

**Briefing Paper on Progress and Opportunities for HathiTrust  
Ithaka S+R for the HathiTrust Strategic Advisory Board  
15 July 2011**

**Executive Summary**

The Strategic Advisory Board for HathiTrust, a library-based initiative to provide preservation and access for library content digitized by Google and other parties, is preparing for a Constitutional Convention in October 2011. At the Convention, representatives from the libraries that participate in HathiTrust will gather to discuss the future direction of the initiative.

To help frame the agenda for the Convention, Ithaka S+R has conducted research on the attitudes and needs of existing HathiTrust partner libraries and libraries that do not yet participate, to help the Board frame the most important issues for the October meeting. Drawing on a survey of HathiTrust participating libraries, interviews with HathiTrust participating libraries and libraries that do not yet participate in HathiTrust, and other research, we see the following as important next steps and Convention discussion items for HathiTrust:

- *Clearly defining objectives for the next 3-5 years, possibly mapping out the rationale for those objectives in the context of a revised mission statement, and enhancing the pipeline of information about HathiTrust's strategic priorities to partner libraries.* Respondents to a survey of HathiTrust partner libraries and follow-up interviewees spoke in particular about the need for clear communication from the HathiTrust staff and leadership, noting in several cases that they have not always felt well apprised of contemplated new initiatives. To the extent that HathiTrust is able to provide a more constant flow of information about proposed initiatives, that would seem to be valued by partner libraries.
- *Advancing changes that allow for nomination and selection to the Executive Committee, and clearly communicating the options to the HathiTrust partner community ahead of the Constitutional Convention.* A number of libraries, particularly those that have contributed content to the archive, noted that they expect heavy content contributors and/or founding members to have a more heavily weighted voice in governance under any future model.
- *Demonstrating to partner libraries the sustainability and feasibility of the new cost model for HathiTrust.* Some current participants raised concern about the increasing costs of HathiTrust, and libraries that do not yet participate worried about a perceived instability in the pricing structure. If HathiTrust elects to move forward with a version of the proposed cost model, there may be value in introducing a stabilizer into the cost model—either asking that institutions commit to membership for a set period, or calculating the divisor in the formula for public domain volumes (the number of institutions) using a rolling average that takes into account anticipated uptake in the next year.

## Introduction

In October 2011, representatives from the partner libraries that make up HathiTrust (<http://www.hathitrust.org>) will gather for a Constitutional Convention, where they will reflect on the progress of the past three years, discuss potential future directions for HathiTrust, and vote on a series of ballot initiatives related to governance, administration, and finance. The Strategic Advisory Board for HathiTrust, a group of library leaders who oversee the strategic direction for the initiative, engaged Ithaka S+R to prepare a briefing paper to help the Board and HathiTrust's other committees and working groups to frame the broad discussion for the Convention.

As part of our working approach for the project, we sought to weigh respondents' feedback and the opportunities that they presented against HathiTrust's broader mission "to contribute to the common good by collecting, organizing, preserving, communicating, and sharing the record of human knowledge," and also against its six subsidiary goals:

- To build a reliable and increasingly comprehensive digital archive of library materials converted from print that is co-owned and managed by a number of academic institutions.
- To dramatically improve access to these materials in ways that, first and foremost, meet the needs of the co-owning institutions.
- To help preserve these important human records by creating reliable and accessible electronic representations.
- To stimulate redoubled efforts to coordinate shared storage strategies among libraries, thus reducing long-term capital and operating costs of libraries associated with the storage and care of print collections.
- To create and sustain this "public good" in a way that mitigates the problem of free-riders.
- To create a technical framework that is simultaneously responsive to members through the centralized creation of functionality and sufficiently open to the creation of tools and services not created by the central organization.<sup>1</sup>

How do libraries' stated perceptions of, and needs from, HathiTrust converge with these goals, and where do they diverge?

In the sections that follow, we draw on our research to address three angles of input from the library community:

- (1) *How participating libraries perceive the value of HathiTrust.*
- (2) *Participating libraries' expectations for operating and governing the initiative moving forward.*
- (3) *The views of libraries that do not yet participate in HathiTrust.*

---

<sup>1</sup> [http://www.hathitrust.org/mission\\_goals](http://www.hathitrust.org/mission_goals)

## Methodology

Ithaka S+R carried out several types of research. First, we administered a web-based survey to the 52 libraries that participated in HathiTrust as of 15 May 2011. The survey was open from 18 May to 3 June, and received 46 responses, for a response rate of 88%. The survey asked respondents a series of attitudinal and factual questions related to HathiTrust, to other preservation and access platforms that the partner libraries may use, and to the libraries' internal policies and spending on digital preservation. The survey was emailed to each library's designated representative for HathiTrust, with a request to forward the survey to the library's director or to someone who could speak to the library's overall perspective on participating in HathiTrust.

We supplemented the survey results with a series of follow-up calls with eight survey respondents, both to clarify their initial responses and to ask follow-up questions. To broaden our understanding of the academic and research library community's perceptions of, and needs from, HathiTrust, we sought interviews with representatives from 25 libraries that do not yet participate in HathiTrust, of which 14 consented to an interview.<sup>2</sup> To get an anecdotal perspective of the ways in which the access and research functions of HathiTrust might be made more useful to scholars, we sought interviews with 15 scholars and subject librarians in history, language and literature, linguistics, and information science, of whom five consented to speak with us. (Although these interviews informed our research approach, they are not discussed directly in this paper.) Finally, throughout the project, our meetings and interviews with HathiTrust stakeholders, including members of the Strategic Advisory Board and the Executive Director, provided a helpful internal perspective.

## Value to Participating Libraries

In the survey and follow-up interviews, participating libraries strongly voiced their support for HathiTrust, citing its value as a digital preservation service, as an alternative to what more than one respondent called "the Google monopoly," and as an initiative growing from, and led by, the research library community. At a high level, there is clearly a great deal of goodwill for the initiative and its mission.

### *How do libraries rate HathiTrust's progress in meeting short-term functional objectives?*

As part of the research process, the Strategic Advisory Board requested that Ithaka S+R use an evaluation of HathiTrust's functional objectives (<http://www.hathitrust.org/objectives>), a list of developments undertaken by the HathiTrust community, as one guidepost for our research. Respondents to the survey were asked to indicate their agreement with a series of statements related to the functional objectives, using a scale from 1 to 10 (where 1 represented strong disagreement with the statement in question and 10, strong agreement).

The average and median responses showed overall satisfaction with the quality of the image scans in HathiTrust and the Trust's capacity for preserving and migrating content. Respondents

---

<sup>2</sup> In general, libraries outside the United States were less responsive to our requests for an interview.

gave slightly lower marks to the HathiTrust user interface and the branding/watermarking of items contributed by member libraries, although the median (5) for each of these two statements falls in the middle of the 1-10 satisfaction range.

*Table One.* Satisfaction with HathiTrust's progress on functional objectives

<u>Survey Statement</u>	<u>Average</u>	<u>Median</u>
HathiTrust is well positioned to migrate the content that my library has contributed to new formats as the need arises.	6.98	7
The quality of scans presented in HathiTrust is generally high.	6.91	7
Compared with my library working on its own, HathiTrust would be more efficiently and effectively able to migrate the content that my library has contributed to new formats.	7.12	8
We are satisfied with the prominence and placement of our organization's branding on individual items that we have contributed to HathiTrust.	5.63	5
The HathiTrust public user interface meets all the needs of our users.	4.74	5

*Branding.* Participating libraries raised two concerns about the local branding of their materials in HathiTrust. First, several respondents voiced a desire for a more prominent marking on the materials that they have contributed to HathiTrust than the marks of ownership that appear on materials now—one respondent called the existing watermarking arrangement “pretty subtle,” and another reported that they would like the materials they contribute to carry their institution's logo or something more visually attractive than plain text.

Second, a small number of participating libraries indicated that they expect their branding to appear on the scans of all items that they have contributed, even if the version being presented to the user was not itself digitized using the library's copy. As one respondent put it: “If [an item in HathiTrust] remains primarily branded by whichever institutions contributed the display copy, it [HathiTrust] will not give us the promotional edge we need.” As libraries further transition from local to networked models for collections and services, they may become more comfortable giving up control over the placement of their name in services like HathiTrust; however, “at this point in time,” one respondent wrote, “it is important to make the institutional branding [in HathiTrust] more prominent in order to help us to justify participation.”

*User interface.* Respondents pointed to a number of incremental fixes to the HathiTrust user interface, noting in particular the importance of reflecting serials holdings accurately in search results: two respondents noted that individual serial volumes need to be indexed correctly,

particularly for those serials that were bound as a single volume by a contributing library and are now reflected in HathiTrust as one item.

*To what extent does participation in HathiTrust help libraries to realize cost savings or cost avoidance?*

*Digital storage.* One potential value for joining HathiTrust is the ability to free up resources (funds, staff time, local server space) that would be devoted to local storage of digitized content and divert them to other functions. The survey asked participating libraries to indicate whether participating in HathiTrust caused them to change their spending on storing, managing, and preserving digital assets using their institution's local infrastructure, and if so, by what dollar amount. Of the 25 participants that reported having contributed content to HathiTrust, only five reported having changed their local spending as a result of HathiTrust. Of those five:

- Two reported increasing local spending; in both cases the respondents cited the staff time associated with preparing HathiTrust files for ingest as the driver.
- Three reported decreasing their local spending, although only one was able to provide an estimate of those savings (\$150,000).

Although the survey did not prompt libraries to estimate cost avoidance, several libraries noted in free response boxes within the survey that they prepared projections comparing the costs of storing digital content locally with the cost of joining HathiTrust before deciding to join the initiative.

A number of respondents made clear in follow-up interviews that participation in HathiTrust is not simply a question of money—the overarching mission, collaboration among the academic and research library communities, and the chance to help prioritize developments for the archive are all important drivers for participation. With that said, the question of how HathiTrust impacts local storage costs seems important because the initiative is considering moving to a new cost model; the model, which is outlined on the HathiTrust website, uses a formula that apportions a fraction of the storage cost for a digitized volume to all partner libraries who have owned a print copy of that volume.<sup>3</sup> As HathiTrust grows and sets new directions, the proposed cost model allows for libraries' participation fees to increase in a way that was unlikely under the existing cost model (which was based solely on the cost per GB of digitized content, a cost that libraries could control): for example, if a library were to join HathiTrust with a significant amount of public domain content, all partner libraries could see an increase in fees. In a period of challenged library budgets, the ability of a service like HathiTrust to demonstrate cost savings for libraries—particularly with the potential for an increase in membership fees—will likely become more important.

*Print collections management.* Another potential source of cost savings or avoidance for HathiTrust participants is the use of the digital archive as a tool for managing print collections, as libraries face the familiar problem of space constraints. However, Ithaka S+R's past research

---

<sup>3</sup> <http://www.hathitrust.org/cost>

has shown relatively cautious interest among in the U.S. higher education community about the speed of a print-to-electronic transition for books: the Ithaka S+R Faculty Survey 2009 found that only 4% of respondents agreed with the statement “Within the next five years, the use of e-books will be so prominent among faculty and students that it will not be necessary to maintain local collections of hard-copy books”; in an Ithaka S+R survey of U.S. academic library directors conducted in 2010, just 7% of respondents agreed with a similar statement.<sup>4</sup>

Those relatively low responses may shed light on why HathiTrust participating libraries seem not to be using HathiTrust as an integral part of their print collections management strategy. While slightly more than half of the responding institutions reported that the existence of a digitized copy in HathiTrust is one factor that they consider in deciding whether to move or dispose of books, no library reported that the existence of a digitized copy in HathiTrust is a prerequisite for those decisions. (Anecdotally, follow-up interviewees reported Google Books, Internet Archive, and other libraries’ repositories as other acceptable sources for a digitized copy when making print collections management decisions.) Similarly, only two libraries reported that the existence of digitized copies in HathiTrust has had an impact on the rate at which they move print books to off-site storage, and only six reported that HathiTrust has an impact on the rate at which they dispose of print books.

