It took **two centuries** for the Library of Congress to acquire its **29 million books** and more than **100 million other items**: manuscripts, motion pictures, sound recordings, maps, prints, photographs.

Today it takes only **15 minutes** for the world to **produce an equal amount of information in digital form**.
A Message from the Librarian of Congress

It took two centuries for the Library of Congress to acquire its 29 million books and 131 million other items: manuscripts, motion pictures, sound recordings, maps, prints, photographs. Today it takes only 15 minutes for the world to produce an equal amount of information in digital form.

This staggering statistic looms over our urgent need to collect and preserve our nation's cultural and historical record that is increasingly being produced digitally and in no other form. Collecting these so-called born-digital materials is essential if the Library is to continue to fulfill its role in the 21st century to "preserve and sustain a universal collection of knowledge and creativity for future generations."

If the amount of information in the Library of Congress is being created nearly 100 times per day, can we collect and preserve it all? I can say definitively that the answer is no. What cannot be answered with such certainty are these questions: What is worth preserving? Who will decide what to preserve? Who will preserve it? And Who will pay for it?

These are the questions we are working to answer as part of our unprecedented National Digital Information Infrastructure and Preservation Program. This program, led by the Library of Congress at the request of the U.S. Congress, is a nationwide cooperative endeavor that is forming a network of digital preservation partners dedicated to collecting, preserving and making available the information that future students, teachers, scholars and lifelong learners will rely on as the nation strives to retain its premier position in a world where globalization has increased competition exponentially.

During fiscal 2005, the program made significant progress toward achieving its goal. Many partners joined the network, and we formed a special working group to study the intellectual property issues libraries everywhere face in handling and serving digital materials to their patrons.

This is not the Library's first time in uncharted waters. In fiscal 1995, when we first began making digitized versions of our rare and unique American historical materials available online, we were in uncharted waters. No other cultural institutions were making large portions of their collections accessible on the World Wide Web.

In fiscal 2005, our National Digital Library Program offered more than 10 million items in American Memory as the award-winning site observes its 10th anniversary. More than 330 thematic presentations are available from the collections of the Library and those of its partners.

Those partnerships we formed with other major libraries and archives gave us a head start when we began to look at ways to build the technical infrastructure to preserve digital materials. We have learned a lot about interoperability of diverse electronic collections originating from diverse institutions and building an effective distributed network.

Thanks to the U.S. Congress, the supporter of the largest collection of information in the history of the world, we are able to look ahead with optimism as we build upon our extraordinary past to meet the future demands for knowledge, the cornerstone of our dynamic democracy.

James H. Billington
The Librarian of Congress

A Message from the Associate Librarian

During fiscal 2005, our national program to collect, preserve and make available born digital materials of historical importance achieved several milestones while it took on new challenges.

The initial group of formal partners in the National Digital Information Infrastructure and Preservation Program (NDIIPP) met together for the first time in January 2005 in Washington. These eight consortia, which comprise 30 institutions, have agreed to collect and preserve specific types of significant digital information as part of our growing network of digital preservation partners. During two days of meetings, the partners listened to presentations from key Library of Congress staff and program advisors, established working frameworks for the future and also met in smaller subgroups to discuss common concerns in the areas of digital content selection and collection, intellectual property, economic sustainability and technical architecture.

We completed the Archive Ingest and Handling Test in June. We and the four universities with which we worked learned much about what happens to large files of information as they are transferred among institutions. The full report is on our Web site at www.digipreservation.gov. Its conclusions will help inform decisions as we continue to build the technical infrastructure.

The partners reconvened in July and learned about our test of the transfer of large and diverse digital archives, our new intellectual property group, our efforts to reach out to the states and the NDIIPP awards we made in conjunction with the National Science Foundation.

Balancing fairly the rights of creators with the rights of libraries and their users is the reason behind the formation of our Section 108 Working Group. Section 108 of the U.S. Copyright law provides exceptions for libraries to enable them to serve their patrons. This equilibrium is much more difficult to maintain when handling easily replicable digital materials; thus, this group will make recommendations to Congress on changes to the law that recognize issues unique to digital media.

In May, the Library and National Science Foundation awarded a total of $3 million to 10 university teams to conduct cutting-edge research in the field of digital preservation. The institutions are matching these awards with in-kind contributions and will complete their projects in one year.

Institutions at every level of government are facing enormous challenges in saving, preserving and making accessible their digital content. The states are no exception. This is why in April and May we convened three workshops with representatives of the 50 states, the District of Columbia and U.S. territories to determine how they could work together strategically to meet expectations from their constituents that state government records are being preserved for future access.

The digital age has changed the Library of Congress in many ways. Through our Web site, we are now able to reach Americans everywhere and share the resources of this extraordinary institution. Yet technology has also presented us with vast challenges. Through the National Digital Information Infrastructure and Preservation Program, the Library of Congress is leading the collaborative effort to build the infrastructure that will sustain the nation's critical digital information.

Laura E. Campbell
The Associate Librarian for Strategic Initiatives
and Chief Information Officer
The Office of Strategic Initiatives (OSI) continued to fulfill its mandate in fiscal 2005 to manage the Library’s digital initiatives and technology requirements. The major focus for the service unit was the National Digital Information Infrastructure and Preservation Program, a congressionally mandated program that the Library is leading for the nation. The program seeks to collect and preserve at-risk digital information in a collaborative framework of partners with defined roles and responsibilities. Because the preservation of digital materials can sometimes run into challenges posed by copyright law, a new independent group was formed in fiscal 2005 to recommend ways that copyright law should be changed to meet the needs of libraries when working with digital materials for preservation as well as other purposes. OSI is also responsible for the management of the Library’s myriad Web sites and its educational outreach programs, which demonstrate to teachers nationwide the value of incorporating electronic primary sources in the classroom.

