

HathiTrust Program Steering Committee's Response to the Collections Committee's *Collection Priorities Survey Analysis Final Report*

May 2016

The HathiTrust Program Steering Committee (PSC) charged the Collections Committee to survey the membership regarding current and future collection priorities. A survey to all members was conducted in the October-November 2015 timeframe, and consisted of both fixed-response (quantifiable) and open-ended (qualitative) questions. Response rate to the survey was nearly 56%. Preliminary results of the survey were reported to the Board of Governors at its December 2015 meeting. A draft report and recommendations were submitted by the Collections Committee to the Executive Director and the PSC in April 2016. The PSC conducted an initial review of the draft report, offering comments and questions about the report's findings and recommendations. Based on this feedback, the Collections Committee submitted its revised report to the PSC in May 2016. Below, the PSC offers its response to the specific recommendations, as well as some additional considerations based on the findings of the report and possible strategies moving forward. Once the Board has had the opportunity to discuss the report, findings, recommendations, and considerations, the PSC is prepared to work with the Collections Committee for further information or clarification needed, and to plan for the next steps of sharing the survey and beginning to implement its recommendations.

Before laying out its response, the PSC would like to commend the Collections Committee for generating and synthesizing this important baseline dataset regarding the membership's viewpoints on HathiTrust's central asset -- its collections. These data and the Collection Committee's accompanying recommendations also serve to inform deliberations of significant questions before the Board regarding strategic direction and investments. Finally, the PSC gratefully acknowledges the member institutions that committed the time to responding to the survey, a non-trivial effort that often involves extensive internal consulting.

PSC Response to Report Recommendations

The survey report offers three overarching recommendations informed by synthesized findings and trends evident in member responses. These recommendations and the PSC response to each follows:

Recommendation #1: Concentrate on Enhancing the Comprehensiveness of the Digitized Print Corpus

The PSC endorses this recommendation, noting the resounding affirmation by members to further strengthen HathiTrust's foundational collection strategy and scope (digitized print). Numerous signals in the survey data urge the partnership to build on HathiTrust's core and distinctive strengths rather than pivot in entirely new directions *at this time* (e.g., non-text formats). There is concern that such investments (e.g., in support of images, audio, video) might come at the expense of delivering a higher quality, more complete digitized print corpus.

Within these signals, there is expression of need to advance HathiTrust's collection development strategies and efforts. In striving to provide valuable service to HathiTrust's user communities, what are the critical collection gaps that need to be addressed in multi-volume sets or series, document types (e.g., theses and dissertations), subject and geographic areas, languages, etc.? What are the barriers to addressing those gaps -- collection analysis and management data, sourcing options, contributor ingest processes? These are among some key questions that the PSC and associated committees and groups, in particular the Collections Committee and the Metadata Policy, Strategy, Use and Sharing Advisory Group, working in coordination with HathiTrust Operations, are well-poised to address.

Recommendation #2: Improve the Quality of the Corpus

The PSC endorses this recommendation, which primarily focuses on the need to remediate quality issues at the surrogate item level, as well as with the metadata that supports discovery, access, citation, and reporting activities. Item level issues of highest concern tend to concentrate around digitization errors and omissions (i.e., legibility, missing pages/foldouts, “rejected” items), which not only pose issues to researchers and other users, but also to the quality and completeness of works in the digitized corpus when, for example, making print retention decisions based on the existence of a digital copy..

This recommendation affirms the timeliness and importance of a number of “quality improvement” activities already set into motion by the PSC and HathiTrust Operations. These include the soon-to-be-released revision of the HathiTrust Commitment to Quality statement, which gives attention to stakeholders, common use cases, issues, and strategies regarding quality in the corpus. In the past year, HathiTrust has benefitted from the creation and documentation of procedures for replacing pages in Google-scanned books in HathiTrust by members of ad hoc quality group. This group is in the process of being formalized into a HathiTrust Working Group by the PSC to extend its work into additional prioritized areas of attention. Finally, the Zephyr Operations team continues to make improvements to metadata ingest and correction processes. The PSC is uncertain at this point whether quality improvement initiatives, beyond what is briefly mentioned here, will require further investment. If that is the case, it will work with the Executive Director to propose activities, as appropriate.

Recommendation #3: Improve Member Services

The PSC is appreciative that the survey surfaced concerns about barriers to access (e.g., related to print disabilities) and the lack of collection management, analysis, and reporting tools that are seen as instrumental in aiding collection development activities at the corpus-level, as well as decision-making at the local collection-level relative to print retention, storage, etc. Recent and welcome news of HathiTrust’s policy change to allow access to all copyrighted material in HathiTrust on behalf of members’ users with print disabilities is a significant lowering of barriers. Specification and development of a HathiTrust analysis and reporting function for members, which the PSC will deliberate among other proposed priorities, is also a prospect with potential high-impact for current and prospective future members. The opportunity for further consideration of this will likely come into play with the advancement of the shared print program and operations.

PSC Further Considerations

In a sense, the survey suggests what might be a “2015 response” to ongoing questions and considerations about whether to expand the HathiTrust corpus to include other formats. Such interests and notions date back to the early years of HathiTrust visioning and planning. The survey response here indicates fairly tepid level of support at this point in time for such expansion relative to other identified needs. However, it is worth noting that there was increased interest in expanding the corpus to include archival and manuscript materials, born digital monographs, and other contemporary text formats (born digital), and maps.

While investments in format expansion may not be the right strategy at this time, the PSC recommends *exploring* an incremental expansion of collecting scope from “print” to “text.” As the boundaries between “book” and other textual formats increasingly blur and the interests in text for research purposes rises (note, HTRC), excluding text due to non-print or non-published origins may become an increasingly artificial distinction. If it were to proceed, such exploration would need to include all aspects of costs, risks, and benefits,

Next Steps

Upon review and discussion of this report and its recommendations with the Board, the PSC proposes to take these next steps:

1. Publish the report on the HathiTrust web site, also distributing it to all members along with planned next steps.
2. Charge the HathiTrust Collections Committee to:
 - a. Develop a plan and proposal to strengthen comprehensiveness of the digitized print corpus.
 - b. Investigate an incremental expansion of collecting scope from “print” to “text,” and issue a discussion paper that identifies findings and considerations related to stakeholder interests, opportunities, benefits, risks, and costs.
3. Publish the revised HathiTrust *Commitment to Quality* statement.
4. Charge and launch the HathiTrust Quality Working Group to address prioritized quality issues within the corpus.
5. Coordinate efforts to specify and propose a HathiTrust analysis and reporting function providing services to members (coordination would involve Collections Committee; Shared Print Initiatives; Metadata Policy, Strategy, Use and Sharing Advisory Group; Zephir, others).