

KILPATRICK TOWNSEND & STOCKTON LLP

Joseph Petersen (JP 9071)
 Robert Potter (RP 5757)
 1114 Avenue of the Americas
 New York, NY 10036
 Telephone: (212) 775-8700
 Facsimile: (212) 775-8800
 Email: jpetersen@kilpatricktownsend.com

Joseph M. Beck (admitted *pro hac vice*)
 W. Andrew Pequignot (admitted *pro hac vice*)
 Allison Scott Roach (admitted *pro hac vice*)
 1100 Peachtree Street, Suite 2800
 Atlanta, Georgia 30309-4530
 Telephone: (404) 815-6500
 Facsimile: (404) 815-6555
 Email: jbeck@kilpatricktownsend.com

Attorneys for Defendants

**UNITED STATES DISTRICT COURT
 SOUTHERN DISTRICT OF NEW YORK**

THE AUTHORS GUILD, INC., ET AL.,

Plaintiffs,

v.

HATHITRUST, ET AL.,

Defendants.

Case No. 11 Civ. 6351 (HB)

**DECLARATION OF JOHN WILKIN IN SUPPORT OF
 DEFENDANTS' MOTION FOR SUMMARY JUDGMENT**

I, John Wilkin, pursuant to 28 U.S.C. § 1746, hereby declare as follows:

1. I am an Associate University Librarian at the University of Michigan (“Michigan”) and also serve as the Executive Director of HathiTrust, which, as is explained in more detail below, is the name of a service provided by Michigan. I submit this declaration in support of the defendant libraries’ (the “Libraries”) motion for summary judgment. Unless

otherwise noted, I make this declaration based upon my own personal knowledge or, where specifically noted, based upon Michigan's business records.

2. As Associate University Librarian for Library Information Technology, I am responsible for, among other things, the online catalog and related technologies for the University Library which is physically spread over numerous buildings and individual libraries (collectively, the "University Library").¹

3. My duties include ensuring the University Library has the necessary technological infrastructure and networked systems to support the library's core mission and services. I have served in this role since June 1, 2002.

4. On June 1, 2012, I assumed responsibility for many of Michigan's publishing activities, including the University of Michigan Press and digital publishing operations.

5. Prior to my current role and responsibilities, I served as the Head of the Digital Library Production Service at Michigan, a position that I held since 1996. In that role I was responsible for campus- and internet-wide Michigan-hosted digital collection services.

6. I have continuously (with the exception noted below) served in various roles in Michigan's library since graduating with a Masters in Library Science from the University of Tennessee in 1986 (with the exception of 1992 through 1994 during which time I served as Systems Librarian for Information Services at the University of Virginia).

7. I have served as the Executive Director of HathiTrust since 2008. In that role, I am responsible for the service's operations, development, budget and the measures taken to ensure the security of the works within the HathiTrust digital library.

¹ As used in this declaration, and unless otherwise noted, the term "University Library" does not include certain other libraries at Michigan including Bentley Historical Library, Clements Library, Kresge Business Administration Library, Law Library, Thompson Library (Flint) and Mardigian Library (Dearborn), among others.

A. The University Library

8. Consistently ranked as one of the top ten academic research libraries in North America, the University Library, which, as a part of a non-profit organization dedicated to learning, is open to students, faculty and the general public, makes available an extraordinary array of resources and services.

9. The University Library holds more than 8 million bound volumes housed in a number of physical locations across the Michigan campus. All of the various libraries at Michigan, including the law and business school libraries, hold more than 11 million volumes.

10. Last year, in fiscal year 2011, the University Library hosted nearly 4 million patron visits.

B. The Core Function of Academic Libraries Such as the University Library

11. In order to place HathiTrust in context, it helps to have some background regarding the basic functions of the University Library, indeed all academic libraries:

- We buy works for academic and scholarly pursuits;
- We curate, maintain, and preserve those works;
- We help scholars and students identify works pertinent to their pursuits; and
- We make those works available and accessible consistent with applicable law.

12. We have been performing these functions for nearly 175 years.

(i) Acquisition of Works

13. Academic libraries such as the University Library acquire works to satisfy anticipated future demand by University Library patrons. When a work is requested by many patrons, and we find ourselves maintaining a waiting list for that work, we will often try to purchase additional copies.

14. Last year, in fiscal year 2011, Michigan's libraries spent more than 24 million dollars on library materials and the vast majority of these sums were spent acquiring new works. Although state appropriations for the university consistently decrease, the University Library's spending for acquisitions continually increases.

15. We spend millions of dollars each year on obtaining access to electronic resources: we spend approximately 12 million dollars² per year in order to acquire the right to display the full text of works (most of which are in-copyright) to library patrons.

16. While I discuss this point in greater detail below, it bears emphasis that our digitization efforts, including those associated with the HathiTrust Digital Library ("HDL"), do not diminish our acquisitions of in-copyright material (digital or otherwise). Each year, we spend millions of dollars to license the right to display the text of copyrighted works and to acquire new books. Moreover, *no portion of in-copyright materials is displayed to patrons through the HDL* (except to students, faculty and staff with certified print disabilities—please see Section J, ¶¶ 100-106, below for a description of this service). In other words, the HDL is not a substitute, in any respect, for our acquisitions of in-copyright material (whether print or digital).

17. There are a number of reasons why academic libraries spend such enormous sums on acquisitions *every year*. Academic libraries acquire works not just for current students and faculty, but also for future generations.

18. Librarians cannot reliably predict which works may be of scholarly interest in ten, fifty, or one hundred years. This is one reason why the University Library acquires an extraordinarily broad range of materials on every conceivable subject.

² This figure includes expenditures by all campus libraries because many such licenses are jointly negotiated or funded for the campus as a whole.

19. Indeed, it is not unusual to hear scholars express pleasant surprise (and relief) to discover that we have a particular work in our collection. Such statements are a testament to our efforts to acquire and preserve a breathtaking number of works. We do so because of the mere *possibility* that a particular work on a seemingly obscure topic may be valuable to a future student or scholar.

20. The imperative that academic and research libraries acquire a broad range of material for future scholarship is magnified by the fact that the library can have no assurance that a work will remain available after it is first published. Indeed, most works go out of print after the initial print run and once that print run is sold out, it can be difficult if not impossible to obtain copies of the work.

21. As a result, academic libraries typically acquire works very shortly after they are published—even before a definite scholarly need has surfaced—and they need to purchase a sufficient number of copies of each work to accommodate anticipated user demands; otherwise, the library may not be able to buy the work later. This is particularly true for books published outside of the United States, for example in developing countries, and most journal issues are out of print soon after the initial issue is distributed.

(ii) Preservation

22. Books, in their physical form, are inherently subject to damage, deterioration and loss. This is particularly true for books published between 1850 and 1990—approximately three-quarters of our entire collection—because books published during this time period were generally published on paper with high acid content.

23. Paper with high acid content degrades far more quickly than paper with low acid content. This is because the fibers that comprise paper degrade when acid meets the moisture in

the air. In what is referred to as an “acid hydrolysis reaction,” the paper fibers are repeatedly split into smaller fragments so long as the source of acid remains in the paper. This process, in fact, produces more acid and the degradation accelerates in a downward spiral.

