

HathiTrust Copyright Review of State Government Documents

Kristina Hall

This article describes a copyright review program at HathiTrust which sought to determine the copyright status of individual US state government documents. The copyright project was carried out from 2014 to November 2019, and more than thirty-five people from different institutions took part.

US state government documents are subject to copyright protection despite the belief they should be public domain just as US federal documents are public domain. 17 USC 105 only states that copyright protection is not available for any work of the United States Government.¹ State publications are not directly addressed within US federal copyright law and therefore receive copyright as a default. This causes uncertainty for libraries wanting to digitize and distribute state publications online.

There has been some effective advocacy on state copyright in the last five years including the State Copyright Resource Center from Harvard and Free State Government Information (FSGI).² The push of information about state copyright has generated interest among rights holders, and HathiTrust has seen an increase in Creative Commons licensing particularly around state agricultural publications.³ Although US state publications by nature contain local and region-specific information, people frequently contact HathiTrust asking for them to be opened.

HathiTrust found that we can open these state publications the most effectively through copyright review.

US State Government Copyright Review

HathiTrust reviewed over 119,000 state publications and found over 101,000 of them to be public domain. The public domain publications are now open for US readers. All states were represented in our work including Alaska and Hawaii, although some states were better represented than others. The higher amounts in some states like California, New York, Michigan, Wisconsin and Illinois correspond to the locations of institutions that

deposited more digital content with HathiTrust. Some state documents were digitized from library collections outside the home state of publication.

Between 1923 and 1977, copyright notice was a required formality in the US.⁴ Many state publications during this time were published without copyright notice and for that reason became public domain. This allowed HathiTrust to open valuable and unique information for researchers, for example state banking reports and agricultural research bulletins. In fact, 1,908 volumes were in response to a user asking that a specific item be opened.

Results and Analysis

Project statistics as of November 27, 2019

- 119,918 in total
- 101,613 public domain in the United States
- 16,768 needs further inquiry to determine
- 1,537 in copyright (1964–1977 with copyright notice)

Over 101,000 volumes were public domain in the United States because they were published without copyright notice. About 18,000 volumes had to remain closed, of these only 3,258 volumes were tagged by reviewers as having copyright notice. Publication with copyright notice holds different significance based on the date. Between 1964 and 1977, renewal was no longer a required formality so the works published during those dates would be in copyright, which for us was a subset of just 1,537. Before 1964, publications with notice could yet be public domain if the copyright was not renewed with the US Copyright Office. We chose not to do the additional research into renewals because of workflow considerations.

The majority of works that would require further inquiry were revised editions or reprints where it was not possible to verify that the original publication lacked copyright notice. The

