

Washington University in St. Louis
Washington University Open Scholarship

University Libraries Presentations

University Libraries

Spring 5-4-2012

HathiTrust -- A Gov Docs Repository?

Brian Vetruba

Washington University in St Louis

Follow this and additional works at: https://openscholarship.wustl.edu/lib_present



Part of the [Library and Information Science Commons](#)

Recommended Citation

Vetruba, Brian, "HathiTrust -- A Gov Docs Repository?" (2012). *University Libraries Presentations*. 4.
https://openscholarship.wustl.edu/lib_present/4

This Presentation is brought to you for free and open access by the University Libraries at Washington University Open Scholarship. It has been accepted for inclusion in University Libraries Presentations by an authorized administrator of Washington University Open Scholarship. For more information, please contact digital@wumail.wustl.edu.

HathiTrust--a GovDocs Repository?

**Brian Vetruba, Catalog Librarian/Germanic Studies
Librarian**

Washington University in St. Louis

bvetruba@wustl.edu

Leveraging Your Strengths: Regional Government
Documents Conference | Federal Reserve Bank St. Louis

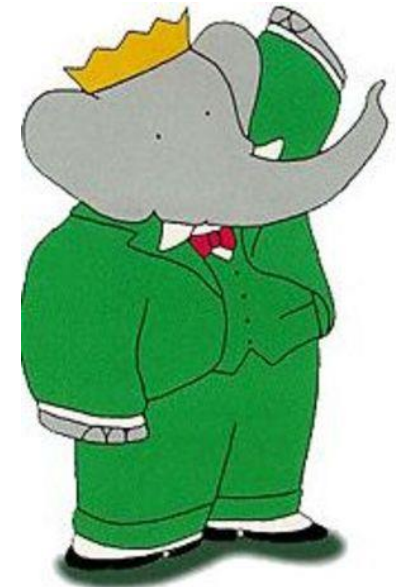
May 4, 2012



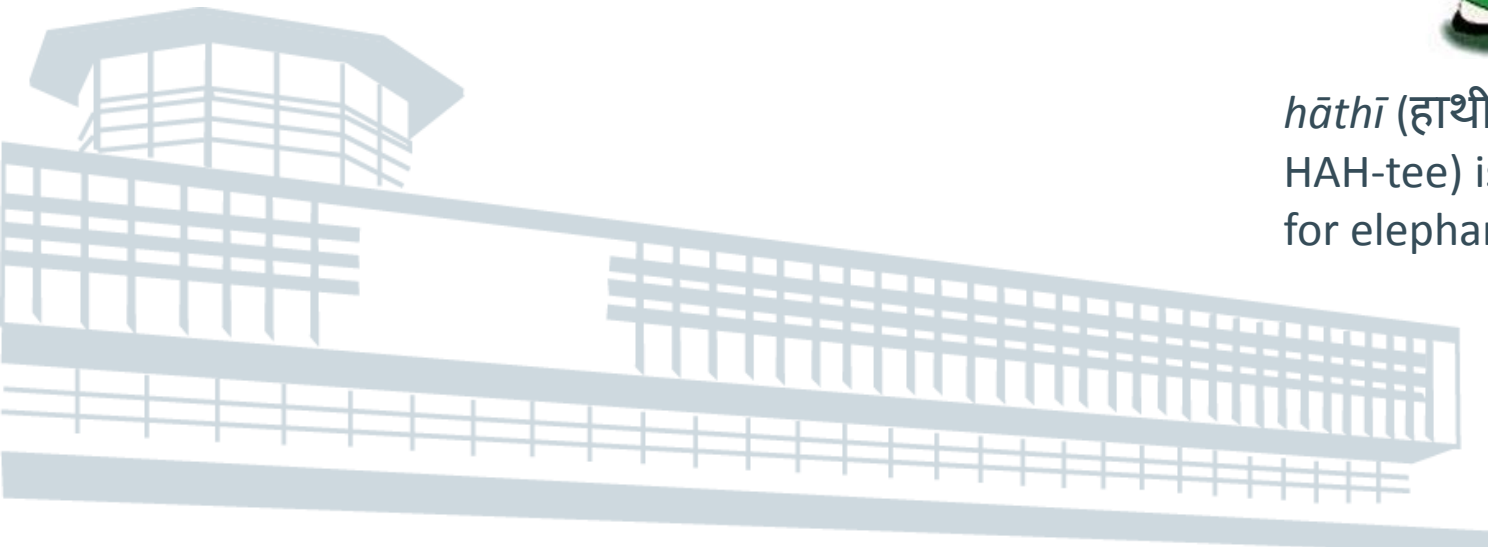
HATHI
TRUST

Overview

- Began in 2008
- Over 10.2 million volumes
- Over 2.9 million public domain (PD) volumes (“full view”)
- Over 60 partners



hāthī (हाथी) (pronounced HAH-tee) is the Hindi word for elephant.

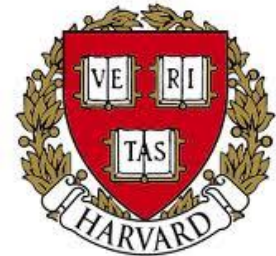


Content

HathiTrust Partners:



HATHI
TRUST



Content by call number, language, date

http://www.hathitrust.org/statistics_visualizations

And more...



US Gov Docs in HathiTrust

- Ca. 300,000 = 4% of all titles in HathiTrust
- 80% of gov docs in HathiTrust in public domain



