

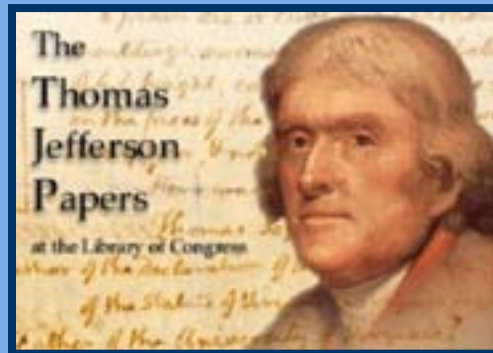


National Digital Information Infrastructure and Preservation Program (NDIIPP)

INTERNATIONAL FEDERATION OF LIBRARY
ASSOCIATIONS

Laura E. Campbell
Associate Librarian for Strategic Initiatives
Library of Congress

LC Online Collections





Digital Information Infrastructure and Preservation Program

- December 2000 -- Enabling Legislation
- October 2002: The "Master Plan"
- January 2003: Congress OKs Plan
- August 2003: NDIIPP Program Announcement
- June 2004: Archive Ingest and Handling Test
- September 2004: NDIIPP Makes Awards of Nearly \$14 Million

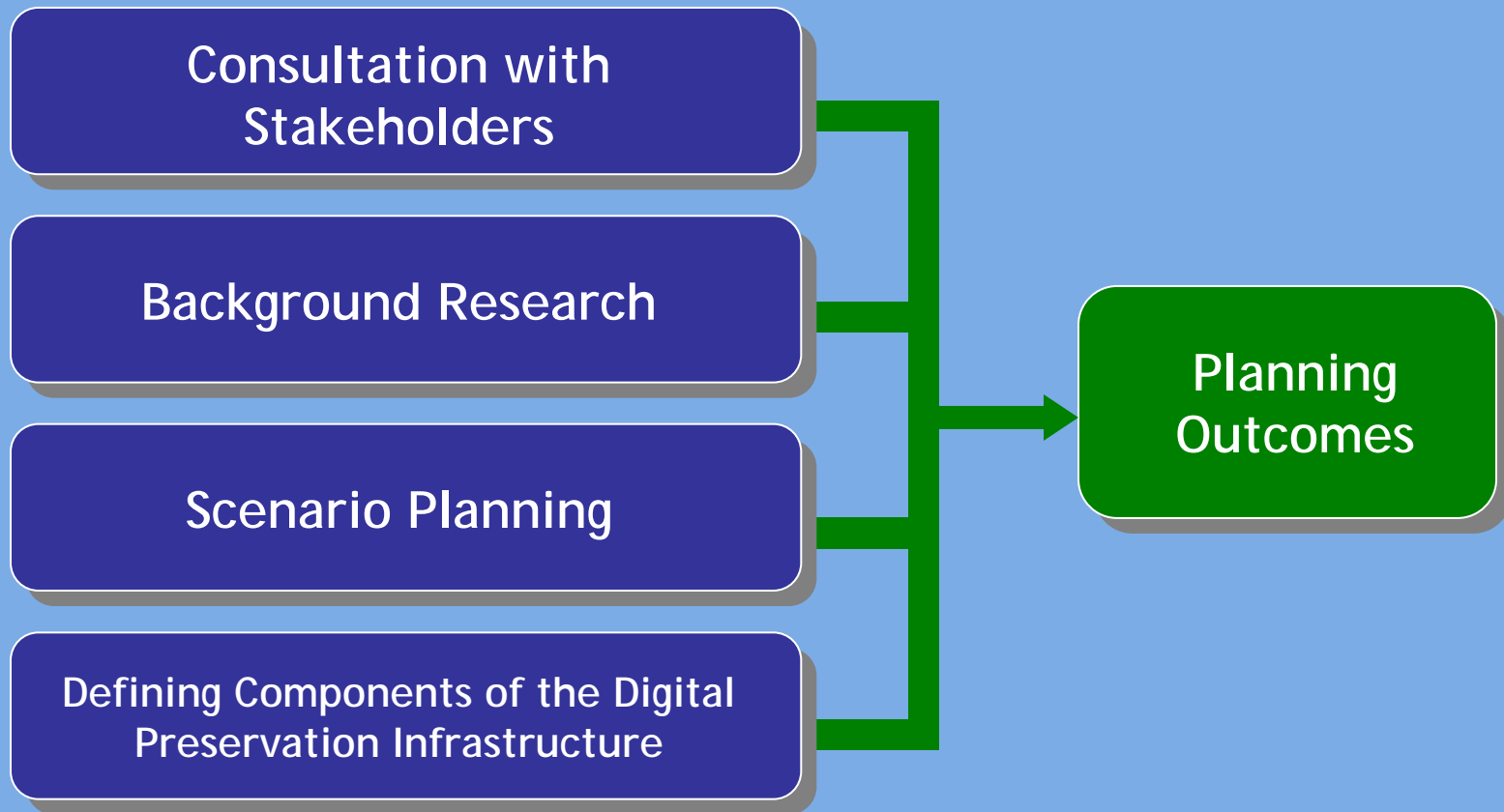


Digital Information Infrastructure and Preservation Program

- April 2005: Section 108 Study Group Formed
- May 2005: NDIIPP-NSF Make Research Awards of \$3 Million
- July 2005: Meeting of NDIIPP Partners
- January 2006: NDIIPP Partners Meeting Draws Largest Crowd