24. As a result, books that are more than 160 years old—that typically were published on rag cotton paper, which is relatively more durable—are usually in far better condition than books that are less than 50 years old.

25. As of 2004, Michigan estimated that about half of its collection—approximately 3.5 million books—was printed on paper with high acid content, *i.e.* on paper that is particularly vulnerable to deterioration and, ultimately, loss.

26. Prior to the digitization project at issue in this proceeding, *the process of searching the University Library’s immense collections to identify deteriorating books took so long that, by the time we identified the most imperiled books from the millions potentially at risk, it was often too late and the books had disintegrated or were unusable.*

27. Our earliest, independent efforts to preserve deteriorating books through digitization were also severely limited by the length of time it took us to digitize them. Indeed, books were deteriorating so rapidly that, even if we could have instantly identified all of the books in our vast collections that were on the brink of deterioration (as noted above, this is an impossible task), we still could not have digitized the collection in time to preserve the content of the works.

28. Indeed, the University Library was the industry leader in the average number works digitized per year. However, we were only capable of independently digitizing approximately 5,000 books per year, which was but a small fraction of the imperiled works within our collection.

29. In other words, prior to asking Google to digitize our collections, we were losing the race to save significant portions of the library's works.

30. Gradual disintegration is not the only threat to books in the academic libraries. Loss from theft, vandalism, fire, and floods presents an ever-looming threat.

31. Hurricane Katrina devastated Tulane University's Howard-Tilton Memorial Library. I understand from published reports that its basement floor (approximately the size of a football field) flooded with over eight feet of water, destroying 90% of the 500,000 volumes in one of the library's collections.

32. Library destruction has not been limited to acts of nature. The most famous example of a loss of a library is probably the destruction of the Library of Alexandria, but millions of volumes have been destroyed in libraries during the World Wars, and the collection of the National Library in Sarajevo lost over 1 million volumes due to shelling in the 1990's.

(iii) Helping Scholars Identify Works of Potential Interest

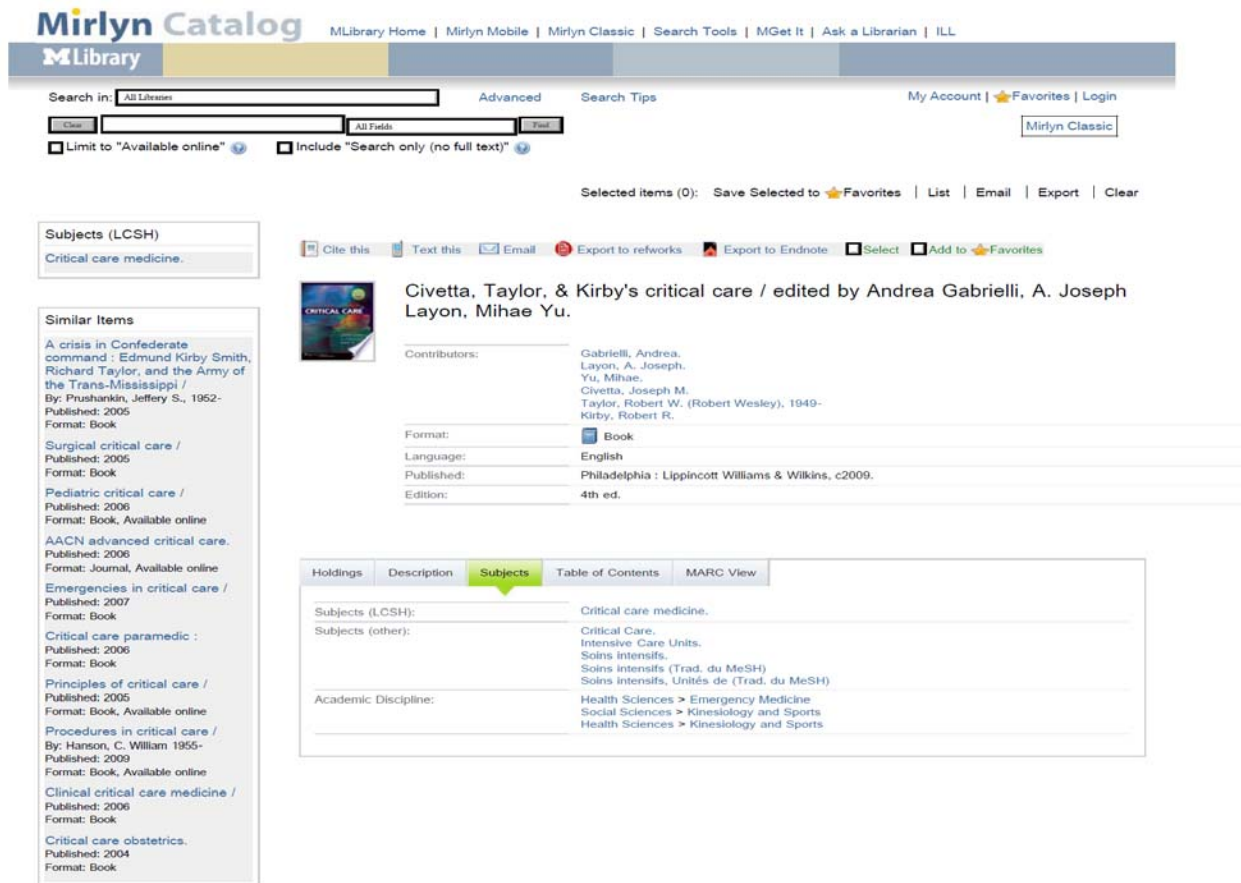
33. Importantly, libraries aid scholars in the identification of relevant works. The immense collections housed by academic libraries such as the University Library would be significantly diminished without reliable and efficient search methods and related technology.

34. Until relatively recently, most searches of a library's collection relied on a physical card catalog. Each card contained limited information concerning a particular work, including its title, author, publication date and publisher and limited information concerning the work's subject matter.

35. In order to automate cataloging, libraries began to share the work of creating bibliographic description in the 1960's. As part of these efforts, libraries created the Research Libraries Group and the Online Computer Library Center (OCLC), a non-profit organization that

developed and maintained Worldcat, the largest online public access catalog (OPAC) in the world. This paved the way for the creation of online catalogs in the 1970's.

36. While converting the physical card catalog to a digital one empowered users to perform broader searches, those searches were still limited to *the work's basic bibliographic data*, such as author, title, subject, etc., as illustrated in the following screen shot from Michigan's online catalog:



37. Even with the advent of the online catalog such as depicted above, the actual content of the works remained closed to searches. Accordingly, a work that contained information of great importance to a researcher would not be discoverable by that researcher unless the work's title, subject headings, or other limited bibliographic data happened to contain certain key words or other evidently pertinent information.

(iv) **Making Works Available Pursuant to Our Understanding of**

Applicable Law

38. One of the most basic functions of the University Library—indeed all academic and research libraries—is lending books and other materials to patrons. Further, the University Library makes works available in a variety of other ways:

- We make copies of works in order to provide equal access to those works to students, faculty, and staff who have certified print disabilities;
- We reproduce and distribute works that are in the public domain; and
- We reproduce and distribute works pursuant to Section 108 of the Copyright Act in the event that a work is lost, damaged, deteriorating, or stolen and a replacement copy cannot be obtained by the University Library at a fair price.