Percentage of Docs in HathiTrust (est.) 1895-2009

Year	Hathi Docs	Docs Total	% Hathi Docs
2009	724	12198	5.9%
2008	1639	11609	14.1%
2007	1546	9389	16.5%
2006	1732	9196	18.8%
1999	2019	14093	14.3%
1998	2252	15846	14.2%
1997	2455	17071	14.4%
1996	2330	16361	14.2%
1995	2727	19107	14.3%
1994	2833	20030	14.1%
1993	2836	20367	13.9%
1992	3206	21056	15.2%
1985	3075	19942	15.4%
1984	3222	20878	15.4%
1983	3009	17550	17.1%
1982	2718	18828	14.4%
1981	3133	18143	17.3%
1980	4552	16443	27.7%
1979	4212	19415	21.7%
1978	4090	20624	19.8%
1977	4090	18439	22.2%
1976	3424	16809	20.4%
1973	2104	16774	12.5%
1972	917	17333	5.3%
1971	824	16986	4.8%
1970	970	18142	5.3%
1969	847		
1968	799		
1967	837		
1966	771		
1965	841		
1964	963		
1963	525		
1962	591		
1961	984		
1960	452		
1959	459		
1958	448		
1957	502		
1956	102		
1955	428		
1954	405		
1953	394		
1952	413		
1951	285		
1950	406		
1949	389		
1948	486		
1947	402		
1946	412		
1945	350		
1944	383		
1943	336		
1942	305		
1941	446		
1940	198		
1939	363		
1938	453		
1937	157		
1936	193		
1935	266		
1934	261		
1933	183	14254	1.3%
1932	262	15200	1.7%
1931	283	14214	2.0%
1930	189	14209	1.3%
1915	0	9127	0.0%
1914	292	10270	2.8%
1913	439	9258	4.7%
1912	428	11033	3.9%
1911	136	8855	1.5%
1910	135	10654	1.3%
1909	0	8167	0.0%
1908	352	6174	5.7%
1907	467	6745	6.9%
1906	0	7485	0.0%
1905	111	7341	1.5%
1904	221	9566	2.3%
1903	432	5890	7.3%
1902	214	9784	2.2%
1901	0	6879	0.0%
1900	310	9464	3.3%
1899	0	5475	0.0%
1898	92	7399	1.2%
1897	0	5793	0.0%
1896	87	7907	1.1%
1895	85	2643	3.2%

Brown (2011)

HathiTrust Compared to Google Books

- More titles found in Google but HathiTrust provides more full-text

Total docs = 385 (1940s)	Titles Found	Full-text
HathiTrust	98	90
Google	181	4

- HathiTrust better for searching serials
 - One record for all issues of a title
 - Title changes noted