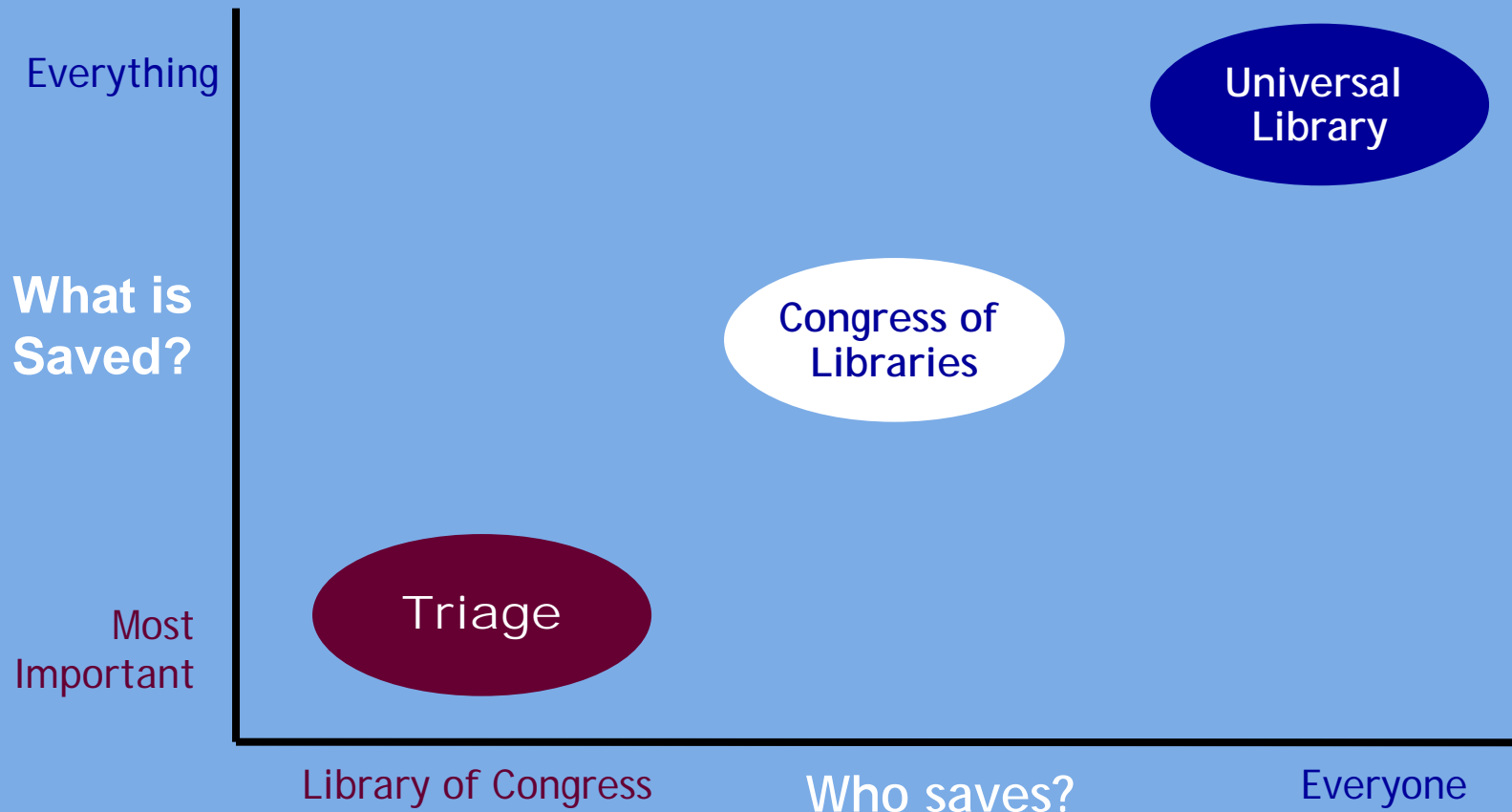


Planning Approach





Plausible Scenarios





NDIIPP Vision

- NDIIPP's vision is to ensure access over time to a rich body of digital content through the establishment of a national network of committed partners, collaborating in a digital preservation architecture with defined roles and responsibilities



NDIIPP Goal

- To develop a nationwide collection and preservation strategy for digital materials in cooperation with the information and technology industry, concerned federal agencies, libraries, research institutions and not-for-profit entities.
- A “Plan for the National Digital Information Infrastructure and Preservation Program,” which we call our “Master Plan,” was submitted to the Library’s congressional oversight committees for approval in October 2002.
- Approval was received in December 2002.



Phased activity, reporting to Congress

- We are making a phased set of investments, what we learn in each phase informs the next
- We are at end of the initial phase and are outlining a second set of investments
- We will make recommendations to Congress in 2010