C. The University Library’s Early Digitization Efforts

39. Starting in the mid 1980’s, the University Library, like many leading academic and research libraries, began investing in the equipment necessary to convert works from print to digital format.

40. We took this significant step because we recognized that, in the decades to follow, basic library functions would increasingly require computing technology.

41. For example, as summarized above, one of the most critical missions served by libraries such as the University Library is the preservation of works for future generations. It was for this reason that in the late 1980’s we began converting at-risk materials to digital format. We knew that by digitizing such works we were ensuring that they would be available for future scholarly pursuits even in the event that the work in physical form was lost and we could not find a replacement copy at a fair price.

42. As noted above (¶¶ 22-29), while we sought to preserve at risk works through digitization, we found that given the enormous size of our collections we could not digitize and,

thereby, preserve deteriorating works quickly enough. In fact, during this time period I understand that we lost irreplaceable volumes which, as a result, have vanished from the academic and cultural landscape.

43. The University Library's early efforts at digitization also allowed for an increased, though still very limited, number of works to be made more readily accessible to those with certified print disabilities, and allowed for some improved search functionality across the digitized works. A truly comprehensive solution, however, required large-scale digitization that the University Library could not possibly accomplish on its own.

D. Google's Involvement in Michigan's Digitization Efforts

44. Prior to Google's involvement in our digitization efforts, at our then rate of scanning, it would have taken us more than 1,000 years to digitize the University Library's then over 7 million volumes.

45. In 2002, we began speaking with Google about its interest in digitizing Michigan's entire library collections in less than a decade.

46. In late 2004, we entered into an agreement with Google under which Google would convert hardcopy books from Michigan Library collections to a digital format and provide digital copies of those books to Michigan. Attached as **Exhibit A** is a true and correct copy of Michigan's current agreement with Google concerning this project.

47. In return for giving Google access to books in the University Library collection, Google was required to give the University Library a digital copy of the works digitized by Google. We bargained for this right because it was important to us that we had the right to control our own uses and satisfy one of our primary missions of providing specialized services to the blind or other persons with disabilities. We knew that by maintaining control over our own

digitized copies of our collection we could ensure that students and faculty with print disabilities had access to works within the HDL on par with their non-disabled peers.

48. Our aim in working with Google was to digitize as much of the University Library as possible.³ If we digitized only select portions of our collections we would not have accomplished our goals.

49. For example, if we limited the project solely to works known to be in the public domain, we would have continued to lose books presumed to be in-copyright to inevitable disintegration and decay and, potentially, to theft, vandalism and natural disaster or calamity.

50. Further, digitization held the promise of providing students and scholars with print disabilities immediate access to our print collections on par with the access afforded other library patrons. That promise would have been largely unrealized if we had limited digitization to the public domain.

51. Finally, *from the very outset of the project our goal was to offer scholars a better, more comprehensive way to search for and discover pertinent works within the collection.* If we only allowed such searches over the portion of works known to be in the public domain, roughly 75% of the library's collections would have been excluded. A search tool that excluded 75% of our collections would be of significantly less value to students and scholars seeking to identify the works most relevant to them.

52. While Michigan was the first academic library to work with Google in connection with what would become the "Google Book Project," it is my understanding that Google ultimately partnered with a number of other universities and research libraries. For example, I am aware that in addition to the defendants named in this lawsuit, Google worked with such

³ In certain instances, rights holders availed themselves of an opt-out program offered by Google in connection with its digitization of works. In those situations, our digital collection does not include such works.

universities as Harvard University, Stanford University, Oxford University, Columbia University, Princeton University, the University of Virginia, and the University of Texas at Austin, among others. The benefits to society—in preserving books, making them accessible to people with print disabilities, and enabling people to find them—increased significantly with each institution that digitized books from its collections.

E. The Formation of HathiTrust

53. In 2008, Michigan formed HathiTrust, named for the Hindi word for elephant, “hathi.” The “hathi” prefix is intended to evoke the qualities of memory, wisdom, and strength symbolized by elephants.

54. The concept underlying the formation of HathiTrust was (and is) simple. It makes no economic or functional sense for each research library to maintain its own digitized collection of works. Rather, we believe that by working together and pooling resources we can better serve our common goals of collecting, organizing, securing, preserving and, consistent with applicable law, sharing the record of human knowledge.

55. Accordingly, pursuant to the HathiTrust mission, participating members combined their digitized collections in order to provide more secure, long-term storage for the works, more comprehensive research and discovery tools, improved access to works in the public domain and improved access to works for students and faculty with print disabilities. Michigan runs the HDL as a service not only for the benefit of Michigan but also for the benefit of all participating institutions and, indeed, all users of the HathiTrust website located at www.hathitrust.org.

56. There are currently more than sixty institutions participating in HathiTrust, including Michigan, and membership is open to institutions worldwide. Attached hereto as **Exhibit B** is a true and correct copy of participating members as of today.

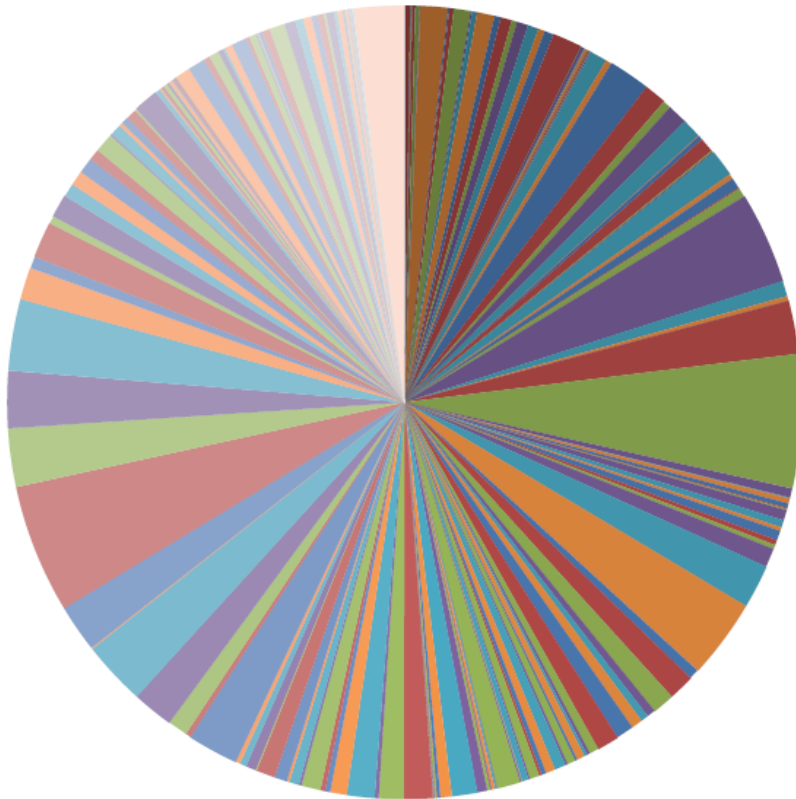
F. The Composition of the HathiTrust Digital Library (“HDL”)

57. The combined corpus of the HDL now totals more than 10 million works and is growing every day.

58. While the HDL corpus contains a very large number of works, we have a significant amount of information regarding the general composition of the corpus.

