EXHIBIT 101
HathiTrust

Sharing a Federal Print Repository: Issues and Opportunities

May 25, 2011

Heather Christenson
Overview

- About HathiTrust as an organization
- Characteristics of the HathiTrust collection
- Collection development & management
- Changing library landscape
- How print management fits in
HathiTrust is a...

- Partnership of 50+ research libraries, founded in October 2008
- Distributed, collaborative enterprise
- Trusted digital preservation repository
- Digital library
- Service provider
- Platform to collectively explore new library models
Goals

- Comprehensive collection
- Preservation...with Access
- Shared strategies
- Collection management, development
- Preservation
- Copyright review
- Efficient user services
- Openness
Partners involved at many levels

Governance by partners
• Executive Committee
• Strategic Advisory Board

In-kind contribution from partners
• Working Groups
• Operations and development

Funding from partners
Working Groups

- Appointed by Strategic Advisory Board and Executive Committee
- Both operational and strategically-focused groups
- Collections, Communications, Discovery Interface, Full-text Search, Usability, Support
- Now 40+ people across the country
- Expertise from across the partnership
Cost Models

#1: “Contributor” model
- Partners pay per volume deposited
- Economy of scale keep costs low

#2: “Sustaining” model, holdings-based
- Partners share in costs of curating aggregate collection
- Share in uses of relevant materials
- Voice in future directions
- Sustaining common resource
- Quality of services increases
- Costs go down
HathiTrust Content

• HathiTrust content has been curated over time by librarians
  – Mirrors collections of large research libraries
  – Focus on quality
• Expanding Non-Google content
  – Public Domain: Copyright Review Management System
  – Content from non-Google sources
    • Internet Archive, image collections, government publications
Content Sources

* As of May 1, 2011
Content over time

* As of May 1, 2011
Content Distribution

- 26% In Copyright
- 74% Public Domain

8,234,081 – Total volumes
2,102,033 – Public Domain
4,527,381 Book titles
202,649 Serial titles

* As of March 5, 2011
Language Distribution (1)

The top 10 languages make up ~86% of all content.

* As of May 1, 2011
Language Distribution (2)

The next 40 languages make up ~13% of total

* As of May 1, 2011

Statistics and Visualizations
Dates

* As of May 1, 2011

Statistics and Visualizations
<table>
<thead>
<tr>
<th>Collection Development and Management</th>
</tr>
</thead>
<tbody>
<tr>
<td>Collections Committee</td>
</tr>
<tr>
<td>Prioritization of collection development activities</td>
</tr>
<tr>
<td>Appropriate principles for duplicate volumes</td>
</tr>
<tr>
<td>Process for decision-making and prioritization for new content types</td>
</tr>
<tr>
<td>Recommendations for tools and services</td>
</tr>
<tr>
<td>Prioritization of copyright review and rights-clearing processes</td>
</tr>
<tr>
<td>Print management proposal</td>
</tr>
</tbody>
</table>
A global change in the library environment

Academic print book collection already substantially duplicated in mass digitized book corpus

June 2010
Median duplication: 31%

June 2009
Median duplication: 19%
Continuing growth of overlap...
And yet every library is different

- Our median rate of overlap may be the same
- But our overlap profiles will differ by library
- Our use patterns differ
- Our risk profiles differ
- Our roles vis-à-vis our constituencies differ
- Thus, the need to act independently on common data
Extending the holdings database

- HathiTrust print holdings database
  - Basis for new cost model (overlap of in-copyright)
  - Basis for lawful uses (e.g., print disabilities, Section 108)
  - A more complete picture than elsewhere

- Print monograph storage proposal
  - Enable partners to register commitments
  - Establish definitions (e.g., environment, use and condition)
  - Build in cost-sharing: collectively fund those that make commitments
  - Communicate information to partners to facilitate decision-making
Next steps?

- Work to develop draft proposal underway by HathiTrust Collections Committee
  - Focus on monographs
  - Dovetail with other efforts
- Consideration (as part of new cost model), at Constitutional Convention in Oct. 2011
Changing Library Landscape

- Rapidly changing landscape
- Libraries are making these decisions but they are more and more collective decisions
- We cannot afford anymore to do work separately that could be done collaboratively
How to find out more

- Web site “About” section: http://www.hathitrust.org/about
- Twitter: http://twitter.com/hathitrust
- Monthly newsletter: http://www.hathitrust.org/updates
- RSS: http://www.hathitrust.org/updates_rss
- Contact us: feedback@info.hathitrust.org
Thank you!

Heather Christenson
University of California, California Digital Library
heather.christenson@ucop.edu