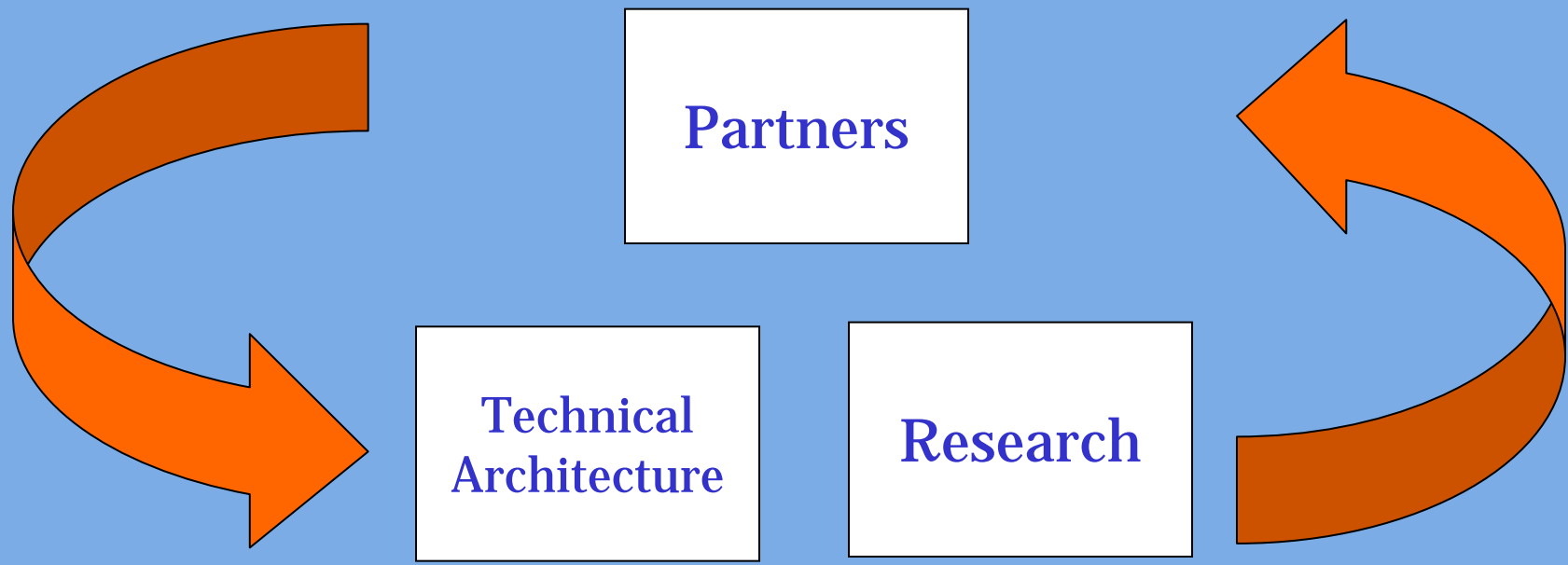


NDIIPP Funding

- Up to \$175 million potentially available
 - \$5 million (approved to develop the “Master Plan”)
 - \$20 million (upon Congressional approval of the plan)
 - \$75 million (subject to \$ for \$ match from non-federal sources)
 - \$75 million (private funds)



NDIIPP Elements





NDIIPP Recommended Initial-Phase Investment Portfolio: \$35M

Building the network of partners	70%
Technical preservation architecture	20%
Basic digital preservation research	10%



Areas of Focus

- Selection and collection development
- Intellectual property
- Technical architecture
- Economics



Current Activities: Network of Partners (36 Institutions)

- **Preservation Partners: At-risk Content**
 - California Digital Library
 - With NYU, U North Texas, Texas Ctr for Digital Knowledge
 - Web archiving tools for government and political information
 - University of California at Santa Barbara
 - With Stanford University
 - Test federated preservation of geospatial images and data
 - Educational Broadcasting Corporation (WNET)
 - With WGBH, PBS, NYU
 - Develop approaches to preserve current-generation digital TV
 - Emory University
 - With U of Louisville, VPI, UVA, FSU, Auburn, Georgia Tech
 - Develop MetaArchive of Southern Digital Culture



Current Activities: Network of Partners

