Collaboration between HathiTrust and University Presses

John Wilkin
HathiTrust Content Growth

- 2008: 2,500,000
- 2009: 5,250,000
- 2010: 8,000,000
- 2011: 10,000,000
- 2012: 12,000,000
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%
Content Distribution

8,625,158 Total volumes
2,297,041 Public Domain
4,722,664 Book titles
209,930 Serial titles

* As of May 1, 2011
Content Distribution

- In Copyright: 73%
- "Public Domain": 27%
- Public Domain (worldwide): 17%
- Public Domain (US): 10%
- Open Access: 0.11%
- Creative Commons: 0.01%

* As of May 1, 2011
Model 1: CC license

• The university press assigns a Creative Commons license to each work that it wishes to make publicly available. The license applied may be the BY-NC-SA version, including, for example, the following explanatory language: “Non-commercial as defined in this license specifically excludes any sale of this work for money, even if the sale does not result in a profit by the seller or if the sale is by a 501(c)(3) nonprofit or NGO.”

• For materials digitized by Google, Google explicitly permits HathiTrust to return to the university press copies of the digital files held by HathiTrust for the press’s own purposes.

• Because the license is a CC license, works are available for full PDF download to all users.

• Implicitly, HathiTrust makes a commitment to archiving the material for non-partner institutions.
Model 2: open license

• The university press grants a specific license to HathiTrust to permit open access to the materials, including full PDF download to all users. This model is similar to Model 1 (the application of a CC license), but does not change the copyright status of the work.

• If content digitized by Google, Google will not permit HathiTrust to return digital files to the press.

• Implicitly, HathiTrust makes a commitment to archiving the material for non-partner institutions.
Model 3: limited open license

• The university press grants a specific license to HathiTrust to permit limited open access to the materials, specifically only 1 page at a time PDF download. The works are open to all users under this model, and not simply authenticated users at partner institutions.

• If content digitized by Google, Google will not permit HathiTrust to return digital files to the press.

• Implicitly, HathiTrust makes a commitment to archiving the material for non-partner institutions.