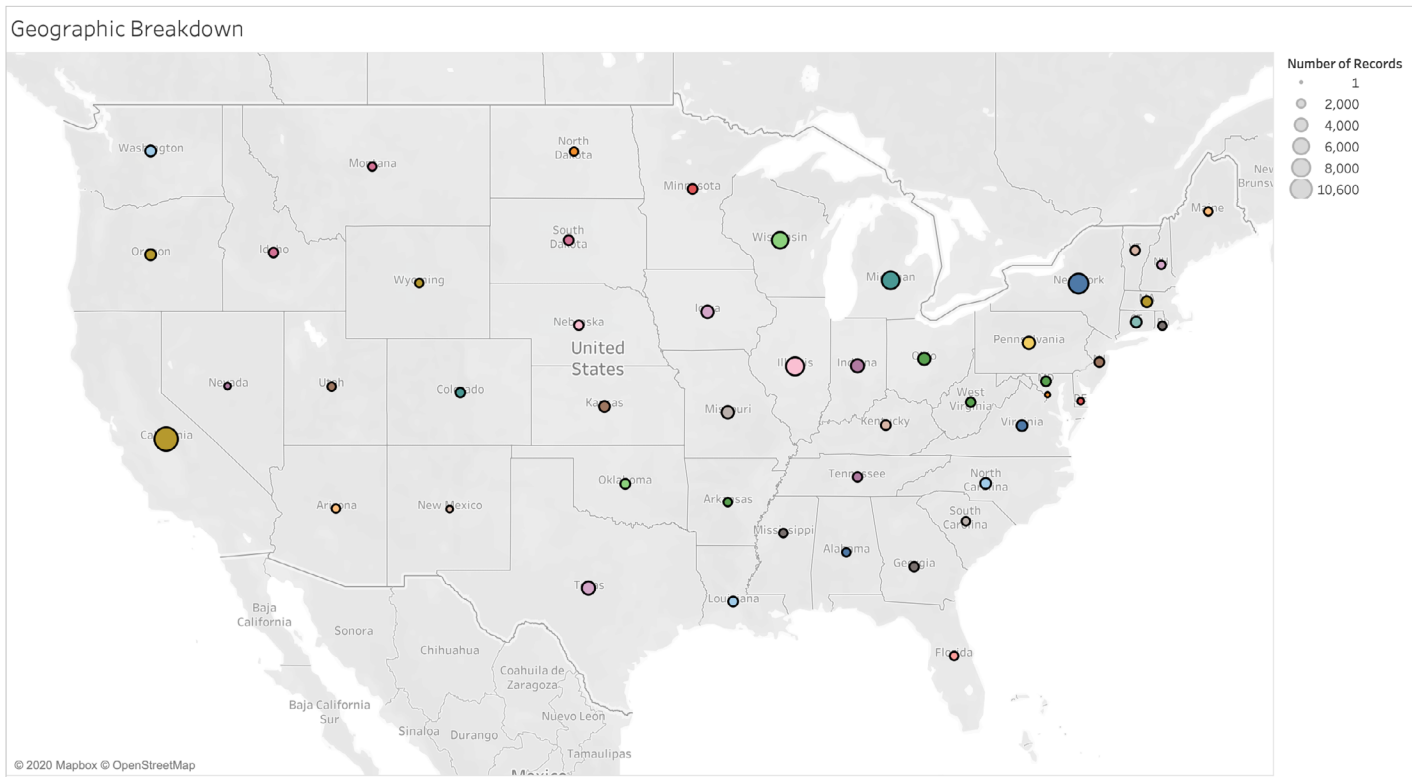


Figure 1. Relative number of documents reviewed by state. © Mapbox, © OpenStreetMap

presence of potentially copyrightable material within an otherwise public domain document was also a notable reason for a work to remain closed. A great resource that goes more into depth about how HathiTrust does copyright review is the book “Finding the Public Domain” which is available open access from Michigan Publishing.⁵

From our results, it appears there was a spike in the 1970s in publications of state governments (or collection holdings within libraries). The presence of copyright notice on works also increased in the 1970s although it still remains a tiny amount of the total. The top five publishers adding copyright notice were St. Paul: West Pub. Co.; Chicago: Callaghan; Atlanta: The State Library; Harrisburg, Pa.: Bureau of Topographic and Geologic Survey; and Topeka, Kans.: R.R. Sanders State Printer. They had each between 65 and 145 publications with notice. The vast majority of the 1,367 publishers using copyright notice had one to ten titles each. Further analysis of the data set might yield an insight into copyright trends within states, but it was not immediately apparent.

Human Resources Needed for This Type of Review

This copyright review work involved thirty-five people from HathiTrust member institutions across the US and Canada.