- **Preservation Partners : At-risk Content**
 - University of Illinois at Urbana-Champaign
 - With OCLC, Tufts, Mich State, & state libs of AZ, CT, IL, NC, WI
 - Tools to help select and preserve various types digital material
 - University of Maryland
 - With GMU, 2 law firms, Internet Archive
 - Preserve digital materials about birth of dot.com era
 - University of Michigan
 - With NARA and several social science data and research centers
 - Appraisal, acquire, and preserve social science survey datasets
 - North Carolina State University
 - With NC Center for Geographic Information and Analysis
 - Geospatial data from state and local government agencies

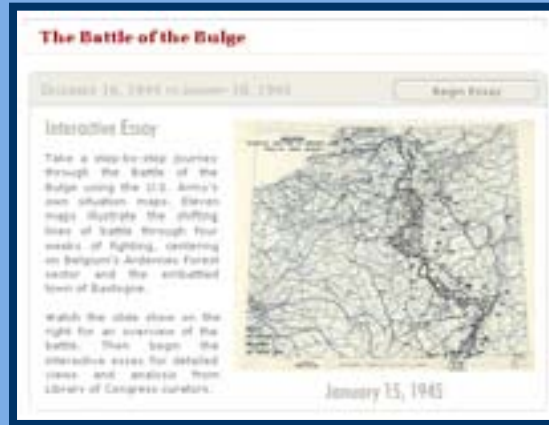


Diverse and Challenging Content

- Social science data sets
- News broadcasting
- Geo-spatial material
- TV programming
- Business records
- Government and political web sites
- Historical multimedia collections



Diverse and Challenging Content





Current Activities: Network of Partners

- Preservation Partners: State Government
 - Focus on state government digital information
 - Workshops with teams from all states representing state libraries, archives, CIOs
 - Developing digital preservation toolkit
 - Report published October 2005
 - *Preservation of State Government Digital Information: Issues and Opportunities*