Sare (2012)

Who can do what



Everyone

- View PD content
- Search PD and copyright materials
- View public collections
- Download *single PD pages*
- Download MARC records

HathiTrust Partners

- Download *entire volumes of PD materials*
- Create private or public collections
- Have a voice in the future of HathiTrust



Advanced Search in HathiTrust

Advanced Catalog Search:

AND ▾

AND ▾

AND ▾

- All Fields
- All Fields
- Title
- Author
- Subject
- Publisher
- Series Title
- Year of Publication
- ISBN/ISSN

Limit To:

Full view only

Year of publication:

During or after ▾

Language:

Original Format:

- All
- Abkhazian
- Achinese
- Acoli
- Adangme
- Adyghe
- Afrikaans
- Afro-Asiatic
- Ainu
- Akan

- All
- Archive
- Audio
- Audio (music)
- Audio (spoken word)
- Audio CD
- Audio LP
- Biography
- Book
- CDROM

Searching & Discovering Content

HT Catalog <http://www.hathitrust.org/>

HT WorldCat Local prototype
<http://hathitrust.worldcat.org/>

Resource Discovery Tools



encore



ExLibris Primo



Searching & Discovering Content

Loading records into local ILS

<http://www.hathitrust.org/data>

Bibliographic and Data APIs

<http://www.hathitrust.org/data>

Widgets

<http://www.hathitrust.org/widgets>





SEARCH RESOURCES

- Catalogs
- eBooks**
- eJournals
- Articles
- Databases
- Google Scholar

- Discovery Catalog eBooks
- HathiTrust
- Google Books

Full view only



[HathiTrust](#)
[What is this searching?](#)

HOW DO I ... ?



- [Get Access from Off Campus](#)
- [Reserve a Group Study Room](#)
- [Find CDs, DVDs, etc.](#)
- [Use Printers and Computers](#)
- [More...](#)

RESEARCH HELP



- [Subject Librarians](#)
- [Research Guides](#)
- [Reference Sources](#)
- [Citation Tools](#)
- [Special Collections](#)
- [Digital Gateway](#)
- [Open Scholarship](#)
- [Copyright and Publishing](#)
- [More...](#)

ABOUT THE LIBRARIES



- [Maps and Directions](#)
- [Services for ...](#)
- [Support the Libraries](#)
- [Library Staff](#)
- [More...](#)

WED, MAY 2, 2012

- Olin (main)
- Monday: 7:30 am - 4:00 am
 - Tuesday: 7:30 am - 4:00 am
 - Wednesday: 7:30 am - 4:00 am
 - Thursday: 7:30 am - 4:00 am
 - Friday: 7:30 am - 4:00 am
 - Saturday: 9:00 am - 4:00 am
 - Sunday: 10:00 am - 4:00 am
- [All hours](#) | [Olin exceptions](#)

NEWS & EVENTS



- [New Book Drop Adds Convenience](#)
- [Latest Issue of Off The Shelf Arrives](#)
- [Need an Article? Reach Out to ILLiad](#)
- [Upcoming Workshops](#)

LIBRARIES: [Olin \(Main\)](#) | [Art & Architecture](#) | [Business](#) | [Chemistry](#) | [Digital Gateway](#) | [Earth & Planetary Sciences](#) | [East Asian](#) | [Law](#) | [Medical](#) | [Music](#) | [Physics](#) | [Social Work](#) | [Special Collections & Archives](#) | [West Campus](#)



Searching & Discovering Content

Embed links

- Public collections
 - <http://babel.hathitrust.org/cgi/mb>
- Individual items

From a LibGuide for a
German literature
course

SELECTED LITERARY JOURNALS

- Deutsche Rundschau (Full text, issues 1879-1920)
From HathiTrust
- Deutsche Rundschau (Print, 1874-1964)
At WU
- Frauentaschenbuch für das Jahr (Fulltext, 1817-1824)
From HathiTrust
- Nord und Süd : eine deutsche Monatsschrift
From HathiTrust
- Vom Fels zum Meer (full text, ~1880s)
From HathiTrust
- 19th century German literary periodicals (Full Text at HathiTrust)
Selected list. Login with WUSTL key to download and/or print. Click on "catalog record" to view all issues available.