The Information Technology Services directorate is also part of OSI and provides technology support to all the service units of the Library of Congress.
The National Digital Information Infrastructure and Preservation Program (NDIIPP) continued to move forward with its collaborative partnerships and other activities designed to bring focus to important issues, such as how best to effect large archives of digital material, how current copyright law affects the process of digital preservation, the establishment of the first digital-preservation research grants program and an examination of how the states and territories of the United States are coping with the preservation of digital state records. The program’s Web site is at www.digitalpreservation.gov.

Digital Preservation Partnerships

The eight consortia comprising 36 institutions, which received NDIIPP awards totaling nearly $14 million during fiscal 2004, met twice in Washington during fiscal 2005. The initial NDIIPP Digital Preservation Partnerships form a model institutional network focused on preserving-at-risk digital information. This work depends on extensive collaboration among the partners and with the Library. To help spur collaboration, the Library hosted meetings in January and July 2005 to share information and establish a community of practice among the partners.

Topical “affinity groups,” comprising both partner and Library staff, are another key aspect of building collaboration among the NDIIPP partners. These affinity groups transcend institutional and project boundaries and explore areas that merit common solutions. They also create synergy among the projects and partners, encouraging the institutions to explore new methods of communication and collaboration while developing common understanding.

The groups’ work focuses on four areas:

- collection and selection of digital materials
- economic sustainability of digital preservation projects over the long term
- technical architecture for preserving digital materials and
- rights and restrictions related to the collection and preservation of intellectual property

All the partners have shown considerable progress during the initial year of their projects, particularly in establishing practical working frameworks and developing digital preservation tools and services. A summary of progress over the past year follows, by project lead institution.

- CHANNEL 13/WMET, NEW YORK (EDUCATIONAL BROADCASTING CORPORATION). With the help of a number of experts, the project team examined technical formats in terms of practicability and suitability for video preservation. Work also focused on selection and appraisal criteria.
- UNIVERSITY OF CALIFORNIA AT SANTA BARBARA. UCSB and its partner, Stanford University, worked to define specific requirements for a long-term digital repository of geospatial data that would serve as a node in the national digital preservation network. The teams developed model agreements for data submission, based on consultations with prospective depositors, articulated the basic principles of data preservation that would underlay the technical architecture; and began work on a registry of geospatial digital formats.
- NORTH CAROLINA STATE UNIVERSITY. The project team designated itself as the North Carolina Geospatial Data Archiving Project to emphasize the scope of its work, which includes long-term

Office of Strategic Initiatives 2005 Annual Review
preservation and access to data, as well as to act as a catalyst to engage state and local agencies in building archival solutions. They have made substantial progress in establishing collaboration with many entities across North Carolina.

University of Illinois at Urbana-Champaign Library. Initial attention has been devoted to developing tools to facilitate the selection of important Web sites, particularly in support of state archives and libraries. Project partner OCLC delivered the first of its Web Archives Workbench tools (Discovery tool and Properties tool).

California Digital Library. During its first six months, the California Digital Library Web-at-Risk Project developed a network of curatorial partners to select and capture significant government Web pages, which are especially volatile and at risk of loss. CDL and its partners convened focus groups and issued an online questionnaire to help define Web content selection criteria.

The number of individuals and organizations that produce digital material is far larger, and growing much faster, than the number of institutions committed to preserving such material.

Emory University. The MetaArchive project partners have built a distributed preservation network based on the LOCKSS (lots of copies keep stuff safe) caching software. Each partner has installed a networked vault computer and prepared at least one collection for the partners to harvest into the distributed archives.

University of Michigan Inter-University Consortium for Political and Social Research. The partners in the Data Preservation Alliance for the Social Sciences (Data-PASS) have developed a series of common guidelines to govern their collaboration: a content selection policy, appraisal guidelines and policies that will govern the processes and responsibilities associated with acquiring the content covered by the partnership. The project has reviewed nearly 200,000 historic database records describing research grants awarded over the past 30 years by the National Science Foundation and the National Institutes of Health to identify important data at risk of loss.

University of Maryland Robert H. Smith School of Business. The Birth of the Dot Com Era project has worked to develop a novel legal and institutional framework to accession legal records associated with 2,000 technology start-ups during the 1990s. The project team assembled an advisory panel of national experts in the fields of confidentiality, privacy and bankruptcy to help them protect the records while seeking traditional waivers of confidentiality from former clients. Based on this success, the project’s initial one-year award was extended for another two years.

Archive Ingest and Handling Test

In June 2005, the Library entered into a joint digital preservation project with Old Dominion University, Department of Computer Science; Johns Hopkins University, Sheridan Libraries; Stanford University Libraries & Academic Information Resources; and Harvard University Library to explore strategies for the ingest and preservation of digital archives. The 12-month project was supported by Information Systems Support Inc.

The Archive Ingest and Handling Test (AIHT) was completed in June 2005. The test was designed to identify, document and disseminate working methods for preserving the nation’s increasingly important digital cultural materials, as well as to identify areas that may require further research or development.

The AIHT partners applied various preservation strategies, using a digital archives donated to the Library by the Center for History and New Media at George Mason University. The archives is a collection of 77,000 digital images, totaling 209 GB video related to the Sept. 11, 2001, events. The transfer of these 12 gigabytes of digital content was being used to emulate the problems that arise in digital preservation and to test possible solutions.

In addition to testing the intake of the archives into diverse systems, the participants studied the difficulties in transferring large and complex digital archives from one institution to another. This is a critical piece of any larger digital preservation effort, as the number of individuals and organizations that produce digital material is far larger, and growing much faster, than the number of institutions committed to preserving such material. Thus, any practical preservation strategy requires mechanisms for continuous transfer of content from the wider world into the hands of preserving institutions.

Results of the project:

• produced a first generation of tools and processes for institutions to obtain sufficient physical control over digital objects to preserve them

• validated with practical experience the need for implementation guidelines for the use of the METS standard (the METS schema is a standard for encoding descriptive, administrative and structural metadata regarding objects within a digital library, expressed using the XML schema language of the World Wide Web Consortium)

• stress-tested six repository implementations and provided useful feedback to the developers of those implementations and

• validated the transfer of data as the point of standardization.