Current Activities: Network of Partners

Business Model Partnerships

- PORTICO: subscription and archiving service for e-journals
 - For publishers: secure backup
 - For libraries: secure copies for the long-term
 - An “insurance policy” for both publishers and libraries
- San Diego Supercomputer Center
 - Service Bureau for long-term archiving of large volume multimedia content



Current Activities: International Partners

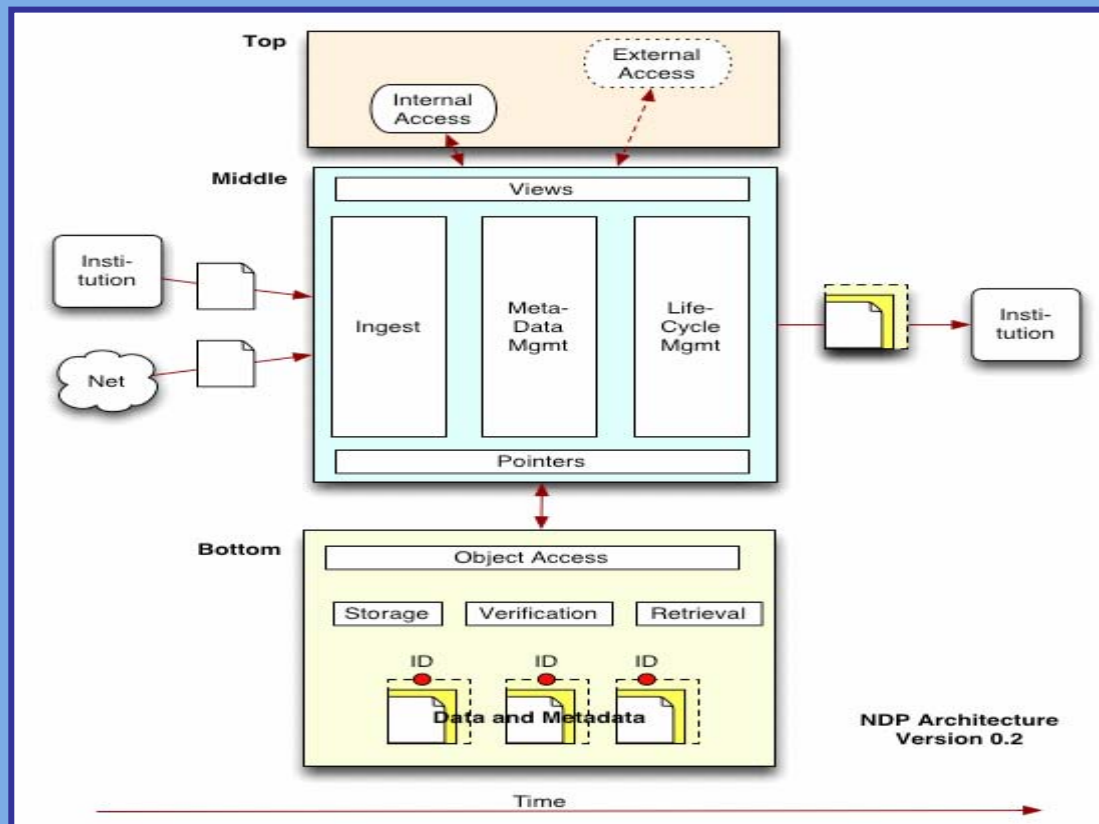
- Digital Preservation Coalition in the UK
- International Internet Preservation Consortium (IIPC)
 - 12 national libraries
 - France
 - Italy
 - Denmark
 - Finland
 - Iceland
 - Canada
 - Norway
 - Australia
 - Sweden
 - United Kingdom
 - Internet Archive, USA
 - Library of Congress, USA
 - Technical specifications for an open source web crawler
 - 5 working groups



Current Activities: Technical architecture

- AIHT
 - 4 institutions and 6 existing architectures
 - Universities: Old Dominion, Johns Hopkins, Stanford, and Harvard
 - Test of ingest, export, and transfer among architectures
 - Identify gaps in functionality
 - Focus on protocols & metadata
- Prototyping with Los Alamos National Lab
 - e-journals, Web sites materials, and multimedia collections

NDIIPP Architecture Version 0.2





Current Activities: Basic Research (10 Partners)