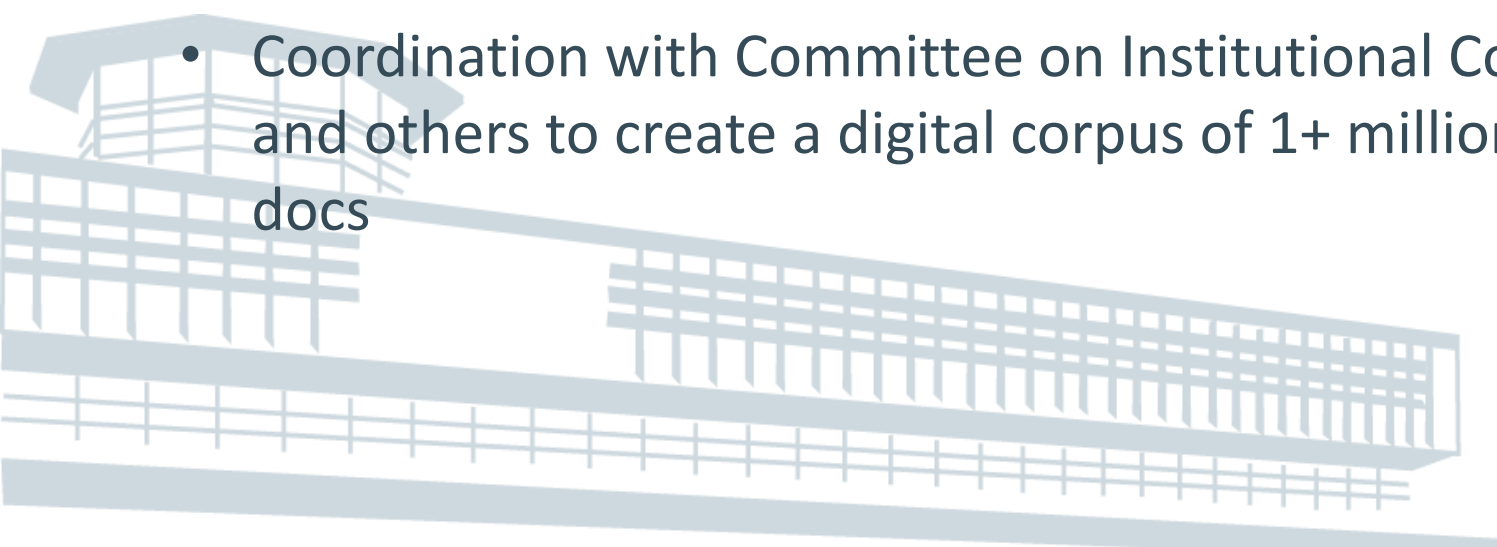
Final Thoughts

CHALLENGES

- Searching/retrieval obstacles (e.g. no SUDoc search)
- Inaccurate copyright statuses impeding access
- Inaccurate linkages and bibliographic info

PROGRESS

- Commitment to expand and enhance access to gov docs
- Research study examining how to improve access
- Coordination with Committee on Institutional Cooperation and others to create a digital corpus of 1+ million print docs