59. For example, an analysis of the Library of Congress call numbers of works provides an overview of the subject matter of the works found in the HDL:

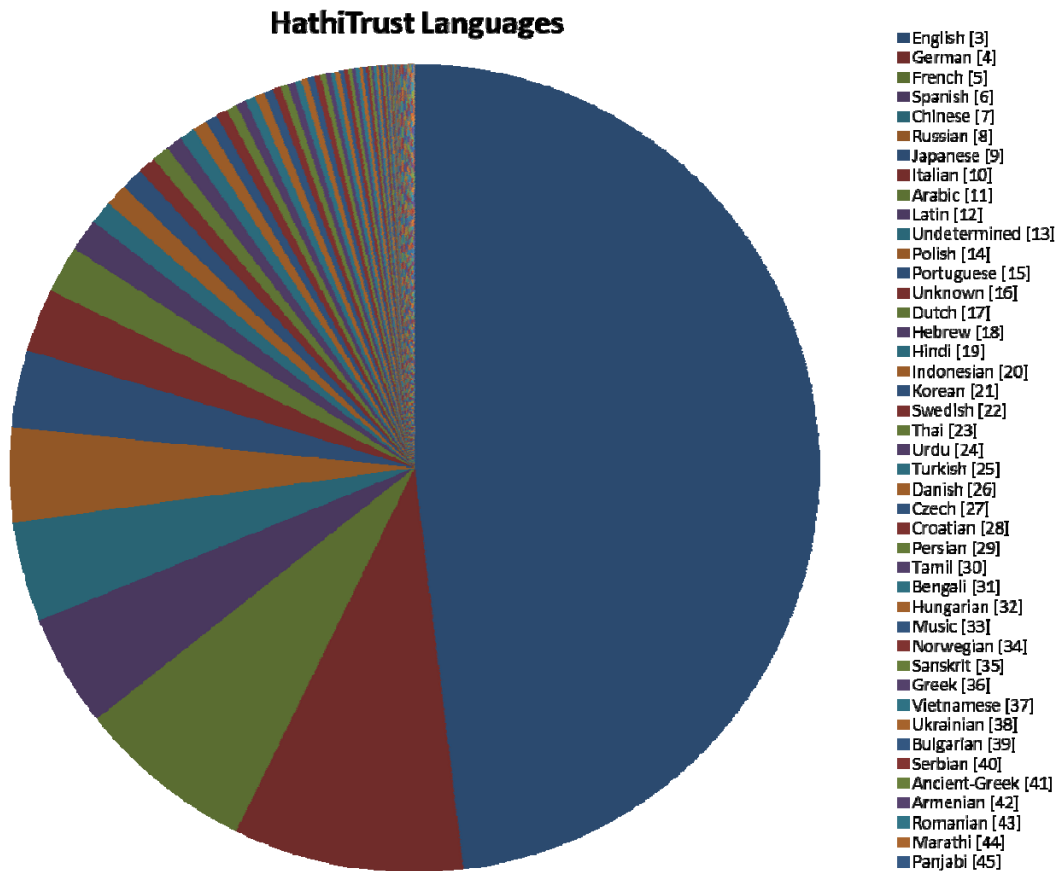
HathiTrust Call Numbers



Legend follows on next page.

