NDIIPP Elements

Preservation Partners

Technical Architecture

Research
Lead Institutions for the 8 Partnerships

- California Digital Library: Web political content
- Emory University: Southern digital culture
- UC Santa Barbara: Geospatial Data
- NC State University: Geospatial Data
- U of Maryland: Dot.com business records
- U of Michigan: Social science data
- U of Illinois at Urbana-Champaign: State government publications, among other content
- WNET/PBS: Digital television
Goals for the Partnerships

- Identify/select/collection content; communicate strategies for doing so
- Probe intellectual property issues
- Collaborate broadly in developing a shared technical architecture
- Study economic sustainability
- Identify and share best practices
- Learn how to build and incrementally improve a preservation network
Network of Networks

- LC pleased to see interconnections between partners, other networks
- Some examples:
  - Association of Research Libraries
  - Digital Library Federation
  - Coalition for Networked Information
  - UK Joint Information Systems Committee
Broad Categories of Challenges

• Technical

• Legal

• Social

• Economic
• Handling complex objects in a repository
  - Multi-format datasets, projects vs. individual datasets
• Preserving spatial databases
  - The whole is more than the sum of the parts
• Managing time-versioned content
Emerging Issues

- Natural tension between collaboration and tradition of institution-specific approaches
- Much interest in joint infrastructure, but developing shared services will take time
- Existing preservation methods are being stretched
- Balancing a focus on project goals while also considering lots of interesting ideas
Brobeck Archive: Records Overview

- Managed Documents (~3.48 million)
  - Word & Word Perfect (~3.34 million)
  - Excel (~54,000)
  - TIFF, PDF, PowerPoint, etc. (~90,000)

- Relational Databases (~7)
  - 1 database table

- Microsoft Outlook Data (~1,700 users; ~363GB)
  - 1 user's mailbox

- Network Share Drives

Brobeck Clients
- Client A
- Client B
- Client C
- Client D
<table>
<thead>
<tr>
<th>Closed Archive Classifications</th>
<th>Documents</th>
<th>Databases</th>
<th>E-Mail</th>
<th>Network Share Drives</th>
</tr>
</thead>
<tbody>
<tr>
<td>Currently accessible under prevailing archival and legal rules</td>
<td>Currently confidential; potentially accessible in the future under prevailing archival and legal rules</td>
<td>Never accessible under prevailing archival and legal rules</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Court-approved partnership documents</td>
<td>Non-consenting clients’ documents</td>
<td>Objecting clients’ documents</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Consenting clients’ documents</td>
<td>Non-court-approved partnership documents</td>
<td>All personal clients’ documents</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Database rows or e-mails relating to consenting clients or court-approved partnership</td>
<td>All databases</td>
<td>Sensitive partnership documents</td>
<td></td>
<td></td>
</tr>
<tr>
<td>None</td>
<td>All e-mail</td>
<td>None</td>
<td></td>
<td></td>
</tr>
<tr>
<td>None</td>
<td>All network share drives</td>
<td>None</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
Preserving Digital Public Television
What will this project do?

- Identify our “at-risk” *digital* holdings (we are not allowed to migrate any analog materials)

- Plan a long term preservation repository for America’s digital public television programs

- Adopt technical standards and selection criteria for digital programs

- Design a test bed to develop operations for file transfers and a model repository

- Look at the finances of such a facility
Learning and Incremental Development

- We need to get broad feedback and continue learning
- Still no “silver bullet” solution to digital preservation
- NDIIPP considering all viable approaches, working toward gradual development of decentralized, interoperable architecture
- Our partners are modeling this approach