At the end of the 12-month test, the Library and its partners were working on a final report, to be made available in fiscal 2006.

Library of Congress-National Science Foundation Research Awards

In May 2005, to university teams received a total of $3 million to undertake pioneering research to support the long-term management of digital information. The Library and the National Science Foundation made the awards as the result of a partnership formed to develop the first digital-preservation research grants program.

Research supported by these awards will help produce the technological breakthroughs needed to keep very large bodies of digital content securely preserved and accessible over many years.

The projects are exploring challenging topics, such as preserving rich oceanographic data from hundreds of deep-sea submersible missions; automating methods to describe digital objects and place them in secure archival storage; testing how to preserve digital video when it is first created; and preserving complex three-dimensional digital content. All the projects are expected to produce study results in one year.

Following are the winning lead institutions, their partner institutions and the subject area of the project:

Institutions: University of California San Diego, Scripps Institute of Oceanography and San Diego Supercomputer Center; Woods Hole Oceanographic Institution. Title: Multi-Institution Testbed for Scalable Digital Archiving. Summary: These two institutions will develop a multi-terabyte digital repository to preserve data from more than 1,600 oceanographic research projects. The collaborating institutions will test processes for automatic archival ingest (acquisition), metadata extraction, validation and access control, and will also explore methods for management of rights-protected data.

Institution: University of Maryland. Title: Robust Technologies for Automated Ingestion and Long-Term Preservation of Digital Information. Summary: This project will explore automated ingest and verification for distributed digital collections. It will also develop and test a preservation architecture that can "evolve gracefully" as technology changes and that is interoperable with different computer platforms.

Institution: Drexel University. Title: Digital Engineering Archives. Summary: This project will work with decades of three-dimensional Computer Assisted Design (CAD) engineering design and production data that currently have very limited preservation options. Researchers will use international standards to convert complex design data into more readily preserved content and will use the results to educate the engineering community about three-dimensional data preservation options.

Institution: University of California San Diego, San Diego Supercomputer Center. Title: Digital Preservation Life Cycle Management: Building a Demonstration Prototype for the Preservation of Large Scale Multimedia Collections. Summary: This project will demonstrate a preservation life cycle management process for video content. Researchers will develop and document a practical preservation process for mixed collection of both legacy and "born digital" video material.

Institution: University of Arizona. Title: Investigating Data Provenance in the Context of New Product Design and Development. Summary: This undertaking will investigate ways to automate metadata capture through an innovative partnership with Raytheon, a commercial defense and aerospace systems supplier. Methods to develop "self-aware self-describing" production and design digital data will be explored.

Institution: University of Michigan. Title: Incentives for Data Producers to Create Archive-Ready Data Sets. Summary: The project will examine incentives for data producers to deposit "archive-ready" data sets. Focus will be on collaboration between producers and archives, including identification of a process for
The intent of this project is to bring together leaders of the Moderate Resolution Imaging Spectroradiometer (MODIS) archive community with computer science researchers to discuss new distributed approaches to managing MODIS satellite data, which currently has a volume of about 2 petabytes.

**INSTITUTION:** University of North Carolina at Chapel Hill  
**TITLE:** Preserving Video Objects and Context: A Demonstration Project  
**SUMMARY:** Development of rich descriptive terms and a process for applying them to digital objects is the focus of this study. Attention will also be given to demonstrating a cost-benefit methodology.

**INSTITUTION:** Johns Hopkins University  
**TITLE:** Securely Managing the Lifespan of Versions in Digital Archives  
**SUMMARY:** This project will study technologies for secure deletion of information to protect personal privacy and provide a mechanism to ensure that any unwanted data is retained along with preserved data.

**Section 108 Study Group**
The independent Section 108 Study Group, convened under the aegis of the National Digital Information Infrastructure and Preservation Program and co-sponsored by the U.S. Copyright Office, held its first meetings in 2005. The goal of the newly formed group, named after the section of the U.S. Copyright Act that provides limited exceptions for libraries and archives, is to prepare findings and make recommendations to the Librarian of Congress by mid-2007 for alterations to current work in digital preservation, the types of issues they face and how these commonalities of interest can be leveraged to advance the NDIIPP collaborative partnership network.

The governments of the 50 states, District of Columbia and U.S. territories today produce much of their information in digital form with no analog equivalent. There is a growing need and urgency to preserve this information before it deteriorates, is altered or forever lost through format or technological obsolescence. State governments are grappling with the same issues that other producers and keepers of digital information share. They want to know how to select from and preserve the ever-mounting volume of electronic information that their states produce.
generously of their time and expertise to provide invaluable input toward the long-term preservation initiative.

Digital Formats Web Site
In November 2004 the Digital Formats Web site (www.digitalpreservation.gov/formats) became publicly available. The site provides information about digital content formats, including the identification and description of formats that are promising for long-term sustainability. The site is intended to support strategic planning at the Library of Congress and to aid other institutions in their digital preservation efforts. It also provides an inventory of information about current and emerging formats, including the identification of the detailed documentation that will be needed to manage content created or received in these formats through the content life cycle. As the site evolves, it will present information pertaining to strategies for sustaining preferred formats, including appropriate practices and tools.

Los Alamos National Laboratory Research Project
Los Alamos National Laboratory’s Research Library was awarded a $50,000 grant from the Library of Congress in November 2004. The grant is being used to support research and development of tools that will help address complex problems related to collecting, storing and accessing digital materials.

The first phase of the project is nearing its successful completion. The goal is to adapt existing standards for transferring content between repositories and to use other standards for digital-object storage. The mechanisms and tools developed in this phase are being tested on digital journal materials of the American Physical Society.

Electronic Deposit for Electronic Journals Project
The purpose of the E-Deposit for E-Journals Project is to build a prototype system by the end of fiscal 2006 that will ingest, archive and create access to electronic journals and other digital content in the Library of Congress collections.