- Cooperative grant program with NSF
 - Univ of Cal, San Diego; Scripps Inst of Oceanography & San Diego Supercomp Ctr; Woods Hole Oceanographic Inst.
 - Testbed for scalable digital archiving, oceanographic research projects.
 - University of Maryland
 - Robust technologies for automated ingestion and long-term preservation of digital information.
 - Drexel University
 - Digital engineering archives, 3d computer assisted design (CAD) engineering design and production data



Current Activities: Basic Research

- Cooperative program with NSF
 - Univ of Cal, San Diego; San Diego Supercomp Ctr
 - Digital preservation lifecycle mgt, large scale multimedia collections.
 - University of Arizona
 - Data provenance in new product design & dev, with Raytheon, defense and aerospace systems supplier
 - University of Michigan
 - Incentives for data producers to create archive-ready data sets



Current Activities: Basic Research

- Cooperative program with NSF
 - Old Dominion University
 - Shared infrastructure preservation models
 - University of Tennessee at Knoxville
 - Globally accessible archive of MODIS (Moderate Resolution Imaging Spectroradiometer) data
 - University of North Carolina at Chapel Hill
 - Preserving video objects and context
 - Johns Hopkins University
 - Securely managing the lifetime of versions in digital archives, secure deletion to protect personal privacy



What we learned

- We knew some things in advance, and these were confirmed by experience
 - The network of partners is distributed (52 partners currently)
 - *This is in the American pattern: for example state governments and US universities operate very independently*
 - Context is one of rapid technology change
 - *For example, when NDIIPP started, there was no iTunes, which (with its cohorts) has strongly influenced the environment for digital content*
 - *There were no blogs, and now people say, "aren't you preserving them yet?"*



What we learned

- We initially thought of the three investment areas as separate, but found that the real projects in all three areas encountered common issues and problems
- Network partners need services and technical tools to do their work



What we learned

- We also found that players in one area were stronger in some elements, while players in another area were stronger in another.
 - Preservation partners
 - Collections, curation
 - Technical architecture
 - Computer science skills, tool making
 - Research
 - Exploration of new ideas and approaches



Other lessons learned . . .

- Complexities in ingesting content
- Challenges regarding the exchange or transfer of content
- Formats: need help to identify, registry
- Tools to make various processes more efficient



About copyright

- Prevalent issue, encountered in many settings
- Became clear that it had to be addressed or little else could happen
- Example: when collecting Web sites, we need to ask permission
 - *We get permissions 24 percent of the time*
 - *People don't say "no," they just don't respond to queries*
- Action: launch the Section 108 study group

About copyright





Current NDIIPP Activity: Copyright Study

- Section 108 is the part of the U.S. Copyright law that governs what libraries and archives can do
- Study Group is a committee of experts charged with developing proposals to update Section 108 in the current digital context.
- Goal is to identify the best possible balance between the rights of creators and copyright owners and the needs of libraries and archives.



Current NDIIPP Activity: Copyright Study

Section 108 Study Participants

- John Wiley and Sons
- Time Inc.
- The Walt Disney Company
- Business Software Alliance
- Association of American University Presses
- JSTOR
- Penguin Group (USA)
- Houghton Mifflin Company
- Universal Mastering Studios
- J. Paul Getty Trust
- Wolff & Godin LLP
- University of North Carolina
- Columbia Law School
- Cornell University Library
- National Library of Medicine
- Columbia University
- American Library Association
- Georgetown University Law Center
- The Andrew W. Mellon Foundation



The BIG lesson

- We need to bring our three groups together for mutual support and next developments
- Play on their individual strengths to develop the collective infrastructure
- Use shared efforts between all to build an improved infrastructure



Synergy more important than new content at this moment

- Building the improved infrastructure through a shared effort is as important as seeking out new content.
- This idea guides our next round of investments.