- A - General Works [2]
- AG - Dictionaries...[5]
- AN - Newspapers [8]
- AY - Yearbooks. Almanacs.[11]
- BC - Logic [14]
- BH - Aesthetics [17]
- BM - Judaism [20]
- BR - Christianity [23]
- BV - Practical Theology [26]
- CB - History of Civilization [29]
- CE - Technical Chronology. Calendar [32]
- CR - Heraldry [35]
- D - History General [38]
- DC - History of France-Andorra-Monaco [41]
- DF - History of Greece [44]
- DJ - History of the Netherlands [47]
- DP - History of Spain - Portugal [50]
- DS - History of Asia [53]
- DX - History of Gypsies [56]
- G - Geography; Anthropology; Recreation [59]
- GC - Oceanography [62]
- GN - Anthropology [65]
- GV - Recreation; Leisure [68]
- HB - Economic theory; Demography [71]
- HE - Transportation and communications [74]
- HJ - Public Finance [77]
- HQ - Family; Marriage; Woman; Sexuality [80]
- J - General legislative & executive papers [83]
- JF - Political instit. & public admin - Gen'l [86]
- JL - Political instit. & pub. admin-Canada; etc.[89]
- JS - Local government. Municipal govt [92]
- JZ - International relations [95]
- KDK - Law (United Kingdom and Ireland [98]
- KG - Law (Latin America-Mex. & C. Amer....)[101]
- KL - Law (History of law-The ancient orient)[104]
- LA - History of Education [107]
- LD - Education (Individual Institutions-US) [110]
- LG - Education (Indiv. Instit.-Asia; Africa; etc.[113]
- LT - Textbooks [116]
- MT - Musical instruction and study [119]
- NB - Sculpture [122]
- NE - Print media [125]
- P - Language and Literature [128]
- PC - Romanic languages [131]
- PG - Slavic langs; Baltic langs; Albanian lang[134]
- PK - Indo-Iranian philology and literature [137]
- PN - Literature (General) [140]
- PS - American literature [143]
- Q - Science (General) [146]
- QC - Physics [149]
- QH - Natural history - Biology [152]
- QM - Human anatomy [155]
- R - Medicine (General) [158]
- RC - Internal medicine [161]
- RF - Otorhinolaryngology [164]
- RK - Dentistry [167]
- RS - Pharmacy and materia medica [170]
- RX - Homeopathy [173]
- SB - Plant culture [176]
- SH - Aquaculture. Fisheries. Angling. [179]
- TA - Engineering (Gen'l). Civil Engineering [182]
- TE - Highway engineering. Roads & pavements [185]
- TH - Building construction [188]
- TL - Motor vehicles Aeronautics Astronautics[191]
- TR - Photography [194]
- TX - Home economics [197]
- UB - Military administration [200]
- UE - Cavalry. Armor [203]
- UH - Other services [206]
- VB - Naval administration [209]
- VE - Marines [212]
- VK - Navigation. Merchant marine [215]
- ZA - Information Resources [218]
- AC - Collections. Series. [3]
- AI - Indexes [6]
- AP - Periodicals [9]
- AZ - History of scholarship[12]
- BD - Speculative philosophy [15]
- BJ - Ethics [18]
- BP - Islam. Bahaism. Theosophy [21]
- BS - The Bible [24]
- BX - Christian Denominations [27]
- CC - Archaeology [30]
- CJ - Numismatics [33]
- CS - Genealogy [36]
- DA - History-G. Britain-Ireland-British Isles [39]
- DD - History of Germany [42]
- DG - History of Italy - Malta [45]
- DK - History of Russia; Soviet Union... [48]
- DQ - History of Switzerland [51]
- DT - History of Africa [54]
- E - History of America [57]
- GA - Mathematical geography; Cartography[60]
- GE - Environmental Sciences [63]
- GR - Folklore [66]
- H - Social sciences (General) [69]
- HC - Economic history and conditions [72]
- HE - Commerce [75]
- HM - Sociology (General) [78]
- HV - Social pathology; Soc'l & pub. welfare...[81]
- JA - Political science (General) [84]
- JJ - Political instit & public admin-N America[87]
- JN - Political instit. & public admin -Europe [90]
- JV - Colonies & colonization. Emigration... [93]
- K - Law [96]
- KE - Law (Canada) [99]
- KH - Law (South America) [102]
- KZ - Law of Nations. International law [105]
- LB - Theory and practice of education [108]
- LE - Education (Indiv. Instit.-Amer. except US)[111]
- LH - College and school magazines and papers [114]
- M - Printed Music [117]
- N - Visual arts [120]
- NC - Drawing. Design. Illustration [123]
- NK - Decorative arts [126]
- PA - Greek and Latin language & lit. [129]
- PE - English language [132]
- PH - Uralic languages; Basque language [135]
- PL - Languages of Eastern Asia; Africa; etc.[138]
- PQ - French lit.-Italian, Spanish, Portuguese [141]
- PT - German Dutch Scand. etc lit since 1830[144]
- QA - Mathematics [147]
- QD - Chemistry [150]
- QK - Botany [153]
- QP - Physiology [156]
- RA - Public aspects of medicine [159]
- RD - Surgery [162]
- RG - Gynecology and obstetrics [165]
- RL - Dermatology [168]
- RT - Nursing [171]
- RZ - Other systems of medicine [174]
- SD - Forestry [177]
- SK - Hunting sports [180]
- TC - Hydraulic engineering - Ocean Eng'g [183]
- TF - Railroad engineering and operation [186]
- TJ - Mechanical engineering & machinery [189]
- TN - Mining engineering. Metallurgy [192]
- TS - Manufactures [195]
- U - Military Science (General) [198]
- UC - Maintenance and transportation [201]
- UF - Artillery [204]
- V - Naval science (General) [207]
- VC - Naval maintenance [210]
- VF - Naval ordnance [213]
- VM - Naval architecture Shipbuilding etc. [216]
- AE - Encyclopedias [4]
- AM - Museums. Collectors [7]
- AS - Academies...[10]
- B - Philosophy; Psychology; Religion [13]
- BF - Psychology [16]
- BL - Religions. Mythology...[19]
- BQ - Buddhism [22]
- BT - Doctrinal Theology [25]
- C - Auxiliary Science of History [28]
- CD - Diplomats. Archives. Seals [31]
- CN - Inscriptions. Epigraphy [34]
- CT - Biography [37]
- DB - History of Austria-Liechtenstein...[40]
- DE - History of the Greco-Roman World [43]
- DH - History of Low Countries-Benelux C's[46]
- DL - History of Northern Europe; Scandavia [49]
- DR - History of the Balkan Peninsula [52]
- DU - History of Oceania (South Seas) [55]
- F - Local History-United States & British...[58]
- GB - Physical geography [61]
- GF - Human ecology; Anthropogeography [64]
- GT - Manners and customs (General) [67]
- HA - Statistics [70]
- HD - Industries; Land use; Labor [73]
- HG - Finance [76]
- HN - Social history & conditions... [79]
- HX - Socialism. Communism. Anarchism [82]
- JC - Political theory-State Theories of state[85]
- JK - Political instit. & public admin -US [88]
- JQ - Political instit. & public admin -Asia; Africa etc.[91]
- JX - International law (obsolete) [94]
- KB - Religious law [97]
- KF - Law (United States) [100]
- KJ - Law (Europe) [103]
- L - Education [106]
- LC - Special aspects of education [109]
- LF - Education (Indiv. Instit. - Europe) [112]
- LJ - Student fraternities; United States [115]
- ML - Literature on music [118]
- NA - Architecture [121]
- ND - Painting [124]
- NX - Arts in general [127]
- PB - Modern languages; Celtic languages [130]
- PF - West Germanic languages [133]
- PJ - Oriental philology and literature [136]
- PM - Hyperborean; Indian; & artificial lang's[139]
- PR - English literature [142]
- PZ - Fiction and juvenile belles lettres [145]
- QB - Astronomy [148]
- QE - Geology [151]
- QL - Zoology [154]
- QR - Microbiology [157]
- RB - Pathology [160]
- RE - Ophthalmology [163]
- RJ - Pediatrics [166]
- RM - Therapeutics. Pharmacology [169]
- RV - Botanic; Thomsonian; eclectic med. [172]
- S - Agriculture (General) [175]
- SF - Animal culture [178]
- T - Technology (General) [181]
- TD - Environmental tech. Sanitary eng'g [184]
- TG - Bridge engineering [187]
- TK - Electrical engineering. Electronics... [190]
- TP - Chemical technology [193]
- TT - Handicrafts. Arts and crafts [196]
- UA - Armies: Org'n; distrib; military... [199]
- UD - Infantry [202]
- UG - Military engineering [205]
- VA - Navies: Org'n; distribution; etc. [208]
- VD - Naval seaman [211]
- VG - Minor services of navies [214]
- Z - Books (Gen'l). Writing Paleography etc.[217]

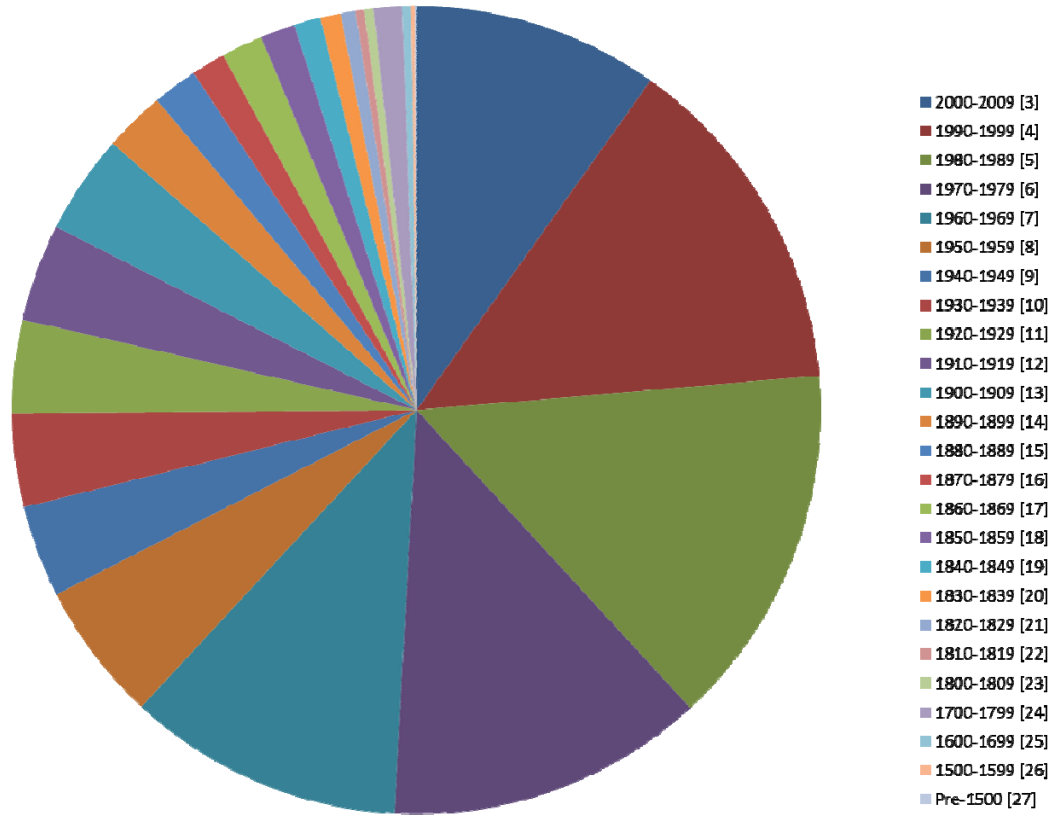
60. Further, the HDL contains works in a multitude of languages as illustrated in the following diagram:



61. Works within the HDL were also published across a broad range of dates, commencing prior to the 15th century and running to the present as illustrated in the following diagram:⁴

⁴ Interactive versions of each of the pie diagrams included in this declaration may be accessed through the HathiTrust website located at www.hathitrust.org.