*Table Two.* HathiTrust’s importance in print collections management decisions

<u>Impact of HathiTrust on Decision-Making</u>	<u>Decision to move print books off-site</u>		<u>Decision to dispose of print books</u>	
	Content contributors	Other members	Content contributors	Other members
Prerequisite <sup>5</sup>	-	-	-	-
One Factor	13	12	12	12
Not a Factor	7	7	8	8
N/A	5	2	5	1

<sup>4</sup> See: Roger C. Schonfeld and Ross Housewright, *Ithaka S+R Faculty Survey 2009: Key Strategic Insights for Libraries, Publishers, and Societies* (7 April 2010), p. 23, <http://www.ithaka.org/ithaka-s-r/research/faculty-surveys-2000-2009/Faculty%20Study%202009.pdf>; Matthew P. Long and Roger C. Schonfeld, *Ithaka S+R Library Survey 2010: Insights from U.S. Academic Library Directors* (5 April 2011), p. 36, <http://www.ithaka.org/ithaka-s-r/research/ithaka-s-r-library-survey-2010/insights-from-us-academic-library-directors.pdf>.

<sup>5</sup> The “Prerequisite” row in this chart captures two response choices in the survey: “The existence of a digitized copy in HathiTrust that is immediately available for viewing by our users is a prerequisite” and “The existence of a digitized copy in HathiTrust, whether or not it is immediately available for viewing by our users is a prerequisite.”

*Table Three. Impact of HathiTrust on rate of changes to print collections<sup>6</sup>*

<u>Impact of HathiTrust on Rate</u>	<u>Moving print books off-site</u>		<u>Disposing of print books</u>	
	Content contributors	Other members	Content contributors	Other members
Increased	1	1	1	-
Maintained	-	-	2	3
Decreased	-	-	-	-
No Impact <sup>7</sup>	19	17	16	15
N/A	5	2	5	1

This is not to say that HathiTrust does not have a continued and vital role to play in reducing the costs associated with managing print collections. In the aforementioned Ithaka S+R Library Survey 2010, U.S. academic library directors were presented with the following concept statement (modeled closely on HathiTrust):

Suppose there existed a robust system for preservation of and access to historical monograph collections. In such a system, the millions of books digitized through the Google library digitization project would be readily available in digital form for use by your community. Digital preservation would be assured by a trusted third party archive, and the paper source materials would be preserved in a suitable number of print archives. Ultimately, discovery and accessibility would be greatly improved while preservation would be assured.

When asked how the existence of such a system might impact their print collections management strategies, 74% of respondents said that withdrawal of print books would be an important strategy for their libraries, and 84% of respondents reported that they would be more likely to withdraw their print book collections under such a scenario if their library faced significant space pressures.<sup>8</sup> In addition, 90% of respondents indicated that they would be likely to withdraw their print book collections if their library could provide guaranteed on-demand access to the print versions through a sharing network. Clearly, there is interest in the library community in using a system like HathiTrust for print collections management, and it may be that either an improvement in the discoverability and accessibility of titles in HathiTrust, the addition of a print preservation and/or access service, or the simple passage of time, will nudge libraries toward greater use of HathiTrust for print collections management.

<sup>6</sup> It bears particular noting that these figures were self-reported by libraries.

<sup>7</sup> The distinction between "Maintained" and "No Impact" was meant here to signify that a library could conceivably choose to maintain, rather than decrease, the rate at which they move or dispose of print books because of the availability of digitized copies of those works in HathiTrust.

<sup>8</sup> Long and Schonfeld (2011), pp. 36-37.

### **Future Operations for the Initiative**

HathiTrust's existing operational structure is designed to leverage cooperative work by its network of partner libraries: the work of a relatively modest central staff at the University of Michigan is bolstered by volunteered time from partner library representatives on working groups and governing committees. In addition, a smaller set of partner libraries are devoting more concentrated time to the development of tools for the benefit of the larger membership, as in the case of the California Digital Library's extensive work on metadata management and the collaborative development of the Research Center by Indiana University and the University of Illinois, Urbana-Champaign, which will allow researchers to conduct text mining across the HathiTrust corpus. The overall model of contributed staff time from partner libraries clearly appealed to the librarians interviewed and surveyed for the project—one interviewee called participation on a HathiTrust working group "one of the most satisfying experiences of my professional career."

The section below is largely reflective of the views of the survey respondents and interviewees, and so it most directly speaks to HathiTrust's existing organizational model. With that said, there remains an overall question about whether a structure that depends heavily on contributed staffing and the buying-out of staff time from partner libraries will be sustainable over the long term. Contributed staff time may be retrenched by partner libraries during difficult budget periods; even in those cases where HathiTrust buys out the time of developers and project librarians at member institutions, the availability of that time may be curtailed as the partner institution's priorities change. (And as one interviewee pointed out, there is not, for example, an infinite supply of developers employed by research libraries for HathiTrust to draw on.) Evaluating whether and to what extent staff positions might need to be brought into the central HathiTrust organization would be difficult to forecast (and beyond the scope of this project), but HathiTrust's leaders and governing organizations will need to keep a pulse on staffing needs and be willing to spend the organization's surplus to support added internal staffing if the need arises.

*Partner libraries' expectations about staffing.* HathiTrust's existing constellation of governance committees and working groups are made up of library leaders whose time is contributed by HathiTrust partner institutions. In addition, and as mentioned above, a smaller number of partner libraries have taken on technological development work for the initiative. On one hand, the decentralized and devolved committee structure allows the HathiTrust community to benefit from the expertise of high-level library leaders with a wide range of competencies; on the other hand, the volunteer-centric staffing structure means that the continued vitality of the initiative depends heavily on the willingness of partner libraries to continue making these contributions.

In the survey administered to the HathiTrust community, partner libraries almost universally reported contributing staff time to the initiative: 43 of the 46 responding partner libraries responded that they devote some level of staff time to HathiTrust; of those, 32 reported a relatively modest allocation of staff time, between 0.1 and 1.0 FTE.

*Table Four. Projected Change in Contributed Staff Time<sup>9</sup>*

<u>Current Staffing Level</u>	<u>Increase</u>	<u>Maintain</u>	<u>Decrease</u>
None	1	2	-
0.1 to 1.0 FTE	17	15	-
1.1 to 2.0 FTE	2	3	-
2.1 to 3.0 FTE	2	1	-
3.1 to 4.0 FTE	1	-	-
More than 4.0 FTE	-	1	-

When partner libraries were asked to forecast how they expect their staffing contributions to HathiTrust to change (if at all) over the next 3-5 years, a roughly equal number reported that they expect to increase their staffing contributions, or maintain their current level (with no increase or decrease). No partner library contemplated decreasing or eliminating staffing contributions to HathiTrust.

*For what functions might libraries expect to contribute staff time in the future?*

Even as the partner libraries give a clear indication that they expect to maintain or increase the quantity of staff time to HathiTrust, planning for the initiative will need to take into account the qualities of those staffing contributions. One indicator for future staffing needs is the extent to which partner libraries prefer that certain functions for HathiTrust be carried out at their local institution (that is, using their own staff and resources, as in the case of contributed staff time to HathiTrust). The following table summarizes partner libraries' responses to a series of questions about their preferences for performing a broad set of functions related to HathiTrust locally:

*Table Five. Preferences for HathiTrust Functions: At the local institution or centralized?  
(On a scale from 1 to 10, where 1 = Local and 10 = Centralized)*

<u>Function</u>	<u>Average</u>	<u>Median</u>
Digital library infrastructure management	5.77	6
Digitization	4.51	4
Outreach and marketing	5.11	5
Research support	7.04	8
Rights determination	7.20	7.5
Tool development	7.11	8
User support	5.11	5

Perhaps unsurprisingly, libraries expressed a preference for functions that require technology development time, like tool development and the formation of research support services, to be accomplished by the central HathiTrust organization; tech development time is costly and likely scarce at those partner libraries that have developers on staff. Rights determination emerged as another function that partner libraries prefer to push up to the central HathiTrust

<sup>9</sup> The survey questions addressing current and projected contributions of staff time instructed respondents to exclude staff time devoted to general digitization activities.

organization (a response that is, perhaps, itself an endorsement of great value of the HathiTrust staff's research on orphan works and other rights issues).

The continued ability of the HathiTrust's leaders and governors to secure the staff time needed for these more complex functions will be crucial to the development of the initiative. If it becomes apparent that HathiTrust partner libraries are not able or willing to contribute staff time at the level needed to develop the project's infrastructure and interface, then the HathiTrust fee structure would need to be revised to reflect these new staffing needs.

*Expectations about representation in governance.* In the survey responses and follow-up interviews, several respondents voiced a desire for a method of consultation and input into the governing committees, including the Executive Committee. Respondents and interviewees noted that, under the initial HathiTrust governance model, there is not yet a process for nominating and selecting representatives to serve on the Executive Committee; they pointed to a process for doing so as a clear priority for the initiative.

With that said, interviewees had different views on the question of who should be eligible to serve on the governing committees: several respondents and interviewees from libraries that have contributed content posited that representation on the Executive Committee might be reserved for those libraries that have contributed content or have devoted significant staff time to the initiative; this has the virtue of rewarding those foundational libraries, but is also clearly at odds with the interests of libraries that have not contributed content but chose to join in part to have input into HathiTrust's strategic direction. (Ten of the 21 libraries in the survey who reported not having contributed content to HathiTrust also noted the opportunity to have input into the strategic direction of HathiTrust as one of their top three reasons for joining.)

If HathiTrust continues to operate under a membership governance model, the method for selecting representation may need to fall between two poles: on the one side, recognizing the contributions made by the heaviest contributors and/or founding members of the initiative, and on the other side, preserving some representation for those libraries with smaller collections that might wish to join HathiTrust in the future. One possible mechanism is a weighted voting system that gives those institutions that have contributed content and fulfill certain other parameters a greater share of voting power; another possibility would be a bicameral governance structure in which (for example) the upper committee would be reserved for founding members and heavy contributors of staff time, and the lower committee would be more broadly representative of the HathiTrust partner community.

*Desire for a clear communication of strategic objectives.* In the survey, respondents' questions about a model for governance and input moving forward are closely tied to confusion about the future priorities for HathiTrust: partner library representatives questioned both how HathiTrust will be shaped in the future, and how those changes will be communicated to the membership. "There is a lot of communication about what's happening already in HathiTrust, but not a lot about what's being contemplated," one participant noted. "For example, the development of the new cost model is something that we wish we could have been asked about: is this being

presented [at the Constitutional Convention] as something that we're thinking about, or as something that we definitely are using?" Although the articulation of a longer-term strategy for HathiTrust is one likely outcome of the Constitutional Convention and will itself serve as a useful opportunity to communicate with the HathiTrust community, the future governance structure for HathiTrust will need to include mechanisms for disseminating news of HathiTrust's priorities to the community, and for soliciting input in return.

*Intersection with other digital library initiatives.* Several respondents to the survey asked how HathiTrust will interact with the Digital Public Library of America (DPLA) initiative.<sup>10</sup> (Indeed, one scholar-interviewee who had not previously been aware of HathiTrust asked whether it was connected in some way with the DPLA.) To the extent that it will be possible at the Constitutional Convention, the Strategic Advisory Board and the Executive Committee should seek to clarify whether and how HathiTrust will interact with that initiative (itself still being defined).

### **Value to Potential Partner Libraries**

As detailed in the Methodology section, Ithaka S+R spoke with representatives from a range of libraries that did not yet participate in HathiTrust, including six ARL institutions, three U.S. liberal arts colleges, and five international libraries outside North America. These include a roughly even number of institutional interviewees who felt they would have content to contribute to HathiTrust, and those who do not currently contemplate engaging in digitization beyond small experiments. All of the interviewees expressed great interest in HathiTrust, and most reported that they had seriously contemplated joining HathiTrust over the past two years. (One of these libraries chose to join HathiTrust during the course of this research.) Their reasons for not joining provide useful context as the HathiTrust community considers shifting models for governance and finance, each of which will have implications for the ability of the initiative to attract new members.