The project is a collaborative effort among three service units of the Library: the U.S. Copyright Office, Library Services and the Office of Strategic Initiatives. The project will consider the near- and long-term needs of library users, as well as the technologies available, by studying other systems and technologies in order to build the most efficient system using available resources.

Electronic journals are becoming an increasingly large part of overall journal resources in libraries today. The need to be able to handle digital collections, both for mandatory copyright deposit and general library acquisitions, is essential for the Library to remain a vital resource for Congress and the American public.

Working groups have been established, and these groups have begun pooling their resources, tallying the content demographics in the digital environment, assessing their interdependencies and planning their collaborative approaches to building a Library-wide ingestion tool.

Web Content Capture Project
During its second year, the Web Capture Team worked to establish broader institutional support for, and participation in, capturing at-risk content from the Web. The team established a regular stream of content identification and selection to harvest the content within a reliable technical environment. The team also worked with custodians of the materials Library-wide to archive the content.

In collaboration with representatives from the Congressional Research Service, the Law Library and Library Services, the Web Capture Team worked with recommending officers and curators to expand the Library’s collection of Web content by 17 terabytes of data comprising more than 3,500 Web sites. To date, the Library has captured a total of 34 terabytes of Web content. The Web materials relate to a wide range of critical events, including the 2004 U.S. elections, the war in Iraq, recent Supreme Court nominations and Hurricane Katrina.

The Library is a member of the International Internet Preservation Consortium and worked through that organization to develop common Web capture tools, processes and practices.

COLLABORATION IN ART

Al Hirschfeld (1903–2003), in a 1996 documentary, was called “The Line King.” He was the king of line drawings because with just a few strokes he could capture the essence of his subject.

Just as Hirschfeld’s works represent a collaboration between the artist and his subject, the Library’s digital initiatives bring together various units of the institution as well as external institutions in programs that provide all Americans with access to the riches of the national library and its partners.
Thomas Jefferson drafted the Declaration of Independence in Philadelphia in June 1776. It was an arduous task. The draft that Jefferson labeled the “original Rough draught,” a portion of which is above, became the foundation of the final document, Revised first by Adams, then by Franklin and then by the full committee, 47 alterations were made to the text before it was presented to Congress on June 28. After voting for independence on July 2, Congress continued to refine the document, making 39 additional revisions to the draft before its final adoption on the morning of July 4. The “original Rough draught” embodies the changes that were made at each step.

The National Digital Library Program represents the work not only of the materials’ creators but also of many Library staff who digitized these materials and placed them online where they could be seen by all.


National Digital Library Program

The National Digital Library Program continues to be one of the premier Web sites of the federal government, with millions of users across the country and around the world. During fiscal 2005, the main home page at www.loc.gov was redesigned with an enhanced interface and site architecture. The new design offers more attractive graphics and is easier to navigate than its predecessor. The American Memory Web site, and its companion Today in History site, also received this visual and architectural upgrade.

Other Web pages such as those devoted to the Poetry and Literature Center, Webcasts, Library of Congress Calendar of Events, Sales Shop, news releases and the search page gateway also received the same treatment. The Library’s site received an “Outstanding Website” award from the Web Marketing Association in September.

American Memory and Global Gateway Web Sites

American Memory, the most visited of the Library’s Web sites, offers 10,374,033 digital items in 133 thematic presentations from the collections of the Library and its partners.

Seven new multimedia collections were added in fiscal 2005:
- “September 11, 2001 Documentary Project”
- “Freedom’s Fortress: The Library of Congress and World War II”
- “James Madison Papers at the Library of Congress”
- “Women of Protest: Photographs from the Records of the National Woman’s Party”
- “World War I Newspaper Rotogravures”
- “The Jedediah Hotchkiss Map Collection”
- “Rochambeau Maps from the American Revolutionary Era”

Global Gateway, the Web site for international materials, added four new collections:
- “Cuneiform Tablets: From the Reign of Gudea of Lagash to Shalmanassar III”
- “Selections from a Polish Declaration of Admiration and Friendship for the United States of America”
- “The Kraus Collection of Sir Francis Drake”
- “France in America” (in collaboration with the National Library of France)

Other American Memory collections were augmented with new materials:
- “Civil War Maps”
- “Panoramic Maps, 1847-1929”
- “The American Revolution and Its Era: Maps and Charts of North America and the West Indies, 1750-1789”
- “America from the Great Depression to World War II: Photographs from the FSA-O.W.I., 1939-1945”

Global Gateway collections that were expanded with new content are:
- “The U.S., Spain and the American Frontier: Historias Paralelas” (in collaboration with the National Library of Spain)
- “The Atlantic World: America and the Netherlands” (in collaboration with the National Library of the Netherlands)
Other collections in various areas of the Web site also added new materials.

The number of digital files in the new collections totals 53,761; the number of files added to preexisting collections was 359,417, for a grand total of 911,317 in fiscal 2005.

**America’s Library**

America’s Library, a Web site for children and families, added 11 features to its “Meet Amazing Americans” section: Martin Luther King Jr., Thurgood Marshall, John Adams, James Madison, Andrew Jackson, James Monroe, John Quincy Adams, Martin Van Buren, William Henry Harrison, John Tyler and James Polk.

**Legislative Information System, LCnet and THOMAS Web Sites**

The Legislative Information System, for use by Congress and the Congressional Research Service, also received a revised interface. LCnet is a new Web site for congressional staff. The site provides Congress with the latest information from the Library that is of particular interest to the legislative body. Users can also request Library tours and constituent meetings through the site.

The THOMAS public Web site devoted to the workings of the U.S. Congress was upgraded with several features, including a new function that allows users to search multiple congressional sessions at the same time. The site was scheduled for another upgrade in early fiscal 2006.

**Internal Web Support Services**

**Library of Congress Employment Web Site**

This site, which features the Library’s new Web interface design, vastly improved accessibility to the Library’s employment information, including job postings, internships, fellowships and volunteer opportunities.

