibition of the famed Waldseemuller Map of 1507, often referred to as "America's birth certificate."

Anticipating the opening of the passageway from the Capitol Visitors Center to the Jefferson Building, the Library has already begun raising the substantial private funds needed to make the New Visitors Experience a reality. It will celebrate the Congress' role in bringing knowledge into the lives of an even larger audience. This will be accomplished with private contributions and without any major reconstruction of Jefferson Building space.

My colleagues and I eagerly look forward to meeting the challenges and opportunities involved in integrating the digital world into our traditional artifactual collections. We strive every day to make increasingly accessible the human knowledge that sustains us as a free people in a dynamic society and economy. The Library looks forward to working with you and the rest of the Congress to furthering human understanding and wisdom, one of the noblest of pursuits and the most precious of gifts to future generations.

###