DAPS ’08
Workshop on Digital Archive Preservation and Sustainability
September 2008

Martha Anderson
Jane Mandelbaum

Office of Strategic Initiatives
Library of Congress

mande@loc.gov
jman@loc.gov
The National Libraries

- The Library of Congress
- National Agriculture Library
- National Library of Medicine
The Library’s mission is to make its resources available and useful to Congress and to the American people and to sustain and preserve a universal collection of knowledge and creativity for future generations.
Jefferson's Library recreated for the Thomas Jefferson exhibition
Photography by David Sharpe
Thomas Jefferson’s ‘Universal Library’

- Jefferson sold personal library to Congress in 1815
- His library was universal in scope
- Its universality guides the collecting principles of the Library of Congress to this day
Library of Congress in the Digital Age

• We can no longer collect “everything”
• We have formed a network of more than 130 partners to
  – Collect and preserve important digital content at risk of loss
  – Conduct digital preservation research
  – Construct an interoperable technical infrastructure to support long-term preservation
<table>
<thead>
<tr>
<th>Year</th>
<th>Event</th>
</tr>
</thead>
<tbody>
<tr>
<td>1990</td>
<td>Library began digitizing its collections for a pilot called American Memory, which was distributed via CD-ROM to 44 schools and libraries across the country.</td>
</tr>
<tr>
<td>1992</td>
<td>Library began making all of its onsite exhibitions available on the Internet; the first two exhibitions were “Revelations from the Russian Archives” and “1492: An Ongoing Voyage.”</td>
</tr>
<tr>
<td>1994</td>
<td>In 1994, the Library officially began its National Digital Library Program with a promise to digitize and make available 5 million items from the Library by 2000.</td>
</tr>
<tr>
<td>1995</td>
<td>THOMAS Web site of information about the daily bills and legislative workings of the U.S. Congress is launched.</td>
</tr>
<tr>
<td>1996</td>
<td>The Learning Page Web site for teachers and students puts the library’s primary resources in the classroom.</td>
</tr>
<tr>
<td>2000</td>
<td>America’s Library, a Web site for children and families, is launched to put primary sources in the home.</td>
</tr>
<tr>
<td>2000</td>
<td>Congress authorizes the National Digital Information Infrastructure and Preservation Program, to capture and preserve material that exists only in digital form.</td>
</tr>
<tr>
<td>2001</td>
<td>A Web site featuring international materials from the Library and its overseas partners, called Global Gateway, features America’s earlier relationships with other countries.</td>
</tr>
<tr>
<td>2004</td>
<td>National Digital Library Program celebrates 10th anniversary and its 10 millionth digitized item in American Memory online.</td>
</tr>
<tr>
<td>2005</td>
<td>Library announces formation of an effort to build a World Digital Library.</td>
</tr>
<tr>
<td>2007</td>
<td>Chronicling America, an historic newspaper archive, is launched.</td>
</tr>
</tbody>
</table>
Vision: To ensure access over time to a rich body of digital content through the establishment of a national network of committed partners

Approach: To form a national preservation partnership network by collaborating with public- and private-sector institutions dedicated to collecting, preserving and making available important born-digital materials that are at-risk of loss
NDIIPP Background

- Created by federal legislation (PL 106-554) in December 2000
- Up to $175 million available with matching funds
- Guided by National Digital Strategy Advisory Board
- *Preserving Our Digital Heritage*, plan approved by Congress in December 2002 for the National Digital Information Infrastructure and Preservation Program
- Focuses on at-risk born-digital content
NDIIPP Goals

- Work with industry, concerned federal agencies, libraries, research institutions from for-profit and nonprofit sectors to build a network of preservation partners to
  - Develop a national digital collection and preservation strategy
- Support development of improved tools, models and methods for digital preservation by
  - Constructing an architectural framework for preservation
  - Making investments in digital preservation research
DIGITAL CONTENT TYPES MANAGED BY THE LIBRARY

E-JOURNALS
- Scholarly electronic journals
- Small-circulation on-line journals

STATE LEGISLATION & AGENCIES
- Laws, bills, proceedings, committee reports
- Rule-making and regulatory documentation
- Judicial decision, opinions, reports, & rules

GEOSPATIAL
- At-risk state, regional, and local Gov't data (e.g., infrastructure maps, Jurisdictional boundaries)
- Aerial and satellite imagery
- Atlantic and Pacific coastal imagery

WEB SITES & BLOGS
- International political movements
- Elections, Congressional Confirmations, Legal blogs
- Politics & social movements in the Western US
- Hurricane Katrina aftermath

SOCIAL SCIENCE DATASETS
- Regional and national polling data
- Social and economic surveys
- Output of government-funded research (e.g., NSF, NIH)
- Surveys of international opinion (e.g., USIA datasets 1952-1999)

VIDEO
- Foreign news programming (SCOLA)
- U.S. television news
- Public television (e.g., Frontline, Nature)

BORN DIGITAL CONTENT
Functions can be shared or distributed across institutions.
Preservation Architecture Projects

- Archive Ingest and Handling Test
  - Harvard, Stanford, Johns Hopkins and Old Dominion conducted ingest, export, format migration and exchange on a common archive

- San Diego Supercomputer Center
  - Testing utility and building trust in a third-party storage and repository for digital image, text and harvested Web content

- Content Transfer and Exchange
  - Designing, testing and developing best practices for exchange of “packages” of content among partners
LONG-TERM PRESERVATION:
WHAT IS LONG-TERM FOR LIBRARIES?
Long-Term Preservation in Libraries

- Libraries have been in the preservation business for centuries.
- Libraries preserve for the current and future generations of users.
- Libraries don’t go out of business.
• Libraries manage objects and their contents.
• Libraries attempt to manage both objects and contents through metadata -- aka “the catalog”.
• Libraries look backward – they keep the history and provenance of what they have.
• Libraries look forward – they hope to keep what future generations will want.
• Libraries are stewards of their objects and content for current and future generations.
• Each steward needs to be able to turn over their responsibilities to the next steward.
• Libraries have curatorial stewards and technical stewards.
• Libraries deal with objects and metadata at multiple stages of a life-cycle.
• A specific steward will not deal with an orderly sequence from beginning to the end of a life-cycle for a specific object.
• Life-cycle consists of:
  • Select
  • Get/acquire
  • Process
  • Preserve and sustain preservation
  • Access and sustain access
Long-Term Concerns of Libraries: Effects of Disasters and Other Losses

- Libraries have always lost objects through disasters (fires, floods, conflicts).
- Libraries also lose objects through “de-accessioning”.
- Libraries have used geographic distribution of objects in threatening times.
Long-Term Concerns of Libraries: Building on Disasters and Loss

- Libraries build and rebuild.
  - LC fires of 1814 and 1851
  - National Library of Brazil

- Libraries use catalogs and surrogates or copies to rebuild.
  - Thomas Jefferson’s library collection
Library of Congress Rebuilt After Fire of 1851
..And Then…Library of Congress – 1890’s
The Libraries of Congress

Long-Term Concerns of Libraries: Sustaining Trust and the Common Good

- Libraries traditionally rely on a community of libraries for the common good.
- Libraries share and exchange metadata and content.
- Libraries have built common standards for the strength and sustainability of the organizations and the objects.
Long-Term Concerns of Libraries: Preservation of Digital Content

- Should libraries and library users have different expectations for digital content?
- Preservation of Objects
- Stewardship
- Effects of Disaster and Loss
- Building on Disaster and Loss
- Sustaining Trust and the Common Good
Join over 2000 preservation-minded readers and sign up for the monthly newsletter!!!