In rough terms, the libraries that have not yet joined can be split into three categories:

*Libraries that only have interest in contributing special collections.* These institutions may have a small or nascent digitization program focused on rare and unique collections, and see a clear value in aggregating those materials with the larger HathiTrust corpus.

- These libraries have a strong interest in the mission and preservation function of HathiTrust, and possibly in using HathiTrust as a factor in making print collections management decisions.
- Among our interviewees, those at libraries that have examined HathiTrust in the past voiced a perception that the initiative does not yet have an effective ingest mechanism for special

---

<sup>10</sup> A useful summary article on the planned Digital Public Library of America is here: Gautam S. Kumar and Sirui Li, "Harvard Leads Digital Library Initiative," *Harvard Crimson* (12 April 2011). <http://www.thecrimson.com/article/2011/4/12/project-committee-library-harvard-google-smithsonian-digitize/>

collections content, although one library noted that they sense from conversations with HathiTrust staff that this is changing.

- Cost was cited as a major concern for libraries in this broad category. These libraries will be charged in the new cost model based on their overall print holdings overlap with HathiTrust. For smaller libraries, this may not sum to a huge fee, but these library interviewees noted a break in the linkage between their perception of the service they would get from HathiTrust (hosting and preservation for a subset of special collections material) and the price they would be asked to pay (which will be linked to overall print holdings). It is important to note that these interviews were conducted before the universities of Michigan, Wisconsin, and Florida announced their intention to begin making orphan works available to their authorized campus users via HathiTrust; libraries in this category may well judge the value of HathiTrust differently now that there may be additional access benefits to HathiTrust beyond the existing preservation and other benefits.

*Libraries that participate in regional or consortial digital library projects, and see HathiTrust as duplicative.* These libraries may have robust digitization programs, including some scanning of general collections, but they receive repository services from a consortial or regional digital library. The fees for these solutions may be wrapped into the participation fees for the consortium, or, as in at least two cases with our interviewees, the costs may be paid by a governmental body.

- Those libraries active in consortia voiced some hesitance to make a move without the consent and participation of their partners in existing arrangements: “We’ve invested heavily in our consortium, both financially and philosophically.” There is a flip side to this, in that two such interviewees also noted that they would expect to receive a consortial discount on HathiTrust participation fees.
- The cost of participation in HathiTrust is also a concern for this group, particularly for those that receive free or low-cost digital storage through a governmental body. “For all the emotional reasons, I would like to support [HathiTrust],” one interviewee said. “But what do I show my faculty and staff we’re getting for the money?” As the process for making orphan works accessible in HathiTrust becomes clearer, and as the Universities of Indiana and Illinois launch the HathiTrust Research Center, designed to facilitate text mining across the corpus, libraries will be able to more clearly weigh the access benefits when making a decision about membership.
- Yet, these same libraries noted some interest in using the availability of digitized copies in HathiTrust as a print collections management tool, and a general interest in supporting a community-driven initiative. (One interviewee stated this in opposite terms: “We don’t want to be seen as a free-rider.”)

*Libraries that do not have a strong preservation emphasis.* These libraries simply may not have a strong interest in devoting funds to an initiative around legacy print collections. As an interviewee at one such library put it: “Our priorities are on supporting emerging modes of scholarly communication,” with the implicit belief that HathiTrust is primarily a preservation solution for legacy materials.

- These libraries may have an interest in new initiatives like the Research Center and, perhaps, in the ability to open up access to orphan work..

Several concerns cut across these categories:

- *Confusion about the proposed cost model.* In particular, library representatives wondered what would happen if other libraries were to exit HathiTrust—which they see as a distinct possibility in what remains a difficult economic climate for academic libraries. The number of partner libraries is an inverse driver of costs in the proposed model—that is, in general, the more libraries, the lower the cost for each individual library—but, as several interviewees pointed out, the model does not include a control for the number of libraries used to calculate HathiTrust membership fees. In general, clearer explanation and a great deal of outreach to libraries who are concerned about this would be needed in order to convince them to join.
- *For libraries outside the United States, a concern about joining a U.S.-based initiative.* One library felt that they were precluded from joining HathiTrust by European law; a Canadian research library noted a general hesitancy among their peers to place their content on U.S.-based servers, given the different cultural and legal context. Intensified outreach efforts would be needed to allay concerns like these if the HathiTrust leadership decides that they would like to greatly increase international participation in the near term.

### Other Strategic Concerns

Partner libraries clearly value the preservation function of HathiTrust—in the survey of participating libraries, this was listed most often as a library’s principal motivation for joining HathiTrust.<sup>11</sup> But access to the works was also clearly stated as a value by the library representatives who participated in our research, and this was particularly true for libraries that do not yet participate: several of the libraries do not have a strong research and preservation focus, but want to be able to make a wider volume of the content they own in print available to their users in digital form.

During the course of this research, the universities of Michigan, Wisconsin, and Florida, announced that they would make digitized copies of orphan works available to authorized users via HathiTrust, a move that could theoretically be adopted by other HathiTrust partner libraries, as well.<sup>12</sup> If this capacity can be extended to other libraries, then HathiTrust can clearly address the needs of the larger constituency of libraries that do not have a great deal of their own digitized content to contribute. Indeed, this offers the HathiTrust staff an opportunity to use print holdings comparisons analysis (using WorldCat or other means) to identify a priority set of libraries to pursue for the HathiTrust partnership: those libraries that have local collections with a large proportion of orphan works and that would greatly benefit from a HathiTrust initiative to help libraries open up access to orphan works for which the partner library owns a copy.

<sup>11</sup> See Appendix C, questions 3 and 4.

<sup>12</sup> <http://blogs.uflib.ufl.edu/news/2011/07/14/university-of-florida-libraries-join-the-hathitrust-digital-library-and-expand-digital-access-to-%E2%80%9Corphan%E2%80%9D-works/>

Given the strong statements by interviewees of the potential value of access to digitized works held in HathiTrust, the HathiTrust staff and community may wish to focus their limited outreach time on the libraries that would be identified by this analysis.

In addition to the opportunities provided by the orphan works access initiative, there are at least two other avenues that HathiTrust could consider in order to build on the value of the initiative as a digital preservation solution:

- *More actively pursuing permissions from publishers.* HathiTrust has already experimented with this in a sense, as some books from the University of Michigan Press (itself tightly integrated with the Michigan Libraries) are made openly available in HathiTrust. While establishing a Publisher Relations function within HathiTrust would be a significant shift for the project, pursuing agreements with publishers would allow HathiTrust to open a greater amount of content to end-users while strengthening HathiTrust's value proposition as a preservation service to its library constituents.
- *Focusing on ingesting special collections, possibly including formats other than books.* Libraries are increasingly digitizing special collections content, and it may be that HathiTrust has a role to play in serving as both a preservation and access aggregator for this content. (Indeed, several interviewees from libraries that do not yet participate in HathiTrust wondered aloud whether HathiTrust could accept their special collections content.) There are several contingencies to such a plan. First, several library representatives noted the time-cost of preparing files for ingest into HathiTrust as a concern, so partner libraries would need to be apprised of, and comfortable with, that time burden. In addition, there are questions about how to effectively present special collections content to facilitate use in research and teaching (particularly with multiple format types if that is contemplated). In Ithaka S+R's past client research for primary source collections, we have conducted extensive background research on existing primary source products for teaching and scholarship, and on faculty preferences for such collections. The overarching concern for such products is that, in order to be effective for a broader range of end-users than researchers, there needs to be some form of editorial shaping, in the form of curated collections of content, suggested excerpts for classroom or student assignment use, and/or contextual explanatory material sitting around the primary sources. Merely aggregating a collection of special collections or primary source materials is not likely to provide a coherent experience for the end-user or drive the level of usage that would demonstrate to partner libraries the impact and value of placing their collections in HathiTrust.
- *Developing a comprehensive strategy for tangible materials.* There is already work underway to determine if HathiTrust can provide for the preservation of print originals following their digitization. Ithaka S+R has seen some indications that supporting print preservation via business models organized principally around access to and preservation of digitized versions may be sustainable (similar to approaches taken by LLMC and JSTOR). In addition, based on the responses of library directors to the Ithaka S+R Library Survey, HathiTrust may find it valuable to offer opportunities to provide for on-demand access to

these print originals, at least to HathiTrust members, whether directly or in partnership with print sharing networks.

### **Potential Next Steps and Questions for the HathiTrust Community**

The purpose of this report is not to prescribe action for HathiTrust—“our survey and interviews seem to say *x*, therefore HathiTrust should do *y*”—but rather to surface opportunities and challenges, and offer those to the HathiTrust community in a way that allows the Strategic Advisory Board and the working groups to shape the agenda for the Constitutional Convention. With that proviso, our work suggests that the following items may be helpful for the HathiTrust community to consider. It also bears noting that each of these is interrelated: that is, questions about mission are closely tied to questions about the governance model, the size and composition of the membership, and the sustainability model that will be needed to support the archive.

- *Clearly defining objectives for the next 3-5 years, possibly mapping out the rationale for those objectives in the context of a revised mission statement.* In short, partner libraries want to know where HathiTrust is headed. All of the other questions or next steps that we suggest flow from this fundamental one: what will HathiTrust be in the future? Which of its six goals (quoted on the first page of this briefing paper) will prove to be the ones that drive the collaboration? While these six items are not necessarily mutually exclusive, it is easy to envision HathiTrust taking on initiatives that would align with HathiTrust’s mission goals as stated now—for example, a coordinating function for shared print storage—that might cause current and potential partner libraries to reassess HathiTrust’s value and costs.
- *Enhancing the pipeline of information about HathiTrust’s strategic priorities to partner libraries.* Respondents to the survey and follow-up interviewees spoke in particular about the need for clear communication from the HathiTrust staff and leadership, noting in several cases that they have not always felt well apprised of contemplated new initiatives. One interviewee said that their institution is quite eager to commit staff time to a HathiTrust development initiative, but they do not fully understand the menu and timeline of contemplated initiatives, and how they might effectively “plug in” their developers. Another interviewee tied a perceived lack of communication about new initiatives to a larger question about mission creep (e.g., “Will HathiTrust be taking on new content types? And if so, why?”). To the extent that HathiTrust is able to provide a more constant flow of information about proposed initiatives, that would seem to be valued by partner libraries.
- *If HathiTrust is to retain a membership-driven governance structure, there might be value in advancing changes that allow for nomination and selection to the Executive Committee, and clearly communicating the options to the HathiTrust partner community ahead of the Constitutional Convention.* A number of libraries, particularly those that have contributed content to the archive, noted that they expect heavy content contributors and/or founding members to have a more heavily weighted voice in any future governance model for HathiTrust.

- *Demonstrating to partner libraries the sustainability and feasibility of the new cost model for HathiTrust.* Some current participants raised concern about the potential for increased membership fees over time, and libraries that do not yet participate worried about a perceived instability in the pricing structure. If HathiTrust elects to move forward with a version of the proposed cost model, there may be value in introducing a stabilizer into the cost model—either asking that institutions commit to membership for a set period, or calculating the divisor in the formula for public domain volumes (the number of institutions participating in HathiTrust) using a rolling average that takes into account anticipated uptake in the next year.
- *Ultimately, making a decision based on the most pressing goals and objectives for HathiTrust, about how large the membership for the initiative needs to grow.* The proposed transitional cost model privileges the addition of new members, as each additional member helps to lower the costs for other libraries—an elegant example of network effects benefitting the HathiTrust membership. On the other hand, adding new institutions may introduce new challenges for HathiTrust, particularly if:
  - a wider range of content beyond general collections print holdings is taken into the archive, potentially increasing the overall costs for member libraries even if they only value HathiTrust for preservation of books, serials, or government documents, or;
  - if HathiTrust retains a membership-driven governance structure *and* expands membership beyond the research library community, in which case the need to accommodate libraries with modest budgets and a desire to have a voice in leadership may be at odds with the transitional cost model and a challenge to a governance structure that may privilege the voice of heavy content-contributing research libraries.

We raise these questions in order to prompt conversations around the future shape of the initiative and the role that current and future partner libraries will play in governing and sustaining HathiTrust. We thank all of the interviewees and survey respondents for their input throughout the project.