HathiTrust Dates



62. Approximately 30% of the corpus consists of material that is clearly within the public domain. It bears emphasis that although we treat the balance of the works as if they are in-copyright, this does not in fact mean that 70% of the corpus is in-copyright.

63. It is notoriously difficult to determine whether a particular work remains in-copyright. For example, it is my understanding that works published between 1923 and 1963 entered the public domain unless they were renewed. Other copyright determinations may rely on the death date of authors about whom very little is known or documented.

64. While the vast majority of works from this period were not renewed, determining the renewal status of works from this period is an extraordinarily difficult task. The Copyright Office's records prior to January 1, 1978 are not completely or reliably digitized at the present time. Therefore, the process of confirming whether a work was renewed involves the laborious

task of checking the physical records at the Copyright Office in Washington, D.C. Of course, even if a search confirmed that the work had been renewed, there is no guarantee that a subsequent search would identify the current rights holder.

65. Accordingly, we err on the side of classifying works as in-copyright *even though we are confident that many of those works are, in fact, in the public domain.*

66. While precise calculation is difficult given the size of the HDL corpus, it is my understanding that the vast majority of works in the corpus are now out of print (and, in fact, for older works within the collection, have been out of print for decades).

67. Based upon an analysis of the call numbers within the archive, less than 9% of the HDL corpus consists of prose fiction, poetry and drama. The remainder, approximately 90% of the corpus, is likely to consist of factual works such as books and journals in many disciplines of the arts, humanities, social sciences and sciences.

G. The Limited Uses of the Works within the HDL

68. We permit extremely limited use of the works in the HDL presumed to be in-copyright. Specifically, patrons can *only* make the following uses of such works:

- **Full-Text Search:** Through the Internet, the University's students, faculty, and staff, as well as the general public, may search for a particular term across all works within the HDL. For those works that are not in the public domain or for which the copyright holder has not expressly authorized use, the search results indicate only the frequency and page numbers for which a particular term is found within a particular book or periodical. (Unlike Google's service, the results do not show portions of text in "snippet" format.) At no time does the user have digital access to any of the actual written content within such works (unless he/she is afforded access as a certified print disabled user). *In other words, none of the work's text is ever displayed on the computer screen or available for print, not even one word.*
- **Preservation:** As noted above, before Google assisted with our digital conversion, we were losing works every year as a result of the natural disintegration of books (particularly the large segment of the collection written on paper with high acid content). There was also the ever-present risk of other more sudden forms of loss such as those occasioned by fire, flood, or theft. The

HDL now constitutes an extremely valuable protection against the prospect of such loss and permits us to make copies pursuant to Section 108 of the Copyright Act in the event that a work is lost damaged, deteriorating, lost, or stolen and a replacement copy cannot be obtained by the University Library at a fair price.

- **Access for people with certified print disabilities:** For decades, the University Library has converted works in its collection to alternative formats for the blind and other persons who have disabilities that prevent them from accessing printed materials. Because digitization has significantly improved the quality of access for print-disabled readers, the HDL was designed specifically to enable libraries to make their collections accessible to such readers in digital format.

69. It is important to emphasize that given the very limited uses made of in-copyright works in the HDL, our digitization of such materials *does not diminish our purchases of in-copyright works*.

70. Indeed, in my opinion, if the HDL has any impact whatsoever on the University Library's acquisition of in-copyright material, *it has a positive effect on our purchasing*.

71. Experience and basic common sense tells me that scholars, students, and other patrons are *more likely to discover and use works that they can locate through digital search*. Such increased demand for works invariably translates into increased purchases. This is because, as noted above (*see ¶ 13*), if a work is frequently requested by patrons, and we find ourselves maintaining a waiting list for such works, we try to acquire more copies of that particular work to meet patron demand.

72. For instance, the University Library includes in its print collection a work called *Television Program Master Index*. This work contains an index of critical and historical information regarding over 1,000 television shows contained in hundreds of books.

73. We digitized *Television Program Master Index* and, since it is presumed to be in-copyright, we only permit HDL users to search the text of the work (*i.e.*, the text is not available except to users who have print disabilities). The HDL search functionality does quickly allow a

user to determine whether a particular show is covered in the work and, if it is, users typically follow up by borrowing a print copy of the work from the library.

74. After we made the *Television Program Master Index* searchable through the HDL, the usage of the University Library's copy of the work increased dramatically and we decided to acquire two additional copies of the work to satisfy the increasing demand.


H. The Benefits of the HDL's Full-text Search Functionality

75. Full-text searching easily constitutes the most significant advance in library search technology since the 1960's.

76. Rather than combing through electronic cataloging records and attempting to discern which works in our collection may be of interest, scholars can access the HDL website and search *the actual text* of over 10 million books and journals. They can then immediately access those works that are in the public domain or for which HathiTrust has the rights to display the work in full text mode.

77. The Libraries, through the HDL, have made it possible for university students, faculty, and staff, as well as the general public, to search the combined digital collections contributed by the HathiTrust members. The search results display bibliographic information—including title, author, publisher, and publication date—for books containing the search term, as well as the frequency and page numbers for which the term is found, giving some clues as to how useful the book might be.

78. For example, as of June 8, 2012, a search for “anaphylactic shock” identifies 38,239 works. If the user selects the work titled *Allergy and Tissue Metabolism* by W.G. Smith – in which the term “anaphylactic shock” appears – the following bibliographic information is displayed:


Help Feedback

Home About Collections My Collections

Catalog Search

Catalog Search

All Fields
 Full view only Full
[Advanced Catalog Search](#)
[Search Tips](#)

Cite this
 Export to Endnote

Similar Items

[Life and letters of Ythomas Kilby Smith, brevet major-general, United States volunteers, 1820-1837.](#)
By: Smith, Walter George, 1854-1924.
Published: (1898)

[Allergy and anaphylaxis as metabolic error.](#)
By: Godlowski, Zbigniew Z.
Published: (1962)

[The Metabolism of oral tissues](#)
Published: (1960)

[Clinical immunology and allergy](#)
By: Crippe, Leo Hermann, 1896-
Published: (1969)

[Ascorbic acid metabolism in adrenal gland and reproductive tissues of the guinea pig.](#)
By: Devine, Marjorie Mealey.
Published: (1967)

[Metabolic bone disease : cellular and tissue mechanisms /](#)
Published: (1969)

[Adipose tissue : regulation and metabolic functions /](#)
Published: (1970)

[Energy metabolism : tissue determinants and cellular corollaries /](#)
Published: (1992)

[Lipid metabolism in tissue culture cells /](#)
Published: (1967)

[Lipid metabolism in tissue culture cells.](#)
Published: (1967)

Allergy and tissue metabolism [by] W. G. Smith.