Next-phase NDIIPP Investments

Strengthen and broaden the content preservation group:

- State and local government digital information
- Commercial content
- Technology support for building the overall network and infrastructure



Next-phase NDIIPP Investments

Strengthen the technical architecture group :

- Develop a distributed storage and exchange platform to test preservation strategies
- Expand current tool development into services for the network
 - Tools to ingest data
 - Support for the transfer and exchange of data



Next-phase NDIIPP Investments

Strengthen the research group :

- Additional basic research grants tied to needs of network partners
- Focus on more efficient ways to manage digital content in all formats

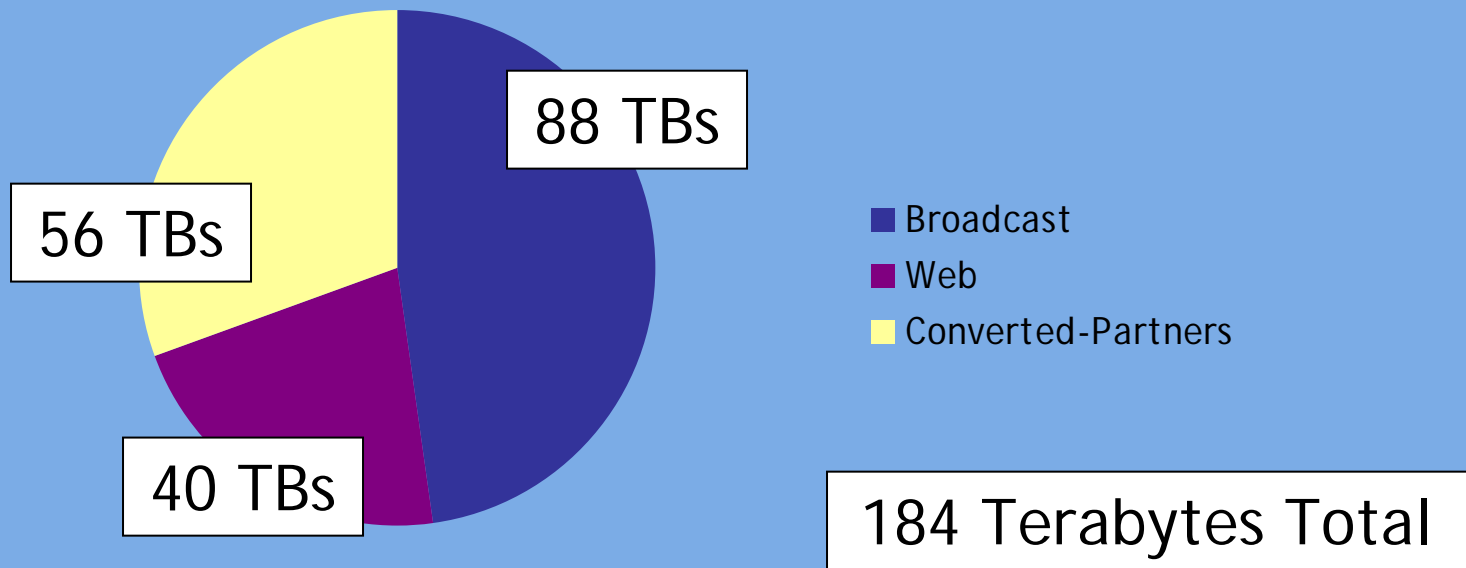


Our role: LC as a Catalyst and a Coordinator

- NDIIPP partners expressed desire for LC to:
 - Serve as coordinator, integrator
 - Focus collaborative activities
 - Shape strategic agenda for cross-cutting legal, technical, economic, collection issues
 - Provide leadership without asserting “top down” ownership