## Appendix A. Interviewees

### HathiTrust Participating Libraries

Bob Wolven, Associate University Librarian, and Janet Gertz, Director of Preservation and Digital Conversion Program, Columbia University (G,M)  
 Carol Hughes, Associate University Librarian, University of California-Irvine (G,M)  
 Heather Christenson, Mass Digitization Project Manager, California Digital Library (G,M)  
 John Wilkin, Executive Director of Hathi Trust and Associate University Librarian, and Jeremy York, Project Librarian for HathiTrust, University of Michigan (G)  
 Lisa German, Assistant Dean for Technical Services and Collections, Pennsylvania State University (G)  
 Rick Luce, Vice Provost and Director of Libraries, and Lars Meyer, Senior Director for Content, Emory University  
 Ruth Jackson, University Librarian, University of California-Riverside (G,M)  
 Thomas Leonard, University Librarian, University of California-Berkeley (G,M)

### Other Libraries

Artemis Kirk, University Librarian, Georgetown University  
 Carla Stoffle, Dean of Libraries and Center for Creative Photography, and Stephen Bosch, Budget and Licensing Librarian, University of Arizona  
 Cathrine Harboe-Ree, University Librarian, Monash University, Australia  
 Christine A. Fowler, Digitisation Manager, University of Southampton, U.K.  
 Harriette Hemmasi, University Librarian, and Andrew Ashton, Director of Digital Technologies, Brown University  
 Jeannette Frey, University Librarian, Lausanne Cantonal and University Library (G)  
 Jeff Hiroshi Gima, Director, American International Consortium of Academic Libraries<sup>13</sup>  
 Jonthan LeBreton, Senior Associate University Librarian, Temple University  
 Judith Russell, Dean of Libraries, University of Florida  
 Mary-Jo Romaniuk, Acting Chief Librarian, University of Alberta, Canada (M)  
 Patricia Killiard, Head of Electronic Services and Systems, and Grant Young, Digitisation and Digital Preservation Specialist, University of Cambridge, U.K.  
 Patricia Tully, University Librarian, Wesleyan University  
 Robert Kieft, College Librarian, Occidental College  
 Sabrina Pape, College Librarian, Vassar College

### Researchers/Subject Librarians

Fred Burchsted, Reference Librarian and Subject Specialist for History, Harvard University  
 Katherine Rowe, Professor of English, Bryn Mawr College  
 Kristin Eschenfelder, Assistant Professor of Information, University of Wisconsin-Madison  
 Joseph Janes, Associate Professor of Information, University of Washington  
 Miles Efron, Assistant Professor of Information and Library Sciences, University of Illinois, Urbana-Champaign

(G) denotes a library that participated in the Google Books project.

(M) denotes a library that participated in the Microsoft Live Search digitization project. (Alberta partnered with the University of Toronto on its Microsoft digitization project.)

---

<sup>13</sup> The American International Consortium of Academic Libraries, or AMICAL, is a group of teaching-focused higher education institutions in Europe and the Middle East.

## Appendix B. *Alternative organizational models*

The Strategic Advisory Board requested that Ithaka S+R provide a brief overview of potential alternative governance models for an organization like HathiTrust. (To prescribe one of these models—or even to argue definitively that the current model is insufficient for HathiTrust’s future needs—is outside the scope of this project.) Below, we suggest three broad organizational models that the Board may wish to consider, arranged roughly in order from most least centralized to most centralized.

*A consortium.* In a typical consortium, a group of institutions pools some level of resources to work toward a set of shared goals, often with some amount of central staffing to support the programmatic initiatives of its members. Consortia with deep programmatic emphases typically serve a fairly fixed number of members working together to pursue shared objectives in a trusted environment. Examples of strong consortia with programmatic emphases include CDL and CIC, both of which are governed at least to some degree at the institutional level in addition to the library level. Other consortia may be governed with equal votes of all the member libraries, or via other mechanisms.

*A membership organization.* Whereas some consortia are governed fairly informally and may not have independent legal status, a membership organization is a formally incorporated not-for-profit. In a membership organization, the constituents who make up the collective have varying degrees of decision-making power and channels to provide input, but they nearly always have the right both to approve bylaws to the organization’s constitution, and to elect members to the organization’s board of trustees. One of the largest and best-known such organizations in the library world is OCLC, which is overseen by a governing board, the members of which are elected by the OCLC membership. (OCLC now counts more than 25,000 institutions in its membership.<sup>14</sup>) At larger membership organizations, there may be nominating committees to help in the election process to the governing board.

*An independent not-for-profit organization.* Like a membership organization, the independent not-for-profit organization is ultimately governed by a board of trustees; in this case, though, that board of trustees is self-perpetuating, electing its own members as vacancies arise. One trait of self-perpetuation is that it can help to provide for somewhat more strategic continuity than certain types of membership organizations. At the same time, independent not-for-profit organizations are not governed directly by the participants and clients they seek to serve.

---

<sup>14</sup> <http://www.oclc.org/us/en/about/cooperation/default.htm>

**Appendix C. Anonymized survey data**

<b>1. Do you view HathiTrust primarily as a platform for preservation, or as a platform for access to content? Answer on a scale from 1 to 10, with 1 being "primarily for preservation," and 10 being "primarily for access."</b>	
1-2	1
3-4	8
5-6	26
7-8	7
9-10	4
Total Responses	46
Avg.	5.52
Median	5

<b>2. Has your library contributed content to HathiTrust?</b>	
Yes	25
No	21

**3. If your library has contributed content to HathiTrust, what motivated your library's decision to do so? Please use the radio buttons to RANK each choice that applies, with 1 being the most important reason, 2 being the second most important reason, etc. (Reflected here as the number of respondents placing each option in each available ranking spot.)**

Rank	I want to ensure the long-term preservation of my digitized collections.	I want our digitized collections to be discoverable as part of an aggregated collection of digital materials.	I want to use HathiTrust to drive my print collections management strategy.	I want to have a role in developing tools in HathiTrust that will enhance the value of my digitized collections.	I want to have a role in influencing the priorities for, and investments in, HathiTrust development.	Other
1	10	8	2	1	3	1
2	7	13	2	1	3	-
3	3	3	4	4	7	3
4	-	-	5	12	7	1
5	3	1	7	7	3	-
6	1	-	3	-	2	2
7	-	-	-	-	-	-

**If you ranked "Other" in the previous question, please describe your reason here.**

- Actually for us it is important for both preservation and access. We want to expose our locally digitized materials where appropriate via Hathi. So, Hathi is our primary repository for files from the Google project. We do not keep those locally. For local digitization projects we keep master files locally but would like to expose them in Hathi as well.
- Participation in strategies for access, discovery, and preservation as well as print collections management.
- I'm interested in using HathiTrust as an infrastructure for my other digital collections.
- We want to have our digitized content included in a larger corpus/aggregate to be used for large scale computer intensive research (text/linguistic analysis, image analysis, etc.) PLEASE NOTE: in ranking for this question: long term preservation is ranked last not because we do not value it but because currently we have our own local preservation strategy that does not rely on Hathi. All items listed about are IMPORTANT in the absolute sense.
- We want to expose our locally digitized materials where appropriate via Hathi. So, Hathi is our primary repository for files from the Google project. We do not keep those locally. For local digitization projects we keep master files locally but would like to expose them in Hathi as well.

4. If your library has not contributed content to HathiTrust, why did your library choose to join? Using the radio buttons, please RANK each reason that applies, with 1 being the most important reason, 2 being the second most important reason, etc. (Reflected here as the number of respondents placing each option in each available ranking spot.)

Rank	To ensure that HathiTrust content and tools remain available to the larger community.	In order to have input into HathiTrust's strategic direction.	Because we think we may contribute content to HathiTrust in the future.	As part of our print collections management strategy.	As part of our digital preservation strategy.	To contribute to the sustainability of the HathiTrust corpus of out-of-copyright materials.	Out of an interest in building tools for researchers on top of the HathiTrust corpus.	Other
1	1	4	3	3	6	2	-	1
2	3	4	4	1	6	1	1	-
3	2	2	5	1	1	5	4	-
4	2	4	6	2	1	5	-	-
5	4	4	1	2	1	4	4	-
6	5	-	2	5	3	-	3	-
7	1	1	-	4	2	2	7	-
8	-	1	-	-	-	-	-	3

**(4, cont.) If you ranked "Other" in the previous question, please describe your reason here.**

- As part of our commitment being a CIC member, we will be contributing materials within the next month.
- Joined as a member institution in a consortium that joined.
- As part of ongoing faculty education on access vs. ownership issues.
- We were members of the CIC (Committee on Institutional Cooperation) and joined in that context.

**5. Which unit(s) of your library advocated most actively for your library to join HathiTrust? Please check all that apply.**

Archives	Collection Development	Dean / University Librarian	Information Technology / Digital Libraries	Operations / Finance	Preservation	Technical Services	Other (please describe)	Not sure
2	27	41	24	1	10	6	6	0

**7. Other (please describe):**

- Scholarly Communication
- Public Services
- Scholarly Communications, Provost
- part of a system-wide decision to join
- Consortial decision
- [Our library] decided to join HathiTrust with the full support of the Library Executive Group, which is chaired by the university librarian. The decision-making forum is composed of AULs and represents public services, collections, technical services, IT, and digital scholarship.

**8. Please offer any additional comments on how and why your library's decision to join HathiTrust was made.**

- Although there won't necessarily be immediate, direct benefits to our library users, we very much wanted to support the work being done by HathiTrust and we wanted to have voice -- however small -- in this work.
- It was done in support of the CIC initiative to create a repository of the scanned Google materials contributed by CIC libraries.
- Initially our cost comparison for storage indicated a clear preference for HT for archive preservation. We plan to use HT as an archive platform for material
- The Dean requested a report on the costs and benefits, and made the decision to join.
- A joint Libraries effort to supplement and make more flexible the results of the Google Book Project and to avoid a Google Monopoly.
- We joined as part of a consortium, the Triangle Research Libraries Network, but would have joined individually if the other members of the consortium had not wished to join as a consortium.
- We were completing the digitization project with Google and we had no resources to provide access and long-term preservation of our digital copy. We decided a very good idea to participate in a cooperative project as HathiTrust.
- Our preservation program is newly emerging. Our University Librarian is very supportive of this effort.
- As part of the CIC, we were one of the founding members.
- We were one of the founding partners for the reasons given above.