**Library of Congress Intranet Sites**

OSI developed and maintained several Intranet sites for the Library, notably those for the Office of Workforce Diversity and the institutionwide Emergency Preparedness Intranet.
A cast of thousands is responsible for every spacelift—from the engineers and the manufacturers, to the crew on the ground and the astronauts. This photo, from 1959, shows “weightless” astronauts in an experimental flight. According to its caption, “Before another 10 years are over, the National Aeronautics and Space Administration hopes to be launching men on missions which may take weeks or months to complete. The Mercury program may prove that men can endure weightlessness for several hours.” No one can predict the future with certainty, especially when technology is involved. Who could have predicted that the Library of Congress would one day serve more patrons online through its Web site than over enter its doors on Capital Hill? The Library’s digital initiatives have enabled the institution to become a national library in the truest and best sense of the word.


In addition to the National Digital Information Infrastructure and Preservation Program, American Memory and Global Gateway, the Office of Strategic Initiatives works with many areas of the Library of Congress on programs such as Web sites and public outreach initiatives.

Wise Guide Web Site

The Library’s Web magazine, the Wise Guide (www.loc.gov/wiseguide) is a portal to the millions of resources from the Library available online. Each month, readers are offered seven “articles” with links to the most interesting materials in all the Library’s many and varied Web sites. The project is a collaborative effort of OSI and the Public Affairs Office.

National Digital Newspaper Program

In March 2005 the Library and the National Endowment for the Humanities announced that six institutions had received more than $1.9 million in grants in the National Digital Newspaper Program (NDNP), a new, long-term effort to develop an Internet-based, searchable database of U.S. newspapers now in the public domain. Two-year projects in California, Florida, Kentucky, New York, Utah and Virginia each will digitize 100,000 or more pages of each state’s most historically significant newspapers published between 1900 and 1910. When completed, digitized newspapers will be made available through the Library’s Web site (www.loc.gov).

The following six institutions received the first NDNP grants to digitize papers in their respective states from the first decade of the 20th century:

- University of California, Riverside, $400,000
- University of Florida Libraries, Gainesville, $320,000
- University of Kentucky Research Foundation, Lexington, $310,000
- New York Public Library, New York City, $352,643
- University of Utah, Salt Lake City, $312,693 and
- Library of Virginia, Richmond, $201,226.

The first phase of the NDNP is expected to result in the Library’s ingest of 60 terabytes of digitized newspaper content from six states. A tool has been developed for the content providers and collectors from the six states to facilitate submission of digital newspaper content to the Library of Congress for long-term preservation. Development of a repository to enable the material to become accessible online has begun.

National Book Festival

In time for the National Book Festival on Sept. 24, OSI developed and produced the second edition of the CD-ROM “The Library of Congress: Inspiring the World with Knowledge.” This new version offers increased interactivity and more in-depth content than its predecessor. Videos, interactive educational games and digital materials from the Library’s collections are among the materials accessible through this disk. The disk was initially made available at the National Book Festival and is being distributed at other venues as well. Educational outreach staff managed the book festival pavilion devoted to Library of Congress initiatives.

Creativity Across America

“Creativity Across America” is a new long-term Library of Congress initiative to be celebrated across the nation. The “Song of America Concert Tour with Thomas Hampson” is one of the first programs being conducted as part of this initiative. OSI developed a new interactive Web site for the tour, with tour schedules, resources for teachers and Hampson performances of selected American song classics. Educational outreach specialists developed educational programming such as a full-day teacher institute and a 60-page book, called “Lyrical Legacy,” on the history of song and poetry in America.

Office of Strategic Initiatives 2005 Annual Review
COLLABORATION IN SPORTS

Christy Mathewson (1880–1925) was one of baseball’s most formidable pitchers, having won 373 games during 15 seasons, mostly for the National League’s New York Giants. He was one of five players inducted into the first class of the Baseball Hall of Fame in 1939, along with Walter Johnson, Babe Ruth, Honus Wagner and Ty Cobb. According to Mathewson’s Hall of Fame plaque, he was the “greatest of all pitchers in the 20th century’s first quarter. ...

... Matty was the master of them all.”

The Library of Congress has a collection of 2,100 early baseball cards, dating from 1887 to 1914, that it received as a donation in 1954 from Lincoln biographer Carl Sandburg. The cards were distributed in packages of cigarettes. Many of the most treasured items in the Library’s collections came through gifts from generous donors who wanted to share these materials widely. The Library’s educational outreach initiative places these materials in context so that they can be used in the classroom. These and other primary sources stimulate critical thinking in young minds.

[Christy Mathewson], 1909–1911, Benjamin K. Edwards Collection, Prints and Photographs Division, on the Web: American Memory Web site, memory.loc.gov

EDUCATIONAL OUTREACH

OSI includes a team of experts in education who assist educators nationwide in the use of online primary sources in the classroom. This staff provides these services through a variety of programs. In addition to the collaborative projects described above, the Library’s education experts managed the following programs and events.

An Adventure of the American Mind

An Adventure of the American Mind (AAM) is a program designed to train in-service and pre-service classroom teachers and college teacher education faculty to access, use and produce curriculum using the Internet and the digitized primary source materials from the collections of the Library. This congressionally mandated program is currently active in Colorado, Illinois, Indiana, Pennsylvania, Virginia, North Carolina and South Carolina and has 23 funded partners. The program was created by Congress and is implemented by the Library of Congress with the Educational and Research Consortium of the Western Carolinas.

In fiscal 2005, new AAM partners were added at Quincy University in Illinois and the Center for Congress in Indiana.

Among the AAM activities, the Library hosted orientation and directors meetings, made presentations at 13 workshops in four states, reaching more than 600 educators; hosted 28 in-house workshops for 660 educators; and contracted with the Center for Children and Technology for a research study on the best practices for the current AAM national program.

