IMLS National Leadership Grant – LG-05-14-0042-14
Project Title: Copyright Review Management System 3 (CRMS Toolkit)
University of Michigan Library

Interim Report - Reporting Period December 1, 2014 to November 30, 2015
Submitted by Melissa Levine, PI
Richard C. Adler, Project Manager
Kristina Eden, Copyright Instruction Librarian
Justin Bonfiglio, Copyright Specialist

This is the interim report for ‘CRMS 3’ covering the work accomplished from December 1, 2014 through November 30, 2015.

CRMS 3 was a third project funded by IMLS to support copyright research on works in the HathiTrust digital library. When works are identified by CRMS as in the public domain, HathiTrust makes them available to the public within the limits of current copyright law in each country. The first CRMS project addressed works published in the United States between 1923 and 1963 (LG 05-08-0141-08). CRMS-World continued that effort while also expanding the scope of copyright review to include works published between 1871 and 1944 in Australia, Canada, and the United Kingdom (LG 05-11-0150-11). CRMS 3, the subject of this interim report, continued the work on works published in Australia, Canada, and the United Kingdom; implemented a parallel project devoted to determining the copyright status of US state government documents; and created a CRMS Toolkit providing recommendations and suggestions to help other institutions create their own CRMS-style copyright determination projects.

Description of Project Partners

As the host of the CRMS 3 project, the University of Michigan had the support of sixteen partner institutions that provided staff resources to service as reviewers to engage in copyright determination over the course of this project. Each institution committed staff time to engage in a rigorous training about the CRMS process, how to use the system, and how to make decisions in the CRMS context. CRMS is designed to help reviewers focus on identifying facts rather than make legal decisions per se. Once a reviewer reached a demonstrated level of competence in the CRMS training sandbox (and with oversight from project staff), reviewers engaged in copyright determinations. This shared work is the core of the effectiveness of this cooperative commitment and one of the keys to the success - and complexity - of this project. All participating institutions were also HathiTrust partners.

1. Baylor University
2. California Digital Library
3. Columbia University
4. Dartmouth College  
5. Indiana University  
6. Johns Hopkins University  
7. McGill University  
8. Northwestern University  
9. The Ohio State University  
10. Pennsylvania State University  
11. Princeton University  
12. University of California Irvine  
13. UCLA  
14. University of California San Francisco  
15. University of Illinois Urbana-Champaign  
16. University of Maryland  
17. University of Minnesota  

The University of Wisconsin also contributed copyright determinations, though it was not a cost-share partner on this grant.

Overview

CRMS 3 has been an exemplar of collaborative work across multiple institutions, with over fifty professionals from sixteen partner libraries performing copyright review at a massive scale. Building on the pioneering work of CRMS-US (2008-2011) and CRMS-World (2011-2015), the CRMS 3 project completed copyright determinations on works in HathiTrust that were published in Australia, Canada, and the United Kingdom between 1874 and 1945, identifying an additional 45,000 works as in the public domain. The project also continued a pilot project that adapted the CRMS review process for rights determination of 61,000 US state government documents in HathiTrust.

Changes (cost-share)

- Clarification regarding cost share commitment by Princeton University *(Phone calls and email, Richard Adler and Marvin Bielawski, March 24-25, 2015)*
- Clarification regarding cost share commitment by University of California Irvine *(Phone calls and email, Melissa Levine and Sandra Toro, August 13, 2015; Richard Adler and Kevin Ruminson, September 4, 2015)*
Activities Completed During the Project

As of this report, we have completed or made significant progress on the goals stated in the original proposal.

The CRMS Toolkit. The Toolkit is the most important component of CRMS 3. The Toolkit provides an overview of the CRMS projects, methods, and tools, with process suggestions and recommendations for comparable institutions of higher learning that may wish to create and design a copyright determination project similar to CRMS. The Toolkit identifies and explores legal, technical, and managerial issues relevant to copyright determination, and provides references to additional resources that are helpful for copyright research. Supplemental material—including white papers, reports, and other resources—will be available online via Deep Blue at the University of Michigan.