Questions about HathiTrust

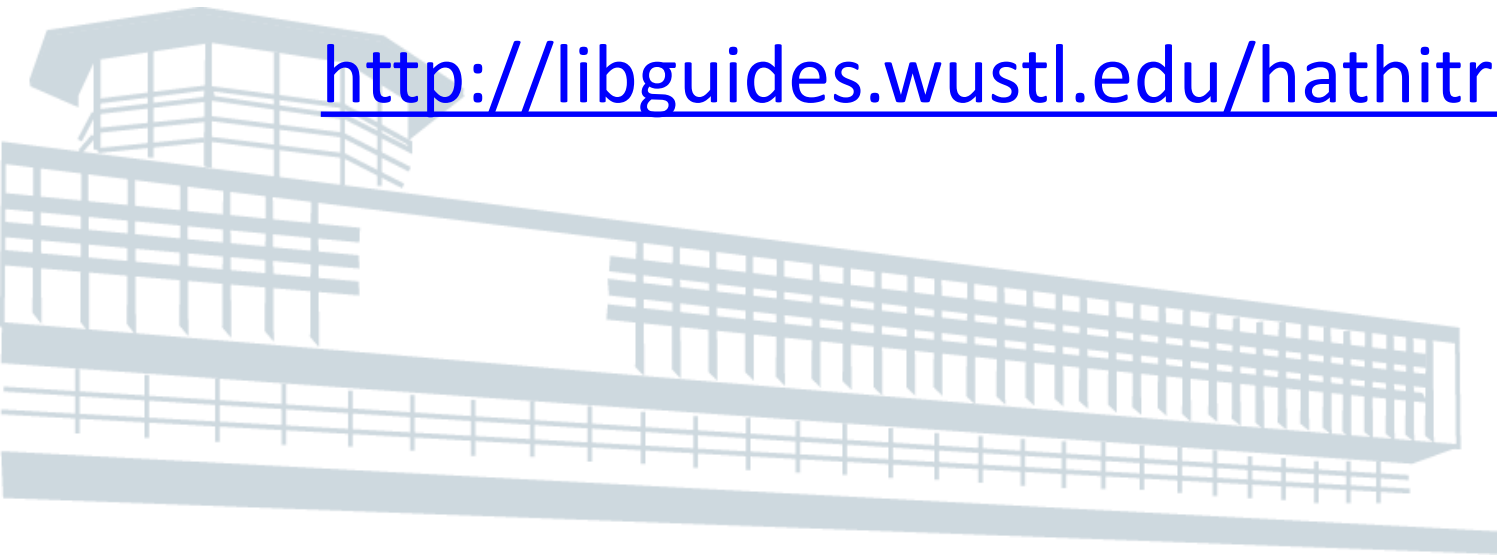
<http://www.hathitrust.org/help>

feedback@issues.hathitrust.org

@hathitrust

More info:

<http://libguides.wustl.edu/hathitrust>



More info on loading records into ILS

Kent State Univ.:

<http://techserv.lib.muohio.edu/ovgtsl11/presentations/Panchyshyn.pptx>

Univ. of Denver:

<http://www.slideserve.com/holleb/harvesting-hathitrust-documents-a-new-model-for-online-access>

Univ. of Colorado-Denver:

Beall, Jeffrey. 2009. "Free Books: Loading Brief MARC Records for Open-Access Books in an Academic Library Catalog." *Cataloging & Classification Quarterly* 47 (5) (January 4): 452–463. doi:10.1080/01639370902870215.



Bibliography

Malpas, Constance. 2011. *Cloud-sourcing Research Collections Managing Print in the Mass-digitized Library Environment*. Dublin, Ohio : OCLC Research. Accessed May 2, 2012
<http://www.oclc.org/research/publications/library/2011/2011-01.pdf>

York, Jeremy. 2012. HathiTrust: Issues and Challenges in Preserving the Published Record [PowerPoint slides]. Accessed April 30, 2012
<http://www.hathitrust.org/documents/HathiTrust-Amigos-201202.pptx>

Brown, Christopher C. 2011. Harvesting HathiTrust Documents: A New Model for Online Access [PowerPoint slides]. Accessed April 30, 2012
<http://www.slideserve.com/holleb/harvesting-hathitrust-documents-a-new-model-for-online-access>

York, Jeremy. 2012. "HathiTrust: The Elephant in the Library." *Library Issues: Briefings for Faculty and Administrators* 32 (3) (January). Accessed May 2, 2012 <http://www.libraryissues.com/sub/LI320003.asp> .

Sare, Laura. 2012. "A Comparison of HathiTrust and Google Books Using Federal Publications." *Practical Academic Librarianship: The International Journal of the SLA Academic Division* 2 (1): 1–25. Accessed May 2, 2012
<http://journals.tdl.org/pal/article/viewFile/5880/5922>

Thank you!

bvetruba@wustl.edu

@bvetruba