Workshops and Presentations

Library staff offered presentations in more than 40 venues nationwide, reaching approximately 7,500 educators. They also staffed an exhibition booth at three major educational conferences: the National Council for Social Studies (4,000 attendees), the Florida Educational Technology Conference (8,500 attendees) and the National Educational Computing Conference (67,000 attendees). Educational outreach staff were also in the Library’s exhibit booth and the annual and midwinter meetings of the American Library Association.

Workshops and presentations held at the Library reached hundreds more educators.

The Learning Page

The Learning Page Web site (www.loc.gov/teachers) provides content specifically developed for teachers and their students. This year, six new Community Centers became available. The activities offer teacher-tested materials based on themes such as elections, political cartoons, the Civil War and poetry.

Four lesson plans were added in the areas of "Recreation Yesterday and Today," "Sea Changes," "Exploring Community" and "Suffragists." Fifteen Collection Connections went online. Collection Connections guide teachers to the most useful materials in the thematic American Memory collections.

Staff of the Learning Page conducted 10 online live chats with educators across the country.
the “segregation in schools was lifted on the second day of the 109th Congress. The search engine for Congressional Research Service (CRS) products was upgraded to a newer version and several new search features were added, among them multiple Congress searching, member searching, state delegation searching and improved topical searching. The Legislative Information System (LIS) was also enhanced in fiscal 2005. A statistical package to dynamically track Alert Service usage and an interface extracting information from the LIS database and inputting it to the Congressional Bills Project in XML format were added. A custom XML-tag library was completed to expedite the transfer of Senate Library-published resources to the LIS. These resources include appropriations tables back to 1983, cloture motion tables back to 1985, Hot Bills, Supreme Court information back to 1789 and veto information.

Another fiscal 2005 ITS accomplishment in support of Congress was the linking of the texts of CRS Products with related legislative bills and laws. These links enable anyone reading the text of a CRS product (e.g., an Issue Brief) to link directly to any bills or laws related to that issue and, conversely, to link from the bill or law to the Issue Brief.

Looking to the future, ITS procured an integrated state-of-the-art search engine as part of an effort, in concert with others in OSI and CRS, to take advantage of XML-based legislative metadata and content used in THOMAS and the LIS.

Service to the Public

The Library’s online computer resources attracted even more users than in previous years. In fiscal 2005, there were 3,68 billion Internet transactions, or hits, on all the Library’s Web sites, compared to 3.36 billion in fiscal 2004. The American Memory Web site continues to be the Library’s most popular, registering more than 155 million hits in fiscal 2005:—more than 208 million above the previous year. The America’s Library Web site for children and families handled 239 million hits in fiscal 2005, compared with 218 million the previous year. The THOMAS public legislative information system and online Library Exhibitions remained popular Web sites, receiving 210 million and 154 million hits, respectively.

In fiscal 2005, ITS continued working closely with its OSI service unit and other Library service units to expand the number of digital collections accessible to the public through the Library’s Web site.

ITS also enhanced searching of the Library’s digitized visual materials by integrating a thesaurus into the Prints and Photographs Online Catalog. Standardized subject, genre and image process terms and their cross-references now enable researchers to look up related terms or to look for pictures indexed with the term.

Support of Digital Initiatives

ITS worked closely with others in its OSI service unit to the nation that it must be unanimous. Warren persuaded some of his associates to change their minds and vote with the majority. Thurgood Marshall, chief counsel for the NAACP argued the case. Attorney George E.C. Hayes won a companion case on the same day. The two received a show of support from fellow civil rights defender James M. Nabrit.

Information Technology Services supports the institution-wide technology needs of the Library by adopting new technologies that are employed throughout the Library.

[George E.C. Hayes, Thurgood Marshall, and James M. Nabrit, congratulating each other, following Supreme Court decision declaring segregation unconstitutional], 1954, Prints and Photographs Division. On the Web: Prints and Photographs Online Catalog, loc.gov/rr/print/pphome.html.

The Information Technology Services (ITS) directorate of the Office of Strategic Initiatives supports the technology needs of all Library service units and personnel. In fiscal 2005, ITS continued to support the Library’s ability to adopt new technology by establishing and enhancing capabilities to monitor, evaluate and respond to changes and trends, not only in technology, but also in business, law and society at large.

Service to Congress

In collaboration with other OSI staff, the Legislative Information System and THOMAS received significant upgrades during the first phase of the THOMAS Redesign Project.

The Legislative Information System (LIS) implemented a wide array of display enhancements across the nine LIS databases for the start of the 109th Congress. The search engine for Congressional Research Service (CRS) products was upgraded to a newer version and several new search features were added, among them multiple Congress searching, member searching, state delegation searching and improved topical searching. The Legislative Information Online Update System (xLIS) was also enhanced in fiscal 2005. A statistical package to dynamically track Alert Service usage and an interface extracting information from the LIS database and inputting it to the Congressional Bills Project in XML format were added. A custom XML-tag library was completed to expedite the transfer of Senate Library-published resources to the LIS. These resources include appropriations tables back to 1983, cloture motion tables back to 1985, Hot Bills, Supreme Court information back to 1789 and veto information.

Another fiscal 2005 ITS accomplishment in support of Congress was the linking of the texts of CRS Products with related legislative bills and laws. These links enable anyone reading the text of a CRS product (e.g., an Issue Brief) to link directly to any bills or laws related to that issue and, conversely, to link from the bill or law to the Issue Brief.

Looking to the future, ITS procured an integrated state-of-the-art search engine as part of an effort, in concert with others in OSI and CRS, to take advantage of XML-based legislative metadata and content used in THOMAS and the LIS.

Service to the Public

The Library’s online computer resources attracted even more users than in previous years. In fiscal 2005, there were 3,68 billion Internet transactions, or hits, on all the Library’s Web sites, compared to 3.36 billion in fiscal 2004. The American Memory Web site continues to be the Library’s most popular, registering more than 155 million hits in fiscal 2005:—more than 208 million above the previous year. The America’s Library Web site for children and families handled 239 million hits in fiscal 2005, compared with 218 million the previous year. The THOMAS public legislative information system and online Library Exhibitions remained popular Web sites, receiving 210 million and 154 million hits, respectively.