Main Author:	Smith, Walter George.
Language(s):	English
Published:	Springfield, Ill., C. C. Thomas [1964]
Subjects:	Allergy Tissue metabolism
Physical Description:	110 p. illus. 22 cm.
Original Format:	Book
Original Classification Number:	QR 185 .A4 S662
Locate a Print Version:	Find in a library

Viewability:

Limited (search-only) (original from University of Michigan)

[About](#)
[Help](#)
[Feedback](#)
[Take-Down Policy](#)

[http://catalog.hathitrust.org/Record/002077547\[6/13/2012 12:47:47 PM\]](http://catalog.hathitrust.org/Record/002077547[6/13/2012 12:47:47 PM])

79. Only the bibliographic information is displayed for this work. As reflected in the above screen shot, the “viewability” of this work is search only; none of the text of the work is available for display or download. (Only certified users who have print disabilities are able to access the text through a secure network.)

80. If the user searches for the same term in this book, a screen showing the page numbers for each use of the term is displayed as follows:

The screenshot shows the HathiTrust Digital Library interface. At the top, there is a search bar with the text 'anaphylactic shock' and a 'Find' button. Below the search bar, there are navigation links: Home, About, Collections, My Collections. A 'Back to page' link is also visible. The main content area is divided into several sections:

- About this Book:** Title 'Allergy and tissue metabolism [by] W. G. Smith. ...', author 'Smith, Walter George.', and copyright information 'Copyright: Protected by copyright law.' There is a link to 'View full catalog record'.
- Get this Book:** A link to 'Find in a library'.
- Add to Collection:** A section with a 'Login' button and a 'Select Collection' dropdown menu.
- Share:** A section with a 'Permanent link to this book' and a URL: 'http://hdl.handle.net/2027/mdp.39015006705670'.

The search results section is titled 'anaphylactic AND shock matched 41 pages in this item.' It includes a 'Limited (search-only)' button and a list of page ranges with the number of matching terms: p.17 - 12 matching terms, p.26 - 13 matching terms, p.30 - 11 matching terms, p.27 - 12 matching terms, p.15 - 11 matching terms, p.88 - 9 matching terms, p.9 - 8 matching terms, p.28 - 8 matching terms, p.18 - 8 matching terms, and p.21 - 8 matching terms. There are also navigation links for the search results: 'Viewing results for: 1 to 10 of 41 pages' and '1 | 11 | 31 | 32 | next ►'.

http://babel.hathitrust.org/cgi/pt/search?id=mdp.39015006705670;view=image;seq=7;ql=-anaphylactic%20shock;start=1;size=10;page=search;orient=0[6/13/2012 1:07:19 PM]

81. Based upon the search results, the user may decide to purchase the book or check out a physical copy from one of the member libraries. The text of books is not made digitally available unless it is determined that they are in the public domain or unless the rights holder has given us permission to provide access to the work.⁵

82. Without the ability to search the entire full text, the content within these resources—as distinct from basic bibliographic information describing that text—is invisible, or nearly so, to the majority of researchers.

83. Moreover, full-text searching is incredibly helpful even with respect to resources that could be identified as potentially relevant through a catalog search. For example, many

⁵ The exception to this is that Michigan students, faculty and staff certified through the Office of Services for Students with Disabilities as having print disabilities are granted access to digitized files of presumed in-copyright works.

libraries, including the University Library, store a substantial portion of their collections in offsite facilities, these materials are not immediately available to the scholar, and their retrieval may require significant library staff time. If full-text searching is available, a researcher can use it to determine whether a potentially relevant off-site work may be pertinent to her research.

84. Indeed, the HDL empowers scholars to perform types of research on a scale that simply could not be performed before the HathiTrust libraries digitized their collections. For example, and as explained in more detail in the Declaration of Dr. Neil R. Smalheiser, a digital research method commonly called “text mining” is already proving itself a vital tool for scholarly research and could potentially have application to works within the HDL corpus.

85. In short, having a digitized library of the magnitude represented by the HDL has the potential to yield breakthrough scientific discoveries – potentially lifesaving discoveries – that simply would not be possible if the service ceased to exist.

I. HathiTrust’s Efforts to Preserve the Libraries’ Collections and the Cultural Record

86. One of the primary goals of HathiTrust is the preservation of the published record of human knowledge through the creation of reliable and accessible electronic representations of the works within the corpus.

87. The use of redundant storage in geographically remote locations ensures long-term preservation of digital data by protecting against the total loss that would otherwise occur from the failure or destruction of the primary storage.

88. The HDL corpus is stored at Michigan with a “live” mirror site located at Indiana University’s Indianapolis campus.

89. The existence of the mirror site allows for balancing the load of user web traffic to avoid overburdening a single site, and each site acts as a back-up of the HDL collection in the

event that one site were to cease operation (for example, due to failure caused by a disaster, or even as a result of routine maintenance).

90. To allow for recovery of the HDL collection in the event of a disaster, automatic tape backups are created. Two encrypted backup tapes are created and are stored separately from each other, as well as from the two “live” storage instances, and the tape backups are not connected to the Internet. In the event of a disaster causing large-scale data loss to the primary HDL corpus at Michigan and the mirror site at Indiana University’s Indianapolis campus, the backup tapes could be used to restore the lost data.

91. The HDL has been certified as a trustworthy digital repository by the Center for Research Libraries (“CRL”) through their rigorous Trustworthy Repositories Audit and Certification (“TRAC”) assessment program, which included an in-depth preservation audit of the HDL. Attached as **Exhibit C** is a true and correct copy of this certification report.

92. This audit began in November 2009 and was completed in December 2010. Only a small number of digital repositories have been granted this certification.

93. In addition to protecting the digital data in the HDL from loss through disaster or mechanical failure, Michigan employs many levels of security to control access to the content in the HDL. The security employed by Michigan with respect to the digital library meets, and in many ways exceeds, the specifications developed by the parties in the Google Books proposed settlement.

94. First, Michigan maintains, and requires the University of Indiana to maintain, rigorous physical security controls. HDL servers, storage, and networking equipment at Michigan and Indiana University are mounted in locked racks, and only six individuals at Michigan and three at Indiana University have keys. The data centers housing HDL servers,

storage, and networking equipment at each site location are monitored by video surveillance, and entry requires use of both a keycard and a biometric sensor.

95. Second, network access to the HDL corpus is highly restricted, even for the staff of the data centers housing HDL equipment at Michigan and Indiana University. For example, two levels of network firewalls are in place at each site, and Indiana University data center staff do not have network access to the HDL corpus, only access to the physical equipment. For the backup tapes, network access is limited to the administrators of the backup system, and these individuals are not provided the encryption key that would be required to access the encrypted files on the backup tapes.

