Against the Grain

2016

Curating Collective Collections--HathiTrust Launches Its Shared Print Monograph Archive

Thomas H. Teper
University of Illinois, Urbana-Champaign, tteper@illinois.edu

Bob Kieft
University of Alabama Libraries, rhkrdgzin@gmail.com

Follow this and additional works at: https://docs.lib.purdue.edu/atg

Part of the Library and Information Science Commons

Recommended Citation

Curating Collective Collections — HathiTrust Launches Its Shared Print Monograph Archive

by **Thomas H. Teper** (Associate University Librarian for Collections and Technical Services, University of Illinois at Urbana-Champaign) <tteper@illinois.edu>

Column Editor: **Bob Kieft** (688 Holly Ave., Unit 4, St. Paul, MN 55104) <rhkrdgzin@gmail.com>

Column Editor's Note: HathiTrust's *print monograph repository project is the largest in geographical scope of any shared collections project to date and has the potential over the next few years to become the largest in terms of volumes committed to retain. Because of its sheer size, the roles* ***Hathi*** *members play in state and regional resource-sharing partnerships, and the efficiencies and technologies needed to declare so many retention commitments, the repository will have substantial influence on the development of other shared print programs, the development of a national structure for print access and preservation, and the services* ***OCLC*** *offers their members in support of shared print agreements. My guest columnist this issue,* ***Tom Teper****, was Chair of the task force that created the shared monograph plan; he and the other members of the task force,* ***Clement Guthro****,* ***Erik T. Mitchell****,* ***Jacob Nadal****,* ***Jo Anne Newyear Ramirez****,* ***Matthew Revitt****,* ***Matthew Sheehy****,* ***Emily Stambaugh****,* ***Karla Strieb****, and your column editor met by phone and in person over the course of several months and consulted with* ***HathiTrust*** *Program Steering Committee Chair* ***Bob Wolven****,* ***Hathi*** *members in October 2014, and* ***Hathi*** *Executive Director* ***Mike Furlough*** *in the course of their deliberations.* ***Tom*** *reported on the* ***Hathi*** *project at the* ***Print Archive Network Forum*** *in January and June 2015, and I am grateful to him for taking keyboard in hand and offering the readers of* ***ATG*** *this account of the project's main features. —* ***BK***

Recent decades have witnessed significant, even transformational changes in research libraries. **Chris Ferguson** has characterized these changes as a "massive and rapid shift from print to digital information resources, from on-site services to virtual services through the network, from an emphasis on our values and visions to those of others…."[1] Longtime independent operators that collaborated when good and necessary in a world of physical materials, research libraries' operations and collections are now beginning to mimic the networked environment of the virtual world in which they house their collections and offer service.

At my own library, the **University of Illinois at Urbana-Champaign**, spending continues to skew in the direction of electronic delivery, with the total materials allocation being 75% spent on electronic resources. This commitment certainly carries implications for staffing and the types of services offered, but it also requires rethinking the management of print materials, especially those that are less used, highly duplicated, and available online. Locally and consortially, this rethinking has manifested itself in multiple efforts in the last several years to develop retention agreements for printed serial literature. The **Committee for Institutional Cooperation (CIC)**, **Scholars Trust**, and **Western Storage Trust (WEST)**, for example, are groups of libraries engaging in establishing retention commitments that enable member institutions to draw down duplicated print serial titles if they wish.

Such print retention commitments for serials focus on the volume of space that members may recover and have met with considerable success. In concentrating on serials, however, they leave vulnerable to loss far less commonly held titles, including many serials to be sure but also a range of other material types, in particular, monographs, for which distribution patterns, rates of duplication, traditional management strategies, and the prospect of space gain are very different from serials.

To address this situation and advance its stated goals of "develop[ing] "partnerships and services that ensure preservation of the materials in the **HathiTrust** and the entire print and digital scholarly world" and "reduc[ing] the long-term capital and operating costs of storage and care of print collections through redoubled efforts to coordinate shared storage strategies among libraries,"[2] the **HathiTrust's** governing board approved a ballot initiative in 2011 calling for the development of a print monographs archive. In the spring of 2014, it charged a Print Monographs Archive Planning Task Force to design a program that would serve both the **HathiTrust** membership and the broader scholarly community (*https://www.hathitrust. org/print_monographs_ archive_charge*).

Working between June 2014 and March 2015, we on the task force completed a report that summarizes recommendations with respect to the composition, phased construction, operating systems, business model, and governance of the **HathiTrust Shared Print Monograph Program (HTSPMP)**. Reviewed by the **HathiTrust's** Program Steering Committee, our report was published in June 2015 with an initial set of recommended actions and an implementation calendar of five years (*https://www.hathitrust.org/files/ sharedprintreport.pdf*).

The HathiTrust Shared Print Monograph Program

Our report and recommendations outline a cooperative program that rests on several assumptions, perhaps the most important of which are that, at least in its first five years, the print monograph repository

- will mirror the monographic holdings of **Hathi's** digital archive,
- be built from and regionally distributed amongst members' general, circulating collections,
- be housed in a combination of active shelving and high-density facilities, and
- be lendable.