In fiscal 2005, ITS continued working closely with its OSI service unit and other Library service units to expand the number of digital collections accessible to the public through the Library’s Web site.

ITS also enhanced searching of the Library’s digitized visual materials by integrating a thesaurus into the Prints and Photographs Online Catalog. Standardized subject, genre and image process terms and their cross-references now enable researchers to look up related terms or to look for pictures indexed with the term.

Support of Digital Initiatives

ITS worked closely with others in its OSI service unit...
Strategies for Sustaining the Infrastructure

Numerous IT infrastructure-related projects were completed by ITS. All were designed to increase the efficiency of Library operations. These include:

- Development of the Centralized Network Monitoring System, which provides a view into the operation and health of the Library's network. It also supplies automated notification and resolution of predefined problems.
- Network standardization to produce a homogeneous network communications environment (50% complete).
- Centralized workstation management through the OSN system to foster a consistent, efficient, maintainable and secure environment using a standard baseline configuration.
- A Server Architecture Assessment that identified replacement requirements for Madison Data Center server components over the next two to three years.
- A Storage Architecture Assessment that will help ITS forecast Madison Building Data Center needs for storage, architecture and infrastructure growth over the next two to three years.
- An Architecture Assessment that recommended the IT infrastructure needed for the National Audiovisual Conservation Center in Culpeper, Va.
- Retirement of the token ring network and the establishment of a fully functional gigabit Ethernet network for the entire Library.

ITS continued to provide units throughout the Library with state-of-the-art technical support services in the areas of trouble desk support, cable installation, network support, workstation configuration and server support.

During fiscal 2005, ITS responded to more than 50,000 requests for workstations-related products and services, including help desk support. In addition, ITS phone operators responded to more than 54,000 incoming calls.

The ITS Digital Scan Center increased its support of approved digital conversion projects for Library Services, scanning 20 collections for inclusion on the Library's Web site. Significant scanning services were also provided for the Congressional Research Service, the Copyright Division and the Interpretive Programs Office. By the end of the fiscal year, the Digital Scan Center had produced approximately 100,000 high-resolution digital images. The center now offers a variety of products for sale to the public.

Over the past several years, the Library of Congress has initiated numerous efforts designed to ensure the availability, reliability and integrity of the Library's systems.

In fiscal 2005, the ITS Assistive Technology Demonstration Center continued to provide assistance to physically challenged Library employees. Approximately 30 staff members were provided technological accommodations to assist them in performing their work. A number of physically challenged Library patrons were also provided assistive technology accommodations.

Optimization

During fiscal 2005, ITS focused on the implementation of best practices in the areas of Systems, Development Life Cycle (SDLIC), Project Management (PM) and IT Security. Managers and staff were trained in all three areas to further the Library's mission.

In February 2005, ITS was reorganized to improve alignment and support of the Library's mission, Strategic Plan and Digital Strategic Plan. Two new ITS divisions were created, Operations and Research and Development. The Operations Division consists of the four technology groups providing computer operations, help desk support, laboratory and multimedia services. The Research and Development Division consists of five development groups focused on the Library's service units and a database administration group. The reorganization also included the creation of a security group and strengthening of the technology assessment group.

and with Library Services on a range of research and development projects to support the Library's digital initiatives for the 21st century. These efforts include the National Digital Information Infrastructure and Preservation Program, as well as all new collections and materials added to the Library's Web sites.

Two of the more ambitious projects started in 2005 are the National Digital Newspaper Program (NDNP) and the Web Collection Management System (WCMS). NDNP is a long-term effort to develop an Internet-based, searchable database of historically significant U.S. newspapers in the public domain. It is an outgrowth of the soon-to-be-completed U.S. Newspaper Program, a Library-coordinated effort by individual states to inventory, catalog and microfilm local newspapers. WCMS will provide curatorial and technical staff with the online capability to track the nominations, selection, permissions and capture of Web content that supports the Library's collection policies and responsibilities. NDNP and WCMS mark the first Library attempts to develop applications using the Open Source development model.

For the Integrated Library System, ITS continued to make improvements in methodology for identifying and managing the significant increases in database queries from public users.

Business Applications and Support of Staff

ITS personnel continue to provide subject matter expertise to both internal and external organizations to further the effective exchange of information among organizations and agency lines. ITS involvement with the Unicode Consortium and with numerous internal working groups, including the e-Deposit of e-Journals and the Internet Operations Group Search Subcommittee, will facilitate the development of a robust, flexible and reliable digital environment that efficiently and effectively serves the Library of Congress.

ITS worked with the Copyright Office to produce significant improvements in the format, structure and display of COPICS (copyright registration) data, making it more accessible and easier to use. These changes allow searching of copyright registrations and recorded documents using the same software used to search other Library collections.

ITS staff built a special test environment to support the implementation of an initial release of the Copyright Business Process Reengineering system. This release, called Pacer One, allows all motion picture registrations to be processed through the new system. A special test partition for the new Copyright Pre-registration process was also implemented.

The number of enterprise applications supporting the Library's service units and infrastructure staffs that ITS maintains has grown to more than 710. These applications are continually being improved and upgraded to newer versions by the Resources and Development Groups. One major effort this past year included fine-tuning the performance of the Momentum financial application deployed at the end of fiscal 2004 and migrating all of the financial applications, including reporting and time and attendance, to a newly created, separate financial hosting environment.

Library of Congress

Office of Strategic Initiatives 2005 Annual Report

25
of capture techniques that address virtually any scanning need and is recognized throughout the Library for its high-quality scanning work. As part of the establishment of the ITS Operations Division, the Digital Scan Center was incorporated into the ITS Multimedia Group.