- We joined as part of the CIC Google Partnership to manage the data produced by that project.
- 10 campus libraries agreed to join as a group to leverage completed mass digitization projects done with Google and Internet Archive.
- Part of the original CIC conceptualization and consortial commitment to instantiating the HathiTrust precursor, Shared Digital Repository (SDR).
- We are moving our collections online in a deliberate and concerted way, and acquiring member-level access to HathiTrust's corpus was a big part of what drove
- Our institution decided to participate as a consortium.
- It made very good sense that the CIC invest collectively in a long-term preservation and access solution.
- As a member of CIC, we considered this fundamental to our Google Books participation
- Saw HT as an important component of long-term digital management strategy; it is difficult for an individual library to do a good digital preservation program on its own. It is also an effective way to keep access to digital / digitized resources open.
- We joined as part of the CIC Google Partnership to manage data produced by that project.
- Our institution is a Google Partner Library and we wanted to make our digitized works from that collaboration available through HathiTrust, as well.
- We are a founding partner, and a primary motivation is to bring together collections that have been digitized across the University libraries by different

<b>9. Please estimate the total budget for your library during the last fiscal year (in American dollars).</b>	
Avg.	\$24,886,473
Median	\$26,154,573

<b>10. Please estimate your library's expenditures on third-party digital preservation services, including HathiTrust, CLOCKSS, DuraSpace, Portico, etc., in the last fiscal year. (Include only membership/data storage fees; exclude staff time spent preparing files for ingest, etc.)</b>	
Avg.	\$62,983
Median	\$39,000

<b>11. Approximately how much does your library spend annually on storing, managing, and preserving your digitized materials using your institution's local infrastructure? Include the costs of staff salaries in your estimate.</b>	
Avg.	\$292,725
Median	\$225,000

**12. Please note any functions that you could not attach a cost estimate to in the previous question, but which are important for storing, managing, and preserving your digitized materials using your institution's local infrastructure.**

- We simply do not have a handle on these costs. Some of the library's costs involve contributions to other campus entities for storage, infrastructure, which they provide for the university.
- I estimated storage costs for material digitized locally from central IT. The figure provided above may be on the low side.
- We have not tracked separate costs for digitized materials as a subcategory of all digital assets.
- Difficult to estimate answer to #14 because of difficulty separating creation and access from storing, managing, and preserving.
- [Our regional consortium's] membership (one of the benefits is the ability to store digital content in an off-site dark archive) = \$50,000
- Equipment, staff salaries, all other software
- No break out
- We currently have access to a substantial amount of free digital storage through another agency. We are not able at this time to estimate the cost of this storage when it is no longer available to us at no charge.
- We have disaster recovery services which would include digitized materials along with other services (~\$50K). We also have various staff involved in digital library project management some aspects of which are related, in part, to managing digitized materials who are not included above.
- We have significant support from our central IT organization on our data storage and data curation components. We have tried to account for this in our numbers but there is significant expertise that is not accounted for in this number.
- Expense level above does not include costs for physical plant overhead: HVAC, electricity, rent (space fees), or other central charges. Also does not include non-operational strategy planning staff expenses.
- Opportunity cost.
- Distributed content curation.
- Please note that our digital preservation/repository program is fully imbedded into a comprehensive library IT infrastructure so it is impossible to pull out these numbers in an absolute way. Activities not included: digital infrastructure development (e.g. born digital system developments), scanning/digitization lab activities, digitization project management, etc.
- Service level agreement with enterprise-level Office of the CIO and our participation in [our regional consortium] as part of our content preservation strategy
- Costs of staff and students performing digitization and metadata work.
- Computer Security.
- Indirect cost of supporting network and other institutionally provided infrastructure.

<b>13. Has your library's spending on storing, managing, and preserving your digitized materials using your institution's local infrastructure changed as a result of joining HathiTrust?</b>	
Yes	5
No, but we are planning to change our spending in the coming fiscal year as a result of joining HathiTrust	10
No, and we are NOT planning to change our spending in the coming fiscal year as a result of joining HathiTrust	31

<b>14. If you answered that your annual spending on storing, managing, and preserving your digitized content using your institution's local infrastructure has changed as a result of joining HathiTrust, has this figure increased or decreased?</b>	
Increased	2
Decreased	3

**15. By approximately what dollar amount has this annual spending changed?**

- \$150,000
- Not Available
- N/A but not a large figure
- \$50,000
- Actually, spending on local infrastructure was repurposed so it didn't increase or decrease.

**16. If you answered that your library *has* changed, or is planning to change, its annual spending on storing, managing, and preserving your digitized content using your institution's local infrastructure as a result of joining HathiTrust, please discuss which aspects of HathiTrust led your library to do so.**

- Locally storing and managing the vast amount of Google based images from our Google project would have significant costs for local hardware and managing the files. Not to mention developing access.
- We will make minor adjustments to our redundancy strategy and investment in discovery of these materials due to capabilities (and redundancy) provided by HathiTrust. As Hathi develops a more robust preservation strategy and considers additional types of content - the potential cost savings will grow from minor to significant.
- We are planning to emphasize digital preservation rather than traditional preservation methods, and we are also planning to migrate locally digitized material from local servers to HathiTrust. We plan to use HathiTrust instead of local solutions.

- Storing our digitized book and journal content with HathiTrust is easier and more cost effective than storing and preserving locally. We also save by using HathiTrust for access in addition to Internet Archive and Google.
- We are in the process of investigating whether we can contribute content to the HathiTrust. If so, we may opt for a more economical local storage option.
- Less attention on preserving books and book-like objects, and more attention on formats not yet supported in HathiTrust.
- Our institution shifted from preserving its LDC from Google with its own Preservation Repository to Hathi's repository
- Disclosure: I think the last question was poorly worded. We *may* change our spending as a result of HathiTrust. We need to conduct a requirements exercise and compare those requirements to what HathiTrust could offer in terms of preservation capabilities.
- We forecast new local cost avoidance opportunities and possible expenditure reduction as we consider shifting digital preservation and access strategies to HathiTrust that are currently supported by locally developed and managed infrastructure.
- We will be spending more money on the process of transferring content to Hathi, and (in the long run) less money on hosting and preserving that content ourselves.
- Currently has silos of digital data. HT gives us the opportunity to place digital objects in a general collections space. [We] will be discussing whether to migrate digital objects that currently reside on category specific platform to HT
- We've needed to allocate more resources to ensure that our digitized content will interoperate with HT.
- Preservation coupled with access. We had not developed discovery and access services although we did have preservation capacity.

**17. If you answered that your library *has not* changed its annual spending on storing, managing, and preserving your digitized content using your institution's local infrastructure as a result of joining HathiTrust, please describe your reasons for not doing so (including barriers to decreasing these costs).**

- We've just been able to ingest some materials into Hathi -- our experience with this partnership is relatively new. We've not yet done the thinking about future contributions or the role of Hathi in our Library's plans, though we are beginning to do so.
- There doesn't appear to be overlap between what is being preserved on local digital platforms and what would be stored in Hathi Trust.
- As a founding member of the [our state's digital library network], we are committed to continued development and support of [our own regional preservation network]. In addition, we have audiovisual, still image, and other materials we must store, manage, and preserve. Our budget reflects these commitments.
- We digitize a lot of material that the HathiTrust is not able to take at this time. Until they can take audio materials, we must continue to support our digital preservation.
- We are focusing our local efforts on supporting faculty projects with non-published materials and with special collections materials. The collective our institution's budget may be affected somewhat but we will still expect to support our institution's digital preservation activities.
- Our institution joined Hathi Trust in the Fall 2010. We will be evaluating whether we can/can't increase spending in the next fiscal year.
- Since we have recently joined we do not have immediate plans to change annual spending, but it is a possibility for the future.
- We simply have not completed our digital preservation planning on this. This decision should be considered as temporary.
- N/A, as we are not contributing content to Hathi

- We have only recently joined HathiTrust and are in the process of contributing our first content. It is too early for us to assess potential budget changes as a result of joining.
- We will not change because the resources spent on storing, managing, and preserving are also used for other purposes
- Implications of having copies of local digital content in HT are not well known enough at this point to impact spending.
- We see our work with HathiTrust running in parallel with our local preservation efforts.
- We are not contributing enough content yet to allow us to change our local practices.
- We have always planned on a tiered storage infrastructure and to feed preservation quality content content to a third party entity or cloud service.
- We are just starting to work on shared print/digital storage strategies.
- Locally digitized materials are maintained in our local depository. We are currently doing a pilot to add some of those files to Hathi by having them picked up as a digifeed by Google and then having Google make them available to Hathi to pick up automatically. Once we have a process that works we will evaluate what to add to Hathi from past projects and what current or new projects will produce files that are appropriate to add. As we have more experience with depositing files with Hathi we may choose to not maintain local copies but at this point we bear the responsibility for preserving master files for locally digitized materials.
- Our system offers several services which may be more cost effective for local content [...].
- It is too early for us to judge how HathiTrust membership might change our outlook on this.
- We currently spend very little on local infrastructure. We haven't tried to ingest local collections into HathiTrust so it's uncertain if it will meet our access needs. Hopefully we will be able to test this next year and then make appropriate changes.
- The Library is in the process of implementing a digital preservation repository system (Fedora) in order to preserve and manage digital objects that are not deposited in HathiTrust. Some of these objects might become candidates for deposit into HathiTrust in the future, as Hathi develops the capability to ingest additional formats. However, there will always be a need to have some form of local preservation and access infrastructure.

<b>18. How would you characterize your understanding of HathiTrust's capacity to manage digital preservation?</b>	
1-2	2
3-4	7
5-6	14
7-8	13
9-10	10
Total Responses	46
Avg.	6.46
Median	6.5

**19. How well do the following statements describe your library's perspective? (1 = strongly disagree; 10 = strongly agree)**

	The quality of scans presented in HathiTrust is generally high.	HathiTrust is well positioned to migrate the content that my library has contributed to new formats as the need arises.	Compared with my library working on its own, HathiTrust would be more efficiently and effectively able to migrate the content that my library has contributed to new formats.	The HathiTrust public user interface meets all the needs of our users.	We are satisfied with the prominence and placement of our library's branding on individual items that we have contributed to HathiTrust.	We are satisfied with the prominence and placement of our organization's branding on collections that our librarians and faculty have built in HathiTrust using the Collection Builder tool.	We believe that HathiTrust's committee-driven governance structure will scale as new institutions join.
1-2	0	2	2	4	5	5	1
3-4	4	1	3	14	1	2	8
5-6	12	14	11	16	13	16	21
7-8	22	14	12	9	4	5	10
9-10	6	12	14	0	7	1	3
Total Responses	45	43	42	43	30	29	43
Avg.	6.91	6.98	7.12	4.74	5.63	4.79	5.63
Median	7	7	8	5	5	5	6

**20. Please feel free to add any comments on the functionality of HathiTrust's public user interface here.**

- We're very new members and have not contributed content, so most of the above questions don't apply and/or we haven't yet formed an opinion.
- The answer to the Collection Builder tool is more accurately characterized as to-be-determined as we recently added content and have not seen much activity from local users in this area.
- The locally developed default interface is superior to the OCLC WCL interface. The proposed SAB services committee can help further develop the locally developed interface.
- Comments from some of our librarians: There really needs to be a better way to see all the metadata from the results page or even the item page. I found one title that looked interesting, but no publisher was listed. It took several minutes of searching before I discovered this was a dissertation (I still do not know from which university). While there are a number of impressive titles in the Hathi collection that would appeal to historians, there is also a great deal of material that is widely available in other formats. There needs to be better options (or more obvious options) to sort a result list into something more meaningful. There MUST BE a "Search Tips" link on the Advanced Catalog Search page! My problems have to do with the scanning of serials, in that if several issues of a serial have been bound into one volume, that one volume is treated like a monograph book and it can be very difficult to find individual issues within that volume. If for serials, the Jump to Section could go to the beginning of each individual issue that would make the serial publications much more usable. The thumbnail view sort of addresses that problem, but takes a long time to load. The drop down box Jump to Section that used to be the table of contents is hard to find, especially for those familiar with the old interface.
- Our patrons are still confused with the holdings of the serial titles.
- Our Institution has not contributed content and cannot answer questions related to branding.
- We have not contributed content yet, so could not answer all questions.
- Very poor marks for displaying page numbers rather than snippets for post 1923 holdings, i.e. non-public domain.
- I'm sorry to provide 5 for every response. Given that we are one of the newest members and we have not contributed content yet, these questions are not relevant at the moment. N/A would have been a more appropriate response. There has been some local discussion about extensions to the HathiTrust user interface but nothing that would prompt me to make a specific rating or comment.
- Since we are just in the process of contributing our first content, it is too early to assess branding.
- Would like previewing and enhanced subject browsing and navigation capability. These are aspects that we are building into our local discovery services.
- There are very few cases where the page images were not quite good in cases where illustrations, photos etc are of lesser quality, it might be the originals that are the culprit. My co-worker probably has more to add, but I have to say that generally I'm quite pleased with the scans in HathiTrust. The interface for search is quite good, I think. I have just a couple of criticisms of other aspects of the interface. However, there is one glaring lack in the interface, from my perspective -- there is no link in the catalog record view or full file view that easily leads you back to your list of search results. Sometimes hitting the back button enough times gets you back to your search results, but in some cases not. Particularly if you have logged in and then try to go back, IE gets stuck at the authentication page. I find myself executing the same search multiple times when I lose my search results. And, while I understand the reason it is done, when you download an entire PDF (allowed after

authenticating by (shibboleth) you are often greeted with a delay in downloading of 30 seconds. Our users may not be quite as patient as I am about this, and I'm generally not an extremely patient person. Aside from those issues, the interface is intuitive and response time is good. The only thing that I can think of that would make branding stronger would be if our visual identity marks could be used at least in some places where branding occurs.