Moreover, the model we outline calls for the repository to be governed, managed, developed, and financially supported by **HathiTrust**, not by a subset of members; in other words, membership in **HathiTrust** will equate to membership in the repository, although not all members will be required to hold materials on behalf of the group. Built, collaboratively stewarded, and supported by the membership as a public good for as rapid and comprehensive a standup as possible, the proposed repository is at once conservative in its goals for retention of print monographs and highly transformational in its implication for the local and collective management of collections.

Among the more challenging topics we discussed was whether the repository should rely on a risk model, that is, whether the archive should seek to acquire a threshold number of copies for each volume in order to establish a minimal level of risk for each title, or whether the repository should overthrow convention and seek some other mechanism for achieving the **HathiTrust's** goals. Adopting the risk model faced two significant impediments — scope and feasibility. First, with over 6,800,000 monographs ingested in the digital repository, a distributed repository of multiple print copies could quickly reach

numbers upwards of 18,000,000 commitments with no effective model for telling individual institutions that **Hathi** sought retention commitments for some volumes and not others. In discussing the development of such a model, the task force quickly identified a new risk — the risk of **HathiTrust** falling prey to such a severe case of analysis paralysis that it would delay development and implementation of the retained collection beyond the point where the conversation would still be relevant to the membership.

The second issue confronting the risk model centered on the issue of feasibility when it came to reaching copy-number thresholds. For example, while many volumes are widely held across member libraries, many others are held in ones and twos. This scarcity creates a long tail of material uncommonly held and can be seen across much larger bodies of member institutions. It would present **HathiTrust** with a perennial challenge: could an item held by only one member be considered preserved if there was only one retention commitment made?

In considering these and other questions, we recommended establishing an initial target of matching 50% of the digital collection of monographs, roughly 3,000,000 titles at the time the report was written. The initial build-out of the repository will not be driven, as noted, by a risk model that might focus on scarcely held titles or the establishment of a minimum number of copies but rather will develop rapidly from large commitments from a group of volunteer libraries. Most of the repository's early requirements in terms of retention commitments, materials housing standards, verification, and access or delivery mechanisms will be relatively lightweight as the repository is built. Relying as did the implementation of the digital repository on the recognition of common interests and values, low-cost voluntarism, and services currently in place among members, the retained monograph collection will therefore be able to develop quickly and more readily play the transformative role envisioned for it.

Indeed, as we concluded our work, we were keenly aware of this initiative's urgency in the minds of many in the community. Implementation of the repository program will require that some institutions balance human-resource and financial commitments to the **HathiTrust's** effort with commitments to other state and regional resource-sharing and print-retention agreements, yet we do not believe that the **HathiTrust** initiative should seek to supplant those agreements. Rather, we believe that it will serve as a complement to those initiatives that are growing at the state or consortial level.

Beginning with the release of the report for public consideration in June 2015, the **HathiTrust** Program Steering Committee and Governing Board set in motion the process of launching the monograph repository. In late 2015, **HathiTrust** began recruiting for a Program Officer to lead the project. In March 2016, **HathiTrust** announced that

Lizanne Payne, a nationally recognized expert in shared collections management, will join **HathiTrust** as the Program Officer for Shared Print Initiatives. Starting in May 2016, **Payne** will have responsibility for planning, organizing, and implementing the shared print monograph archive, a program that will have a lasting impact on the future stewardship of the print record.

The Transformative Role of the Repository Program

The task force's recommendation that the retained collection be persistent accords with the mission of the **HathiTrust** digital archive and the historic mission of many **HathiTrust** members. Moreover, the size, prominence, geographic distribution, and multiple affiliations of **HathiTrust** members and the position **Hathi** itself occupies in the library community establish the repository as a component of nationwide and potentially international programs of print retention and management. Developing the repository requires a new, broadly usable technical infrastructure that enables disclosure of retention commitments, easy discovery of those commitments, and local/consortial collection management decisions based on commitments. For these reasons, the **HathiTrust's** distributed monograph repository will help shape the shared print policies, governance structures, and financial regimes that enable the coordination of retention and archiving efforts across consortia and regions; most important, it will encourage a community-wide approach to managing the collective collection by producing a critical mass of public retention commitments, defining new preservation and collection management standards, and catalyzing enhanced service development.

From the point that it initiated its work, **HathiTrust** sought to be a leader in helping to transform how libraries manage their collections. In our recommendations, we sought to develop a program that would position **HathiTrust** and its members to lead in print retention as they did in digital preservation and service provision. Through the repository, **HathiTrust** stands poised to make further, transformative impacts on the management of libraries and the mechanisms by which they provide services to their constituents. The repository will fulfill **HathiTrust's** purpose of creating a "new paradigm by which research libraries and other academic libraries can develop shared reliance on a scholarly print record that is collaboratively stewarded and supported as a public good."[3]

Endnotes

1. **Chris Ferguson**, "Whose Vision? Whose Values?" in Reflecting on Leadership (Washington, DC: Council on Library and Information Resources, 2004): 20.

2. **HathiTrust** "Mission and Goals at *http://www.hathitrust.org/mission_goals*.

3. **HathiTrust** "Print Monograph Archive Planning Task Force Charge" at *http://www.hathitrust.org/print_monographs_archive_charge*.