96. Web access to the HDL corpus is also highly restricted. Access by users of the HDL service is governed by primarily by the HDL rights database, which classifies each work by presumed copyright status, and also by a user's authentication to the system (e.g., as an individual certified to have a print disability by Michigan's Office of Services for Students with Disabilities).

97. Michigan staff who have been granted web access to in-copyright works in the HDL in order to perform a job function for HathiTrust (e.g., because they are involved in-copyright status determinations or image quality research) must use a specific, approved IP address *and* successfully authenticate to the system using their username and password. In addition, this web access is encrypted using Secure Socket Layer, a cryptographic protocol providing communication security over the Internet.

98. Even where we do permit a work to be read online, such as a work in the public domain, we make efforts to ensure that inappropriate levels of access do not take place. For example, a mass download prevention system called "choke" is used to measure the rate of

activity (such as the rate a user is reading pages) by each individual user. If a user's rate of activity exceeds certain thresholds, the system assumes that the user is mechanized (e.g., a web robot) and blocks that user's access for a set period of time.

99. A proxy detection system is also used. The proxy detection system consults real-time blacklist services to identify users that appear to be employing known proxy or anonymization servers to falsify a physical presence in the United States (in an attempt to subvert HDL use limitations that restrict access to some materials to users in the United States). If the proxy detection system identifies such a user, that user's access is blocked. Web access to the HDL is also logged by IP address and, when available, by username, the HTTP request string, and whether or not the volume requested is identified as presumed in-copyright in the HDL rights database. Such logs are regularly reviewed.

J. Access for Individuals with Print Disabilities

100. One of the primary goals of HathiTrust has always been to enable people who have print disabilities to access the wealth of information within library collections. We constructed the archive with the objective of making the world's first accessible research library. Access for people who have print disabilities is a part of our agreements with HathiTrust members and it is one of the core services around which the archive is built, along with preservation and search.

101. For centuries, libraries have been inaccessible to people who have a broad range of disabilities because library collections have not been available in accessible formats. As a result, individuals with print disabilities simply do not have equal access to library collections and are denied the full promise offered by libraries. This is particularly true in the academy, where access to the written record is at the heart of most scholarly pursuits.

102. For instance, given the number of works a student must review to write a typical term paper, she may have to wait weeks or even months to get access to converted works—in Braille or audio recording format—that she has not yet even been able to determine whether or not she will use. Even digitizing the works on a case-by-case basis can take weeks, which makes pursuing an education that much more difficult for a student who has a print disability.

103. Accessibility has been a hallmark of our digitization efforts from the earliest days and, when we discussed digitizing the library in connection with Google, it was one of our primary objectives to make our library collections immediately accessible to people who have print disabilities.

104. In fact, in 2005, when the National Federation of the Blind first contacted us about Google Books, we invited them to campus to demonstrate the accessible library we had already developed.

105. Our accessible library works like this:

- A person who has a print disability seeks certification from a qualified expert.
- The expert informs the library when a particular patron has a print disability for which digital access is a reasonable accommodation.
- We explain the digital library to the patron, we describe appropriate uses of the service (including warnings about copyright infringement), and we enable the patron to get secure access to the accessible library.
- If we have a digital copy of a work, the authorized patron with a print disability will have immediate access to that work in a format that can be made accessible through a variety of adaptive technologies. For example, the disabled user can enable software that translates the text into spoken words.

106. For our patrons who have print disabilities, this service makes it possible for these individuals to achieve their full academic and scholarly potential.

K. The Orphan Works Project

107. As noted above, we believe it is clear that scholars, students, and other patrons are more likely to discover and use works that they can access in a digital format.

108. Therefore, we strive to find ways to provide lawful access to digital works we have digitized and we spend millions of dollars each year to license access to digital content for our campus. Unfortunately, with orphan works, because a rights holder cannot be identified, there is no way to license digital access.

109. With this in mind, we developed a project that we called the “Orphan Works Project” (or “OWP” for short). The goal of the OWP was to identify orphan works and then to enable limited uses of those works to students, scholars, and walk-in patrons.

110. The OWP, as initially contemplated, had two steps: (i) first, we identify potential orphan works through a diligent, reasonable process that eliminates works that are claimed by a putative rights holder or that are otherwise found not to be orphans; and (ii) based on the results of the first step, we planned to enable limited uses of orphan works by Michigan students, scholars, and walk-in patrons.

111. As contemplated, the OWP would have allowed users access to orphan works for the purpose of online review, with the number of users permitted to view a given work limited at any one time to the number of copies held by the library. Readers would have been reminded, through watermarking and other explicit notices, that the books are subject to copyright.

112. After completing its initial process to identify potential orphan works, Michigan posted a list of *potential* orphan works on the Michigan library website on or about July 15, 2011 and provided a link to the list on the HathiTrust website. The public posting was a conscious

part of the identification process. With more eyes on possible orphan works, it was our intent to increase the accuracy of the identification process.

113. The public scrutiny component of the OWP worked as intended in instances when works identified as potential orphan works were claimed by the putative rights holder or the rights holder was identified by a third party. Had the project moved forward, these works would not have been treated as orphan works.

114. In evaluating the methods used to determine potential orphan works, we concluded that there were flaws in our pilot process and that we needed to remedy those flaws before moving ahead with the OWP. We therefore suspended the process and never proceeded to the second step of the project (i.e., we never proceeded to enable limited uses of putative orphan works).

115. Michigan, which is the only member of HathiTrust that has actively engaged in the work of the OWP, is continuing to study ways to improve the candidate identification process. In fact, we reached out to plaintiff The Authors Guild and other associations (including the Association of American Publishers) to invite their input on ways to improve the candidate identification process. After initially expressing interest in speaking with us and participating in this process, the Authors Guild thereafter abruptly filed this lawsuit.

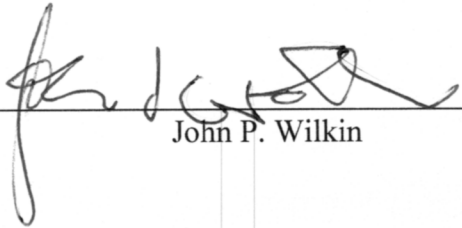
116. Not a single patron has been given access to a work through the OWP and at present, we do not know whether or how the OWP will continue.

117. In the event that Michigan decided to move forward with the OWP and provide access of works to users through the project, it would seek to comply with the requirements of section 108(e) of the Copyright Act.

118. While I am not a lawyer, it is my understanding that section 108(e) of the Copyright Act authorizes distribution of copyrighted works provided certain requirements are satisfied, including requirements that: (i) the library first determine, on the basis of a reasonable investigation, that a copy of the copyrighted work cannot be obtained at a fair price; (ii) the copy becomes the property of the user; (iii), the library or archives has had no notice that the copy would be used for any purpose other than private study, scholarship, or research; and (iv) the library or archives displays prominently, at the place where orders are accepted, and includes on its order form, a warning of copyright in accordance with requirements that the Register of Copyrights shall prescribe by regulation.

I declare under penalty of perjury under the laws of the United States that the foregoing is true and correct.

Date: June 28, 2012



John P. Wilkin