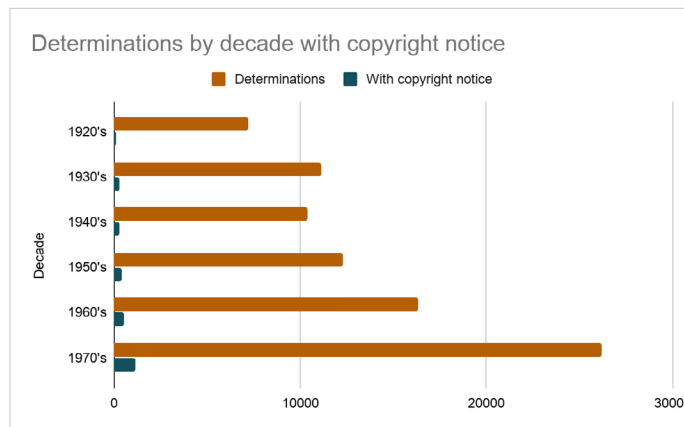


Image 2. Works reviewed vs. works with copyright notice, by decade

They worked on average three hours a week apiece on copyright reviews from 2014 to 2019; in total 8,125 hours of work. This accounts for two independent copyright reviews per volume, and a third review if the first two were in conflict. It does not include time spent in training and project management. The work was done by library staff from HathiTrust member institutions. Their institutions gave between three to six hours per week for a year or more. The library staff often did not have formal copyright expertise and they participated in a HathiTrust copyright review training program. Training could take upwards of three weeks which is why new reviewers

were accepted in an annual cycle instead of ongoing basis. By working with digital objects, significant time savings occurred through not pulling books from shelving.

Use Cases

HathiTrust does not have analytics on usage of these documents, however the copyright reviewers were asked what potential use cases they saw. After looking at thousands of these documents, here is what they suggested:

- Text mining state financial and banking reports to assess the impacts of state spending.
- The importance to women's studies of home economics and family science information originally being published as a part of agriculture science.
- Allowing state employees to digitally access older publications of their own department.
- A valuable image and graphics source for artists.
- Creating a central portal of agricultural information for urban farming where industrial scale methods are not used.
- Research into the history of K-12 educational methods and educational-related metrics.
- Labor and policy documents post WWII, and economic planning at local and regional levels.
- Making interlibrary loan delivery easier to researchers in remote locations.

Opportunities for Additional Work

This project focused on state government documents cataloged as BK format (book) only. We inadvertently tested our ability to work on SE format (serials) when several thousand unknowingly made it into the candidate pool. Despite serial format in general being considered a more complex copyright inquiry than monographic works, these state government serials are very straightforward with the main problem being many items bound together in the same digital scan. For these 'bound-withs' it is difficult to find the start and end pages of all internal parts, and this seems to be a problem unique to working with a digital object. There are approximately 104,250 unreviewed state documents in serial format in HathiTrust that we could work on in a future project.

Another possibility is to expand the date range and look at more state government documents published between 1978-1988. Copyright notice still was not required during these years, but an additional search would have to be taken for rightsholder registration of copyright within five years. Not much is

known about how often this remediation was taken advantage of, however HathiTrust's prior work on state government documents suggests it is very unlikely that a state agency publishing without copyright notice would have submitted a registration to regain copyright.

Conclusion

HathiTrust is proud to have collaborated with our member institutions to achieve this large set of public domain state publications. Our expertise happens to be in copyright review, however we also recognize advocacy efforts of others trying to influence state legislators or seek Creative Commons licenses. US state publications are an important body of work that whatever the means, time invested by libraries in clarifying the copyright status can have some satisfying results.

Kristina Hall (keden@hathitrust.org), Copyright Review Program Manager, HathiTrust.

HathiTrust copyright review of state government documents by Kristina Hall is licensed under a Creative Commons Attribution 4.0 International License.



References

1. Subject Matter of Copyright: United States Government Works, 17 U.S. Code § 105. Legal Information Institute, <https://www.law.cornell.edu/uscode/text/17/105>.
2. "State Copyright Resource Center," Harvard Library Copyright Advisor program, <http://copyright.lib.harvard.edu/states/>; and "Free State Government Information," <http://stategov.freegovinfo.info/>.
3. Kristina Eden, "HathiTrust Members: Opening State and Local Agriculture Documents," HathiTrust Perspectives (blog), May 7, 2019, <https://www.hathitrust.org/blogs/perspectives-from-hathitrust/hathitrust-members-opening-state-and-local-agriculture-documents>.
4. US Copyright Office circulars, "Circular 3: Copyright Notice," <https://www.copyright.gov/circs/circ03.pdf>.
5. Melissa Levine, Richard C. Adler, Justin Bonfiglio, Kristina Eden, and Brian S. Hall, Finding the Public Domain: Copyright Review Management System Toolkit, <https://quod.lib.umich.edu/c/crmstoolkit?page=home>.