- Right now, all branding is simple text, no logos or other identifying marks. At this point in time it is important to make the institutional branding more prominent in order to help us to justify participation. At first I wasn't sure what they [the survey authors] were talking about, but I think they are referring to the text in the grey bar at the bottom of the page when you are looking at collections... It's pretty subtle. Potentially the structure can work. Providing APIs and other methods of access to the information is at least equally important. Also, the Hathi public user interface should allow links back to the print copies in library collections to allow ILL when desired.
- Must strive towards the ease of use of Google Book Search, while maintaining options of more sophisticated access suitable and often require in support of academic and scholarly purposes.
- Being a new member, it is too early to accurately assess our experience.
- We're not really using it yet so can't answer these questions knowledgeably
- No complaints: interface is intuitive and response time is fast.
- The interface for full text search is most in need of improvement. The metadata interface is fine and not the primary interface for our users.
- You need a N/A category in the above! We have not yet had a chance to see many of our materials in Hathi so cannot answer two of the questions above.

21. As you think about the future of HathiTrust, to what extent would you prefer that certain aspects of HathiTrust-related work be done at your local institution versus in a centralized HathiTrust organization? (1 = at my local institution; 10 = in a centralized HathiTrust organization)

	Digitization	Digital library infrastructure management	Outreach and marketing	Rights determination and management	Tool development	Structuring of fees	Research support (i.e., services that provide expertise in bringing computational tools and methods to the HathiTrust corpus to undertake specific analytical tasks)	User support
1-2	8	2	5	2	2	-	3	9
3-4	18	9	10	2	2	-	4	11
5-6	9	17	19	7	13	-	6	11
7-8	8	11	9	24	22	-	19	13
9-10	1	5	2	12	8	-	14	3
Total Responses	45	43	44	46	46	-	46	46
<b>Avg.</b>	4.51	5.77	5.11	7.20	7.11	-	7.04	5.11
<b>Median</b>	4	6	5	7.5	8	-	8	5

<b>22. Excluding staff who work on general digitization activities, approximately how many FTEs of staff time did your library devote to HathiTrust-related activities during the <i>past</i> fiscal year?</b>	
None	9
0.1 to 1.0 FTE	31
1.1 to 2.0 FTE	3
2.1 to 3.0 FTE	2
3.1 to 4.0 FTE	0
More than 4.0 FTE (please estimate if possible)	1 (Estimate N/A)

<b>23. Excluding staff who work on general digitization activities, approximately how many FTEs of staff time do you anticipate your library will devote to HathiTrust-related activities during the <i>current</i> fiscal year?</b>	
None	3
0.1 to 1.0 FTE	33
1.1 to 2.0 FTE	5
2.1 to 3.0 FTE	3
3.1 to 4.0 FTE	1
More than 4.0 FTE (please estimate if possible)	1 (Estimate N/A)

<b>24. How do you anticipate that your ability to contribute staff time to HathiTrust will change over the next three years?</b>	
We expect to increase the level of staff time that we contribute to HathiTrust.	23
We expect to maintain the level of staff time that we contribute to HathiTrust (no increase or decrease).	19
We expect to decrease the level of staff time to HathiTrust.	0
We expect to end contributions of staff time to HathiTrust.	0
We do not currently contribute staff time to HathiTrust, and we do not expect to begin doing so in the next three years.	3

## 25. How well do the following statements describe your library's perspective?

	<b>We are satisfied with the responsiveness of the HathiTrust central staff at the University of Michigan to our questions and needs.</b>	<b>We are satisfied with the division of labor between HathiTrust central staff and our library's contributed staff time.</b>	<b>We have confidence in HathiTrust's governance because the initiative is operated by libraries within a research/academic library setting.</b>	<b>We would have less confidence in HathiTrust's governance structure if the governing body included representatives from non-library organizations</b>	<b>We want greater input into decision-making for HathiTrust.</b>
1-2	0	0	0	2	4
3-4	0	0	1	9	3
5-6	5	14	6	10	17
7-8	21	11	21	17	13
9-10	18	16	18	7	8
Total Responses	44	41	46	45	45
Avg.	8.09	7.54	7.93	6.27	6.22
Median	8	8	8	7	6

**26. Please note any specific aspects of HathiTrust's governance or operations for which your library would like to have greater input or decision-making influence.**

- There needs to be more financial transparency of the costs associated and income generated by Hathi Trust for accountability purposes.
- We would like to have a method of consultation and input to Executive Committee decisions.
- Interface design priorities, branding, and user testing
- Rights management, development of non-PDF text deliverables, long-term preservation strategy, development of tools to enable integration into library services.
- A voting system of the membership for new initiatives and fees.
- Data curation
- We are interested in HT's relationships with other critical organizations working in this domain, e.g., JSTOR, CRL where we would not like to see their role diminished. In other instances, such as OCLC we have concerns about deep alliances. We wonder if KUALI offers any benefits to HT for future common needs.
- I'm currently on the Executive Committee and one of my associate deans [sits on another committee]. We are well represented in governance and operational decisions.
- Local digitization concentrates on materials that are not necessarily appropriate for Hathi. This can include archival materials, also special collections materials that would not be allowed out of house, and brittle materials that require special handling as well as institutionally important materials that may be exposed in specially selected and constructed interfaces For materials that are not appropriate for Hathi, local infrastructure is necessary. Where access permissions are an issue with archival or other materials it may not be legal to deposit with Hathi. Hathi does not support all types of materials. The library may be asked to support locally important datasets or other types of things that are not appropriate for Hathi. More local resources can be available for these specialized things if run of the mill materials are preserved via Hathi, but the specialized local files are not going to go away so local infrastructure is required. It is necessary for us to publicize digitized materials as they become available. It is unlikely our community would regularly look at Hathi for information so it needs to be pushed out via local channels. For materials deposited via Google project, and other locally digitized materials that are deposited, it would be good to have Hathi manage rights. But there are many materials like archives some media of field research, perhaps data sets where it may not be possible for Hathi to manage rights. In this case there is a question as to whether those files can or should be deposited with Hathi For materials deposited at Hathi, Hathi should develop tools that can be used by our users. The library subscribes to access to many resources. Fees should be through the library not directly to our end users. In some cases Hathi can provide research support. But in some cases it will be more appropriate for Hathi to provide the raw data and allow local work on computational tools. We have a group interested in developing such analyses and it would be too limiting for them to have to use Hathi and not be able to work on this themselves. There will be scholars working on their own research where they will build their own tools. It would be a mistake to require them to use Hathi built tools. I doubt Hathi could keep up with their needs and capabilities. What Hathi might want to do is to make it easy for them to share their work via a software depository or something like that
- None at this time.
- I may not have a current understanding of the way in which the CIC institutions have influence in the HathiTrust governance structure, so my following comment may not be completely accurate. For the most part, the CIC institutions share a core or common perspective on Hathi

development directions. At some points along the way, however, there has been frustration at our institution about the ability to have a voice at the table to influence and to have a current understanding of the real issues being decided around the standards, the technology development, and the partnerships with other organizations beyond library membership. With the current arrangement there does not seem to be a channel for more "development partners" to be in the mix beyond Michigan and CDL. We think there could be benefit to opening this up to more local, distributed development, particularly in developing out services to support ingest of more formats, specialized user applications.

- Preservation planning and strategies, research services and policies, storage requirements Shared collection development strategies.
- Governance and structure along with collection development functions.
- Strategic directions related to visualization tools and metadata enhancement.
- Development priorities and paths; more robust mechanism for collaborative development and incorporation of members' work.

**27. Please comment on any other aspect of the HathiTrust governance structure that you would like the Strategic Advisory Board to consider at the Constitutional Convention in October.**

- Financial Audit statements.
- My main concern would be to ensure a sustainable governance structure that will accommodate growth and development. I also expect complete transparency with regards to membership pricing and costs.
- As HathiTrust grows, we believe there needs to be a process (e.g., nomination and election) for member input in membership on governing committees, including Executive.
- Hathi should be governed by the institutions that have contributed content and expertise; intuitions unable to contribute in these ways should not expect to shape its direction.
- How fees are determined and how priorities are set. What factors should be considered if membership is allowed in the HathiTrust without content contribution.
- Data curation.
- If the community begins to rely on HT for collection decisions, we would like to give more input.
- None at this time.
- It's not clear to me how much role the SAB has in the governance of HathiTrust. Based on the minutes, it seems to mostly a informational committee rather than an action-oriented committee that moves the strategic mission of HathiTrust forward.
- How to ensure members' interests are fairly and fully represented without creating a burdensome and ineffective structure.

<b>28. How are HathiTrust holdings reflected in your library's general OPAC/catalog?</b>	
Each item in HathiTrust that was digitized using our local collections is indexed in the OPAC/catalog as an individual record.	5
Each item in HathiTrust for which we own a local print copy is indexed in the OPAC/catalog as an individual record.	6
Each item available for viewing in HathiTrust is indexed in the OPAC/catalog as an individual record, whether or not we own a local print copy.	4
Does not apply.	14
Other (please describe)	23

**(28, cont.) Other (please describe):**

- Too soon to know.
- Still in the process of deciding between 2 and 3 above
- Haven't decided our approach yet.
- We plan to migrate to option three to integrate Hathi availability into our discovery environment.
- Technically the item is not indexed in our OPAC. The record representing the print is indexed and there is a link to HT for the digital item.
- Items for which we own a local print copy are displayed via the HathiTrust API. Other full-view items are indexed as individual records.
- Items digitized by our institution are reflected \*only\* in OCLC/WorldCat Local
- Indexed if either (1) local print copy owned, or (2) public domain regardless of local print copy
- We plan to link out to HathiTrust content that is in our OPAC/catalog.
- HT records available in NGM
- "umlaut" search from our holdings for electronic version
- We are still working on determining the best approach for our library.
- Not yet in place but will probably implement option 3 above.
- We have installed the Hathi Trust API
- Not sure what you mean by 'individual record' Our LENS OPAC links to full text scans in Hathi whether or not they came from our collections when we own a print copy. There is not a separate bib record for the Hathi copy. Our LENS OPAC has a record with links to full text scan in Hathi for public domain items where we do not own a print copy.
- Each item in HathiTrust for which we own a local print copy is linked to (through dynamic look-up) from the catalog record of the print copy. To be clear, explicit records to the HathiTrust surrogate are not indexed in the local catalog.
- We're very recent members and are in the process of catalog migration; have not yet loaded any HathiTrust records.
- Under discussion
- Not sure what you mean by 'individual record' Our LENS OPAC links to full text scans in Hathi whether or not they came from our collections when we own a print copy. There is not a separate bib record for the Hathi copy. Our LENS OPAC has a record with links to full text scan in Hathi for public domain items where we do not own a print copy.
- We do not do anything to specifically reflect Hathi holdings

- We don't currently reflect our Hathi Trust holdings per se in our local catalog (Voyager), which is actually our secondary OPAC (our 'classic catalog'. We are currently in the process of implementing an API to link from our Voyager catalog to Google Books holdings, when available, for those items we hold in print. We have not yet determined whether to link to HathiTrust holdings as well from this secondary OPAC.
- We don't currently note any information about HathiTrust but are considering ways to highlight HathiTrust works in the catalog.
- We are using WorldCat Local, and OCLC has created records for HathiTrust items.

<b>29. How are HathiTrust holdings reflected in your library's discovery interface (e.g., Summon)?</b>	
Each item in HathiTrust that was digitized using our local collections is indexed in the discovery interface as an individual record.	2
Each item in HathiTrust for which we own a local print copy is indexed in the discovery interface as an individual record.	2
Each item in HathiTrust is indexed in the discovery interface as an individual record, whether or not we own a local print copy.	13
Does not apply.	14
Other (please describe)	19

**(29, cont.) Other (please describe):**

- Too soon to know.
- Haven't decided our approach yet.
- We are in the process of implementing a discovery interface and have not yet made this decision.
- Items digitized by our institution are reflected \*only\* in OCLC/WorldCat Local
- Indexed if either (1) local print copy owned, or (2) public domain regardless of local print copy
- Same as above
- The records are in system-wide Worldcat Local.
- Same as above.
- It doesn't apply yet
- Not yet in place but we will probably implement option 3 above
- We use EDS.
- Not yet resolved, as our discovery layer is in the process of being procured and implemented.
- Are right now in the process of implementing Primo Central, so it's too early to say.
- Rely on OCLC WCL for discovery.
- An HT subset is searchable in the Libraries' instance of WorldCat Local.
- Currently, we have included HathiTrust as a target in our locally developed federated search portal called Easy Search. It performs a search against Hathi and then provides total hits in the results list.
- Will have some access because our institution is loading to [our regional consortium's] central catalog.