Digital Content Under Management 2006



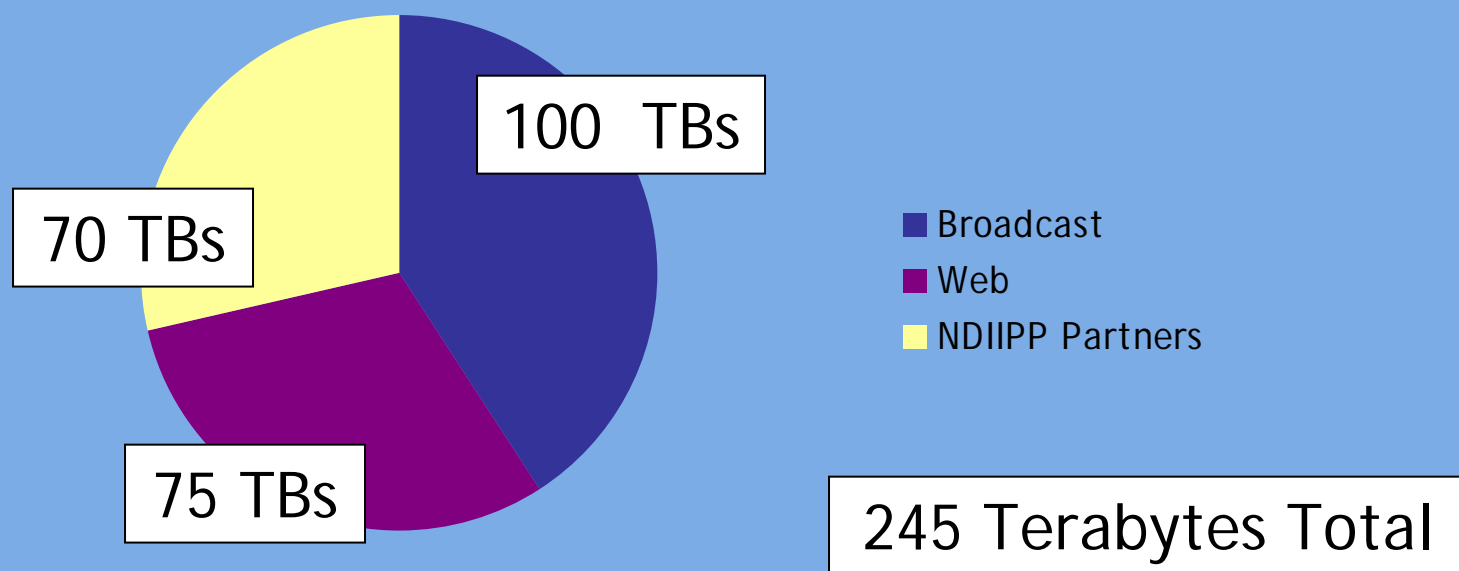


Digital Content





Forecast: Digital Content - LC & Partners 2008





Five Year Outcomes

- 52 partners in a distributed network
- Over 100 terabytes of “at-risk” born-digital content
- Variety of formats: text, numeric, video, images, audio, and geo-spatial
- Multiple Sources: TV and radio, broadcasts, e-journals, private archives, donations, the web
- Preservation standards for digital content ingest and transfer

Digital Universe





The LIBRARY of CONGRESS DIGITAL PRESERVATION NDIIPP Pages [More Search Options](#)

The Library of Congress > Digital Preservation



- [Digital Preservation Home](#)
- [About the Program](#)
- [Technical Infrastructure](#)
- [Partnerships](#)
- [News, Events](#)
- [Contact](#)

Related Resources at the Library

- [Standards at the Library of Congress](#)
- [Digital Collections & Programs](#)
- [U.S. Copyright Office](#)
- [Webcasts from the Library of Congress](#)

Digital Preservation

The National Digital Information Infrastructure and Preservation Program

Information is being produced in greater quantities and with greater frequency than at any time in history. Electronic media, especially the Internet, make it possible for almost anyone to become a "publisher." How will society preserve this information and make it available to future generations? How will libraries and other repositories classify this information so that their patrons can find it with the same ease that they can locate a book on a shelf?

- [Learn more about the Digital Preservation Program](#)

Highlights

The Section 108 Study Group, which is formulating recommendations for changes to copyright law relating to the exceptions and limitations applicable to libraries and archives, will hold public roundtables in March in both Los Angeles and Washington. Participation can be either through attendance at one of the two meetings or by written comment. For more information:

- [Visit the Section 108 Study Group Web site](#)

Representatives of the NDIIPP collaborative development projects, Library of Congress staff, members of the National Digital Strategy Advisory Board and experts in digital preservation met in Berkeley, Calif., in January for the first of their semiannual meetings in 2006.

- [Read the complete report](#)
- [More News and Events](#)

[Site Map](#) | [Contact](#) | [Accessibility](#) | [Legal](#)