Videos produced by audiovisual specialists in the ITS Multimedia Group are now featured on the Library of Congress main home page. An archives of more than 600 video programs, with 182 productions completed in fiscal 2003, is available for viewing on the Webcasts page (www.loc.gov/today/cyberlk/). These popular programs cover a wide variety of subject areas and were sponsored by more than 20 Library work units. The programs include conferences, symposia, lectures, book talks and documentaries about the collections held at the Library.

For the first time, ITS productions were featured through a national distribution network. The Research Channel aired a series of author talks from the National Book Festival packaged and edited by ITS. A public service announcement about the festival was produced and is regularly aired on the ResearchChannel. ITS created a special animated introduction with original music for the opening of book festival programs.

In May 2005, ITS worked with Brigham Young University (BYU) to present the two-day symposium “The Worlds of Joseph Smith.” ITS provided real-time broadcast of the event to several locations. Technical challenges included simultaneous Web streaming in different formats through a back-channel link between the Library and BYU. The satellite uplink used for the broadcast was made possible by the technical infrastructure developed by ITS.

In addition to the standard workflow of pre-production, production and post-production, ITS demonstrated the technical ability to capture, edit, distribute and archive digital video footage in a single real-time workflow. To organize and track processes and media, ITS implemented Studio Workhorse, an item tracking database and workflow scheduling tool for multimedia projects. Integration of the workflows has allowed the Multimedia Group to meet sharply rising production demands with existing resources.

IT Security

Over the past several years, the Library of Congress has initiated numerous efforts designed to ensure the availability, reliability and integrity of the Library’s systems, applications and information. In fiscal 2004, ITS drafted and the Library issued a new IT Security Policy (LCR 1620). ITS issued a dozen directives during fiscal 2005 providing the details that the service units need to fulfill their role in protecting Library resources. ITS also developed a Library IT Security Plan and is currently working with the other components of the Library to draft their individual IT security plans.

As part of the effort to institute a central IT security program, ITS created the Security Operations Center (SOC) within the newly formed IT Security Group. The SOC provides incident handling and response capability, review of audit logs and monitoring of certification and accreditation activities. The SOC serves the entire Library and also provides IT security advisories to all Library IT personnel.

The Library Security Plan mandates annual IT security training. ITS, working with the Library’s Office of Management and Training, created an online course that all Library staff, contractors and other personnel are required to complete annually before gaining access to Library IT resources and retaining access. Moreover, this training includes staff acceptance of a standard Library Rules of Behavior for Using Information Technology Systems.

In order to support critical applications, ITS created an IT Continuity of Operations Plan (COOP) in fiscal 2003. During fiscal 2005, the IT COOP was refined. ITS staff received extensive training on the IT COOP and a Drive-Away Kit was created and distributed to all the ITS managers and those staff with responsibility in maintaining critical systems.

ITS worked with various service units to certify and accredit (C&A) a number of key IT systems within the Library. These systems were thoroughly analyzed to provide the Library with a clear understanding of the level of risk imposed by their operation. This enables the Library to ensure that measures taken to protect electronic information are commensurate with that information’s value to the Library’s mission. Systems that have undergone such certification and accreditation include the Application Hosting Environment (Intranet and DMZ Hosting Environments), the Financial Hosting Environment, Momentum, the Financial Reporting System, WebTAT, eCO (the Electronic Copyright Office) and the Online Learning Center.

As part of the C&A effort, ITS implemented an IT Security Risk Assessment Development Tool. This tool allows analysts to complete NIST SP 800-30 compliant Risk Assessments far more expeditiously than performing the process manually.

United States Congress

$15 million appropriation over five years

FOUNDING SPONSORS

contributions of $5 million or more

Mr. John W. Kluge

The David and Lucile Packard Foundation

CHARTER SPONSORS

contributions of $3 million or more

AT&T—Lead Corporate Sponsor

Ameritech

Bell Atlantic

Cargill Foundation

Discovery Communications Inc.

Donaldson, Lufkin & Jenrette

Eastman Kodak Company

Federal Express Corporation

The William and Flora Hewlett Foundation

Jones Family Foundation

Glenn R. Jones (Jones International Ltd.)

W.K. Kellogg Foundation

H.F. Lenfest

Robert R. McCormick Tribune Foundation

Occidental Petroleum Corp.

Alexander Papamarkos

The Pew Charitable Trusts

Reuters

Laurence S. and Mary French Rockefeller

Suzanne and Walter Scott Foundation

CONTRIBUTORS

Bankers’ Trust Foundation

Compaq Computer Corporation

R.R. Donnelley & Sons Co.

The Ford Foundation

The Hewlett Foundation Inc.

David H. Koch Charitable Foundation

Carl H. Lindner

Lucent Technologies Corp.

Mellon Foundation

Microsoft Corporation

NYNEX Foundation

Shell Oil Co. Foundation

Texaco Foundation

IN-KIND CONTRIBUTORS

International Business Machines Corp.

LizardTech
Looking Ahead

The Office of Strategic Initiatives (OSI) is meeting the technology needs of the Library of Congress and building the infrastructure for the 21st century library.

Through its leadership of the National Digital Information Infrastructure and Preservation Program (NDIIPP), OSI is also building a network of public- and private-sector partners to ensure the survival of our nation's critical information in digital form.

NDIIPP is reaching its halfway mark with many successes to its credit. The program will continue to add partners to the network of institutions. Soon, this network will reach more than 100 members. In 2006, NDIIPP will make new awards to institutions and organizations and, for the first time, these will include creators and distributors from the entertainment industry. Following up on its three meetings with representatives from the states, NDIIPP will also strategically fund state projects that can contribute to a national effort to preserve important state records.

Internationally, the World Digital Library program will work to add partners to a planned common resource for learning about the cultures of the world.

The educational outreach program of OSI will continue to develop a national network of schools that want to make the Library's online primary sources a part of their curricula. These sources have become essential tools for developing the critical thinking skills in today's young minds that are necessary to meet the challenges of today's competitive global society.

After more than 200 years, the mission of the Library of Congress has not changed. The challenge is to meet the expectations of its many constituencies in an ever-changing environment.