- The search box on our discovery interface is currently pointed to a institution instance of OCLC's WorldCat Local. When exporting metadata to accompany the material we ship to Google for digitizing, we also send copies of relevant metadata to both HathiTrust and OCLC. Thus, our HathiTrust holdings eventually appear in WorldCat Local, which serves as our primary OPAC.

<b>30. To what extent does HathiTrust factor into your library's decision to relocate print books/monographs from on-campus shelving to off-campus storage?</b>	
We relocate print books/monographs to off-campus storage only if we know that a digitized copy is held in HathiTrust, whether or not it is immediately available for viewing by our users.	0
We relocate print books/monographs to off-campus storage only if we know that a digitized copy will be immediately available for viewing by our users in HathiTrust (i.e., public domain volumes).	0
The availability of digitized copy in HathiTrust is one factor that we consider, but not a prerequisite.	25
The availability of a digitized copy in HathiTrust is not a factor in our decision-making.	14
Does not apply.	7

<b>31. To what extent does HathiTrust play a role in your decision to dispose of print books/monographs?</b>	
We dispose of print books/monographs only if we know that a digitized copy is held in HathiTrust, whether or not it is immediately available for viewing by our users.	0
We dispose of print books/monographs only if we know that a digitized copy will be immediately available for viewing by our users in HathiTrust (i.e., public domain volumes).	0
The availability of digitized copy in HathiTrust is one factor that we consider, but not a prerequisite.	24
The availability of a digitized copy in HathiTrust is not a factor in our decision-making.	16
Does not apply.	6

<b>32. What direct impact, if any, has HathiTrust had on the rate at which your library relocates print books/monographs to off-campus storage?</b>	
We have significantly increased the rate at which we move print books/monographs to off-campus storage because of the availability of digitized copies in HathiTrust.	2
We have maintained the rate at which we move print books/monographs to off-campus storage because of the availability of digitized copies in HathiTrust.	0
We have significantly decreased the rate at which we move print books/monographs to off-site storage because of the availability of digitized copies in HathiTrust.	0
No direct impact.	36
Does not apply.	7

<b>33. What direct impact, if any, has the availability of digitized books/monographs in HathiTrust had on the rate at which your library disposes of print books/monographs?</b>	
We have significantly increased the rate at which we dispose of print books/monographs.	1
We have maintained the rate at which we dispose of print books/monographs.	5
We have significantly decreased the rate at which we dispose of print books/monographs.	0
HathiTrust has had no direct impact.	31
Does not apply.	6

<b>34. Consider scans that have been generated from your local collections and contributed to HathiTrust, whether the digitization was done by a third party (e.g., Google) or by your own staff. Does your library maintain local copies of those files?</b>	
We generally do not maintain local copies of these files.	10
We would always or nearly always maintain local copies of these files, primarily for risk management purposes.	7
We would always or nearly always maintain local copies of these files, primarily for a reason other than risk management. (Please describe that reason here.)	6
We take some other approach to our decision-making. (Please describe that approach here.)	7
Does not apply: No files generated from our content are included in HathiTrust.	12

**35. We would always or nearly always maintain local copies of these files, primarily for a reason other than risk management. (Please describe that reason here.)**

- Risk mgt and also to participate in local cross-collection searching across campus cultural institutions
- We have shared storage facilities for the system's Libraries.
- Alternative/multiple applications of content, as well as risk management.
- We have not yet begun to contribute digital copies to HT. Once we begin, we will likely maintain copies on a case by case basis. For example, we will likely store digital copies of rare or scarce materials.
- 1) digital preservation and 2) research computing on Harvard corpus
- We have a primary responsibility to do so.

**36. We take some other approach to our decision-making. (Please describe that approach here.)**

- Locally digitized material: we have local copies. Google Project: only in HathiTrust
- All Google scans are hosted by HT. Locally digitized items however are hosted locally at this time.
- Hathi keeps Google and other copies
- In the future we will want to monitor. We do NOT maintain copies of Google scans. We DO maintain copies of local scans. Whether in any circumstances we would not maintain copies of local scans for specific types of materials will be a topic for monitoring and review.
- We maintain a copy in a consortial repository located at another location, for both risk management and other purposes.
- We maintain files for items digitized locally or through the Internet Archive, but not for items digitized by Google.
- We do not maintain the copies of the books digitized by Google. However, for our other digital collections in public domain, we would always maintain local copies of these files, primarily for risk management purposes.

<b>37. To what extent should development time and investment in HathiTrust be focused on improving the custom hathitrust.org interface, and to what extent should it be focused on improving the discoverability of HathiTrust content via third-party discovery services (e.g., Summon)?</b>	
1-2	4
3-4	5
5-6	9
7-8	22
9-10	6
Total Responses	46
Avg.	6.29
Median	7

<b>38. Has your library digitized (or arranged to have digitized by a third party) any book/monograph content that you have not contributed to HathiTrust?</b>	
Yes, we have some amount of digitized book/monograph content that we have not contributed to HathiTrust.	35
No. We either have not digitized any book/monograph content, OR we have contributed all of our digitized book/monograph content to HathiTrust.	10

39. If your library has some digitized local book/monograph content that you have NOT contributed to HathiTrust, why haven't you? Rank any reasons that apply, with 1 being the most important reason, 2 being the second most important reason, etc. (This question is meant to exclude licensed or purchased content. Leave blank if the question does not apply to your library.)

	Cost of HathiTrust fees	Cost of preparing files for ingest into HathiTrust	We feel that we have not cleared the appropriate intellectual property rights	We feel that the content in question would not benefit from aggregation with other HathiTrust materials.	Other (please describe below)
1-2	5	8	5	3	27
3-4	4	3	3	2	0
5-6	1	0	0	1	1
7-8	0	0	0	0	0
9-10	0	0	0	0	0
Total Responses	10	11	8	6	28
Avg.	2.7	2	2.375	2.5	1.25
Median	2.5	2	2	2.5	1

(39, cont.) If you ranked "Other" in the previous question, please list that reason here.

- Just haven't gotten to this work yet.
- Need for cataloging prior to deposit in Hathi Trust.
- We do intend to migrate some or all of our local digitized files but we have not as yet worked through any of the issues
- We have just begun to work on the specifications for submitting these materials and intend to submit them in future.
- Limited resources to devote to preparing the files for ingest; we intend to contribute content when some other high priority initiatives are completed.
- We have some materials that are unique, but probably not enough yet to make it worth the while of either party. Our unique materials are in digitized audio files.
- Contributed items to help HathiTrust develop ingest workflow, but not yet ready to contribute in production (planned for Summer or Fall 2011)
- We have recently joined, so have not investigated supplying content to date. We do believe we have files that we could contribute at a future date.
- We are very new to the Trust and have not yet sorted out how we will integrate the Trust into our strategies.
- CJK or other special materials done by foreign partners and whose eventual access point is yet to be determined.
- Hathi did not feel our volume of content merited the effort to load. We would be happy to contribute.

- Our in-house digitization of brittle books and special collections began prior to our joining HathiTrust.
- There have been local discussions about possibly depositing copies of these materials into HathiTrust though no formal assessment has occurred.
- We have not yet established a workflow for ingesting directly from our local servers to Hathi.
- Our focus has been on our internal preservation interface.
- lack of staffing resources and the need for more time to develop workflows, policies, etc. (since we are relatively new members)
- We haven't investigated what resources we have to contribute yet or the undertaken any cost analysis.
- We haven't addressed the what's local/what's hosted policy question.
- To 38 we want to comment: Interoperability is of primary importance. We are currently evaluation EDS, not Summon. But interfaces can change so interoperability, the APIs provided are critical. We are currently doing a pilot project of locally digitized materials that we have uploaded to Google, and then Hathi can download. Hathi needs to make changes to identify these as locally digitized, not part of the Google project. We have some OCA records in Internet Archive that we have not yet arranged to have Hathi add. We have other locally digitized materials we have not yet loaded. We are working on having an efficient way to add locally scanned files to Hathi where appropriate as we set up local digitization projects We are in the process of doing a pilot project to load locally digitized brittle books. Currently we are waiting for Google to improve its algorithms that identify public domain items and for Hathi to be able to identify these materials as locally digitized, not digitized by Google. We have not yet gone through a definition of the load of the OCA records from Internet Archive. Originally there was some question about the available metadata as the bib records were for the microform copies. Staff time to do the back and forth to get materials loaded has been the major hold up. Also, we are testing a way to just give Google PDFs so we don't have to bundle up a lot of files to send to Hathi. Time for preparing files for submission is a stumbling block so we are looking for the most efficient way to do that before we decide to add some of our previously scanned materials and to automatically add some types of materials in future projects.
- We have not yet developed a contribution criteria.
- Uncertain how to get local material into HathiTrust.
- Eventual benefits and strategic advantages of HathiTrust aren't yet clear, a better picture expected after the governance meetings in the fall; local decision-making structures also in transition.
- We haven't made contributing such content a high priority, and we don't have that many items.
- We are in the process of doing a pilot project to load locally digitized brittle books. Currently we are waiting for Google to improve its algorithms that identify public domain items and for Hathi to be able to identify these materials as locally digitized, not digitized by Google. We have not yet gone through a definition of the load of the OCA records from Internet Archive. Originally there was some question about the available metadata as the bib records were for the microform copies. Staff time to do the back and forth to get materials loaded has been the major hold up. Also, we are testing a way to just give Google PDFs so we don't have to bundle up a lot of files to send to Hathi. Time for preparing files for submission is a stumbling block so we are looking for the most efficient way to do that before we decide to add some of our previously scanned materials and to automatically add some types of materials in future projects.

- We have spent significant effort to build our own infrastructure for locally digitized works and have that workflow in place. We haven't decided to deposit those rare/unique materials elsewhere.

**40. If your library holds digitized book/monograph content that you have NOT contributed to HathiTrust, please list any other platforms or online repositories to which your library has contributed that content for preservation purposes. (This question is meant to exclude licensed or purchased eBook content.) Leave blank if the question does not apply to your library.**

- Local campus wide pilot.
- Internet Archive LOCKKS
- Our local CONTENTdm installation.
- All are preserved locally.
- We deposit our scanned materials into our institutional repository which is a staging area for preservation purposes.
- Our institution's Digital Repository.
- 'bePress' Local SANS storage for archival copies of what is accessible via CONTENTdm
- LOCKSS Portico
- Our institution's [local system] & Internet Archive
- Local digital repositories & University Digital Conservancy and UMedia Archive, related preservation infrastructure.
- ARTstor Shared Shelf
- Our institution's [local system]
- CONTENTdm; local DSpace; local Fedora repositories.
- Local repository [at our institution]: 123TB of digital content
- We have placed monograph content in our own institutional repository and in the Internet Archive
- Internet Archive
- DLXS, DSpace, Internet Archive
- We add digitized works to our in-house repository and just implemented a new institutional repository.

**41. If your library holds digitized book/monograph content that you have NOT contributed to HathiTrust, please list any other platforms or online repositories to which you have contributed in order to provide access (i.e., not primarily for preservation). (This question is meant to exclude licensed or purchased eBook content.) Leave blank if the question does not apply to your library.**

- Local campus wide pilot.
- Our digital repository (DSpace)
- Internet Archive
- Our local CONTENTdm installation.
- DRAM.

