

October 9, 2019

*Email from Karla Strieb, Chair of the HathiTrust Program Steering Committee*

*You are receiving this message because you are either a Library Director, Member Representative, or Primary Contact from a HathiTrust member library.*

I'm pleased to announce HathiTrust's increased commitment to sharing freely the metadata that it collects, produces, and/or manages. The HathiTrust Metadata Policy, Strategy, Use and Sharing Advisory Group (MUSAG) has created a new [HathiTrust Policy on Metadata Sharing and Use](#), and, in June of this year, the Board of Governors approved the policy. This commitment to broader sharing of metadata is significant for research, discovery, and dissemination of the scholarship under our collective stewardship. I am happy to share the Policy with you now, and to outline the steps we are taking to implement the policy and ensure HathiTrust-managed metadata is shared as freely as possible.

## About HathiTrust metadata

HathiTrust gathers and maintains — and sometimes creates or enhances — eleven types of metadata, each managed by informal policies and practices developed by members and staff as services have evolved. These types include, notably, bibliographic metadata in MARC format, rights and source metadata, print holdings and shared print retention data, METS structural metadata, data about personalized collection building in the HathiTrust applications, and market availability data.

In the absence of an overarching policy in the past, HathiTrust set up limited sharing of certain types of metadata. For example, bibliographic data has been shared in the form of “hathifiles” (with selected fields only), via Open Archives Initiative Protocol for Metadata Harvesting (with selected fields for public domain works only), and via queries to a Bibliographic Data API (complete records, but with a severe throttle on quantities). While libraries, their vendors, and scholarly researchers have made use of these forms of metadata provision with excellent results, we have been eager to provide bib data, as well as other types of HathiTrust-managed metadata, for broader uses and in more flexible formats.

## About the new Policy

In 2016, the HathiTrust Program Steering Committee (PSC) charged MUSAG to review HathiTrust-managed metadata and to recommend principles for metadata sharing. MUSAG members conducted an environmental scan of all types of metadata managed by HathiTrust and published its report in 2017 describing each type and its uses. Based on the conclusions from the scan, MUSAG set up a Metadata Sharing Policy team, which reviewed policies and practices from stakeholder organizations and other groups that manage or control metadata. The team developed an initial draft and worked with PSC to refine the draft into the current Policy on Metadata Sharing and Use.

The policy provides a broad framework for sharing and using HathiTrust-collected, -managed and -generated metadata. To encourage discovery of materials, research, and for a variety of current and future uses, the new policy states that *HathiTrust will provide open access to HathiTrust-collected, -managed and -generated metadata where it has been determined that legal, contractual, policy, ethical, practical, and strategic considerations allow.*

I appreciate the dedicated work of MUSAG members, and particularly MUSAG chairs Tim Cole (University of Illinois) and Stephen Hearn (University of Minnesota), and subteam chair Patti Martin (University of California, now retired), and thank them for leading this effort.

## Next steps

Deeper investigation into the legal, contractual, policy, ethical, practical, and strategic considerations is necessary to determine the exact scope of sharing to be provided for each type of metadata. PSC has charged a new group to develop a sub-policy specific to each type. **The Metadata Sharing Policy Task Force** will begin work this fall to increase understanding of the potential impacts and policies needed to ensure the appropriate sharing and use of a variety of HathiTrust metadata. This group will be chaired by Stephen Hearn and report to the Program Steering Committee. A call for interest will be out later this month.

Implementation will require HathiTrust staff to explore the use cases that emerge from the sub-policy development process. Then staff will assess existing HathiTrust mechanisms for metadata sharing, including the [Bibliographic API](#), the [Open Archives Initiative \(OAI\) feed](#), and the nightly updates of [Hathifiles](#), to decide how to either change those tools or create new ones for more flexible use in retrieving metadata. This development work will be planned in 2020 work cycles.

Thanks again to the participation and expertise of members to enable the ongoing growth of our collection and policies governing its use, so that together we can provide the greatest possible access and use of our collective HathiTrust resources. Please contact me at [strieb.2@osu.edu](mailto:strieb.2@osu.edu) if you have any comments or questions about this new policy.

Karla

Chair, HathiTrust Program Steering Committee  
Associate Director, Content & Access  
University Libraries, The Ohio State University  
[strieb.2@osu.edu](mailto:strieb.2@osu.edu)

## Resources

HathiTrust Website: [Metadata Sharing Policy 2019](#)  
[Metadata Sharing Policy Webinar](#), December 2018. Describes the proposed policy and outlines the work conducted by MUSAG.