- Internet Archive
- We deposit our scanned materials into our institutional repository for direct access.
- Google Book Search.
- CONTENTdm, bePress, Mountain West Digital Library
- Internet Archives OAI-enabled for some digitized collections BePress Proquest Digital Dissertations
- Local ETD, Local DPUBS, ContentDM, Internet Archive
- Our local catalogs and digital project pages. LC American Memory project. OAIster.
- Our institution's Merritt System & Internet Archive
- Local digital repositories & University Digital Conservancy and UMedia Archive, related preservation infrastructure.
- ARTstor Shared Shelf
- eScholarship Locally hosted Omeka instance
- Internet Archive; local web-based discovers systems--e.g., IllinoisHarvest portal
- Local digital repository
- Institutional repository,[regional consortium's repository] and Internet Archive
- Internet Archive
- Some of this content is locally hosted.
- Internet Archive
- Our local catalogs and digital project pages. LC American Memory project. OAIster.
- Our locally created repositories serve to manage both preservation and access.

**42. Does your library contribute digitized versions of non-book content types to any access platforms or online repositories external to your institution? If so, to which platforms or online repositories do you contribute these? (Include HathiTrust in your response if applicable.)**

*Audio recordings* (6 responses)

- No
- DRAM
- iTunesU
- Mountain West Digital Library
- ContentDM
- Local Digital Repository (DRS)

**(42, cont.)**

*Born-digital text publications (9 responses)*

- ETD's to DuraCloud
- No
- eScholarship articles, technical reports, born digital journals, dissertations and theses
- Online Library of California
- Internet Archive (webpage publications)
- BePress
- BePress and ETDs to ProQuest
- Local Digital Repository (DRS)
- Institutional repository

*Government documents (5 responses)*

- No
- HathiTrust, Google Books
- bePress
- Internet Archive
- Google, HathiTrust

*Images of maps and other large-format visual media (6 responses)*

- No
- Mountain West Digital Library
- ARTstor, HathiTrust
- Local Digital Repository DRS
- OhioLINK
- ARTstor

**(42, cont.)**

*Images of photographs, paintings, and other regular-format visual media (17 responses)*

- ARTstor
- No
- ARTstor, Flickr
- Online Archive of California
- NIU SE Asian Digital Library
- ARTstor
- Online Library of California
- Mountain West Digital Library
- ContentDM
- [Local repository]
- ArtSTOR, HathiTrust
- ArtStor
- Local Digital Repository (DRS)
- OhioLINK
- CONTENTdm [hosted by regional consortium]
- Flickr
- Online Archive of California

*Manuscripts (9 responses)*

- A few small projects.
- Online Archive of California
- e-Scholarship
- HathiTrust, Google Books
- Digital Scriptorium
- Mountain West Digital Library
- Local Digital Repository (DRS)
- OhioLINK
- Online Archive of California

**(42, cont.)**

*Moving images / video* (6 responses)

- No
- iTunesU, YouTube
- Mountain West Digital Library
- ContentDM
- Institutional Repository
- Kaltura

*Newspapers* (5 responses)

- Internet Archive
- California Digital Newspaper Project
- Mountain West Digital Library and Utah Digital Newspapers
- ActivePaper/Olive software
- National Digital Newspaper Program

*Serials* (7 responses)

- A few small projects
- Internet Archive
- HathiTrust, Google Books
- bePress
- Local Digital Repository (DRS)
- institutional repository [and regional consortium's repository]
- e-Scholarship

**(42, cont.)***Other special collections (9 responses)*

- No
- Webpages in web archiving service
- Accessible Archives, Alexander Street Press, American Memory Project (LC)
- Spencer Sheet Music Collection / Sheet Music Consortium
- Online Library of California
- Incunabula: HathiTrust
- EAD finding aids via Northwest Digital Archive
- [Our local repository]
- IR, [our regional consortium's repository]

**43. If HathiTrust were able to accept a wider range of content types, how likely would your library be to contribute your digitized content in the following formats? (1 = very unlikely; 10 = very likely)**

	Audio recordings	Born-digital text publications	Images of maps and other large-format visual media	Images of photographs, paintings, etc.	Manuscripts	Moving images / video	Newspapers	Serials	Other special collections
1-2	5	3	5	3	4	5	6	-	4
3-4	7	5	3	5	4	2	7	-	4
5-6	13	8	10	8	14	17	11	-	12
7-8	9	17	17	12	13	13	12	-	15
9-10	8	11	10	14	8	5	8	-	7
Total Responses	42	44	45	42	43	42	44	-	42
Avg.	5.86	6.86	6.93	6.56	6.23	6	5.98	-	6.31
Median	5.5	7.5	7.5	7	6	6	6	-	7

**44. Are there other format types (other than those listed above) that you would like to be able to contribute to HathiTrust, but cannot currently?**

- Decisions on all formats would depend on the costs of storing and preserving through HathiTrust vs doing so through our local repository.
- Data-sets
- Mission creep is to be avoided until access to bound materials has proven to be robust.
- Government documents; serials
- Scientific data
- Serials=7
- No
- Data-sets
- Data-sets
- Answer to all above is "unknown" at this time. Local repository service is well developed and HathiTrust would need to be as robust for us to consider a change in strategy for some subset of content. This is a longer term question in my mind.
- Data-sets
- No
- No
- No

<b>45. A small number of publishers currently display in-copyright titles in HathiTrust. To what extent would your library value HathiTrust as a platform for purchasing access to publisher-owned content? (1 = value little; 10 = value highly)</b>	
1-2	4
3-4	12
5-6	13
7-8	10
9-10	7
Total Responses	46
Avg.	5.65
Median	6

**46. Do you have any other comments or suggestions about the existing suite of services offered by HathiTrust?**

- We are increasingly dependent on HathiTrust as a preservation service and continue to see this as the most important service offered.
- I have no idea if HathiTrust is well positioned to migrate the content that my library has contributed to new formats as the need arises and would assume that very few people actually know enough to assess that critically beyond what is stated in marketing brochures.
- Your efforts to identify more material for the public domain is a hugely important service.
- No.
- Continue to obtain access to more materials by clearing copyrights!
- We don't think having HathiTrust sell publisher-owned content should be a priority. There are other sources, it can get complicated and HathiTrust has an advantage over something like Google Books in that it is NOT involved in selling content and does not compete with profit making companies that do. We would like a statement that the locally digitized files that we deposit will receive the same level of preservation as the Google files we are depositing. We think this would be so, but we would like a real commitment.
- Lower-barrier ingest: 1) HathiTrust requires very sophisticated submission packages for ingest; and 2) We would not be able to ingest into HathiTrust our digital master files in their native formats; rather, we would need to convert them to the HT standard master file format. Expensive, at least on the front end of preservation.
- As a new participant, we don't yet understand the existing services well enough to make suggestions.
- More emphasis should be placed on either enabling better discovery by third party services (locally chosen); or providing access to Hathi-specific unique services to support research.
- In early learning phase -- no comment.
- We don't think having HathiTrust sell publisher-owned content should be a priority. There are other sources, it can get complicated and HathiTrust has an advantage over something like Google Books in that it is NOT involved in selling content and does not compete with profit making companies that do. We would like a statement that the locally digitized files that we deposit will receive the same level of preservation as the Google files we are depositing. We think this would be so, but we would like a real commitment.
- Before adding other content types, there needs to be broader discussion of policies regarding rights restricted content and evaluation of access/usage functions for special formats.

**47. Are there any other services that you would like HathiTrust to develop?**

- A way to load individual records into the local online catalog.
- Encourage work with middleware like Shibboleth, Gouper and COmanage.
- Publishing platform, Print on Demand, Mobile/e-book reader functionality.
- Right now, it seems like we have to submit all our IP clearances to a Hathi staff member; it could be useful to eventually have an interface where Hathi participants can update IP status of their own titles directly.
- Ability to create virtual curated collections by combining materials deposited by different libraries. For instance, a digital library of agriculture or classics.
- More assessment surrounding the needs of scholars for services to make digitized books easier to incorporate into their work habits and how HathiTrust compares to other sources of digitized books would be timely and could reveal areas for service development.

**48. Are there any topics that you would particularly like to see addressed or discussed in depth at the HathiTrust Constitutional Convention in October?**

- None beyond what is already being advanced.
- Long-term content curation - planning, prioritization, and strategy.
- How partnering institutions can assist in rights determination and management. A roadmap for including other format types, including born-digital theses and dissertations.
- Long-term content curation - planning, prioritization, and strategy.
- Strategic direction over the next 3-5 years
- Hathi can become a vital service for libraries moving to largely digital collections. But branding for local consumption will be very important as local libraries still need to 'get credit' for what they have digitized. In our institution I would want our users to see our brand on every title contributed by our institution whether the text displayed actually came from our sources. If that is true we can promote Hathi as our own service expanded and enriched by other research libraries. If it remains primarily branded by whichever institutions contributed the display copy, it will not give us the promotional edge we need.
- Future plans.
- Shared storage infrastructure.
- What is the distinction between HathiTrust and the future 'Digital Public Library of America?' What, if any, plans are there to improve branding of materials contributed by libraries?
- Copyright clearance infrastructure as a community-based activity where local institutions have the capability and authority to determine rights and contribute to the overall service. Relationships with other critical organizations in the library landscape and how determinations will be made about their involvement. How to ensure alignment of interests w/ research library goals?
- A statement of the level of commitment to preservation of a library's locally digitized files if they are deposited in Hathi.
- How to modify the "constitution" as more libraries join the Trust.

- Not at this time.
- Cost sharing Relationship to distributed print repository developments.
- A major concern is establishing a governance structure that is both representative and agile and that focuses on the organization of the large-scale cooperative activities that are not/can not be adequately addressed by individual libraries or through other organizations. the nature of its governance structure and the broad sweep of its mandate will significantly affect the intensity of our participation.
- Communications about the HT should be cast to a wider audience among participating institutions.
- Despite its efficiencies, centralization decreases the ability for HathiTrust partner institutions to develop, contribute, and directly manage new services, which in turn serves as a constraint on HathiTrust's potential. We would like to see HathiTrust move to a more distributed model of development and service delivery. The latent capacity in the growing list of institutions who see it in their interests to be part of this alliance is enormous, but this remarkable potential will languish unless organizational and logistical questions about how to distribute the technical work of HathiTrust are addressed.

**49. Do you have any other comments for the Strategic Advisory Board on the strategic direction of HathiTrust over the next 3-5 years? Please enter any comments here.**

- Focus on integrating HathiTrust content into discovery environments via web services and api development.
- Let's not get bogged down in issues like DPLA, everything for everyone and then nothing gets done.
- Inclusion of the partners in the public service is a strategic question. Hathi can become a vital service for libraries moving to largely digital collections. But branding for local consumption will be very important as local libraries still need to 'get credit' for what they have digitized. In our institution I would want our users to see the our brand on every title contributed by our institution whether the text displayed actually came from our sources. If that is true we can promote Hathi as our own service expanded and enriched by other research libraries. If it remains primarily branded by whichever institutions contributed the display copy, it will not give us the promotional edge we need.
- Building reliable access that is simple and visually attractive in both interface and the content of the archive is the prerequisite to any other services or programs. Get the basics right.
- This whole questionnaire has been designed from an "insider" perspective. I am hoping to learn more about the thinking behind strategic issues and choices to date before I would jump in.
- Establish system for direct member input into initiatives being considered and in regard to fee structure.
- Should HathiTrust be working cooperatively with Digital Public Library of America?
- Quality of scans and verification of integrity of item (missing pages problem) is increasingly important if we are to rely on HT for collection management decisions, particularly de-accessioning.
- Governance infrastructure needs to be flexible enough to address changing landscape of preservation and access. Rights to use and to access orphan works will be important to address, so need HathiTrust to offer sustainable options to allow widest possible use and access.
- Not at this time.

- HathiTrust is an invaluable preservation and access aggregation of scholarly content for libraries and their users. Focus should continue to be on the best ways to continue aggregating that content, ingesting different formats, clearing rights (with qualified community contributions), developing API's that 3rd party vendors of local choice can write to.
- Strategy should be aimed at utilization of scale, capacity, and creative collaboration to define research library agendas that may not be otherwise addressed. Thank you.