CRMS-World continuation. When the original grant period for CRMS-World came to a close at the end of November 2014, there were still approximately 83,000 volumes remaining in HathiTrust that had been published in Australia, Canada, and the United Kingdom between 1874 and 1945. The CRMS reviewers continued to make copyright determinations of these works as part of CRMS 3. The interface allowed reviewers to access scans of works in HathiTrust; consult various bibliographic research tools to confirm each work’s metadata; and then render a judgment about the correct copyright status of that work and the reason for that status. Each work received at least two independent reviews, ensuring the reliability of the process. If the two reviewers disagreed, the system presented the work to an expert reviewer who would render a final judgment. Once a work’s rights status was determined, the system exported its updated metadata to the HathiTrust Rights Database. CRMS 3 contributed determinations for approximately 55,000 works, of which 47,000 volumes were found to be in the public domain.

The State Government Documents Project. This project began as a pilot, designed and implemented in 2014, to adapt CRMS-style copyright research for US state government documents. Approximately 61,000 documents were included in the pilot, and work on them continued as part of the third CRMS project. Thus far, approximately 70% of the volumes in this candidate pool have been identified as being in the public domain based on reviews to check for the presence or absence of a copyright notice for US state government documents published between 1923 and 1977.

CRMS Business Plan. Initial planning identified several potential areas of development for a post-grant copyright determination project using the CRMS tools and approach for HathiTrust. Prioritization and implementation of business or sustainability planning can move forward once HathiTrust completes its current round of staff reorganization and resource reallocation in 2016. In the interim, Mike Furlough, Executive Director of
HathiTrust has agreed to retain Kristina Eden to continue to coordinate the review/determination process.

**CRMS Records Preservation Plan.** The University of Michigan library will preserve records pertaining to the administration of the CRMS projects as well as process and results data. Where appropriate, these records will also be made available to the public through Deep Blue, the University of Michigan’s institutional repository.

**Project Results**

*Total World determinations, December 2014 - November 2015:* ~51,000 volumes

*Public Domain determinations, December 2014 - November 2015:* ~45,000 volumes

**What’s Next?**

We have requested a no cost extension to complete the following:

- We are working and planning with HathiTrust leadership to develop sustainable ways to engage in copyright determination efforts.

- Publication and promotion of the CRMS Toolkit.

- We will retain the CRMS primary developer (Brian Hall) to work on stabilizing the CRMS interface and attempt to develop a modular version of CRMS that can be used in conjunction with Hydra and similar systems. This will broaden the likely impact of the toolkit approach, allow for the system to be picked up and tailored more reliably and easily by others for their needs and projects. This will extend the value and long-term prospect for impact.

  - The CRMS code is reliable thanks to over six years of modification and debugging, but the choice to use Perl was decided largely due to its common use in development projects at the time (~2008). If we were building CRMS today, Rails would be a likely alternative because the University of Michigan Library considers it a “best practices” platform for new development. The CRMS pages are relatively static, requiring only an occasional AJAX-style call back to a server for additional computation (for example, when calculating a rights prediction based on an author’s death date in CRMS-World). Most “web languages,” such as Python or Ruby, would be appropriate for building a CRMS-style system. This is an important addition to the utility of the Toolkit for others and key for the sustainability of the tools.
The records preservation plan will ensure that the work we have done will be safely preserved and made as accessible as possible to future projects in publishing, copyright, and digital libraries.

The CRMS suite of projects has led to considerable thinking about new initiatives that could be executed in the future. Here are a few projects that we are thinking about for future consideration:

- Shared data resources with organizations like the Writers Artists and Their Copyright Holders (WATCH) at the Harry Ransom Center in Austin, Texas, with the potential for significantly improving researchers’ ability to identify authors.
- Training for the Library of Congress's Name Authority Cooperative Program (NACO), which would allow CRMS reviewers to contribute authorial metadata, which would improve NACO records for the benefit of librarians and their patrons.
- Digitization and search capability for the U.S Copyright Office's Catalog of Copyright Entries
- Identification of important legislative initiatives pertaining to copyright law in the US Congress
- Extending the CRMS process to non-English volumes in HathiTrust
- Extending the CRMS process to serials

**Grant Products**

The website for the project will be maintained by the University of Michigan Library indefinitely. The website includes links to a wide range of grant-produced products, presentations, reports, and training materials. The primary URL is:

http://www.lib.umich.edu/imls-national-leadership-grant-crms-world

CRMS records appropriate for public access will be made available through Deep Blue, the University of Michigan Library's institutional repository.

Aside from the copyright determinations themselves, the primary product of CRMS will be the forthcoming CRMS Toolkit, which will document the CRMS process including legal, technical, and managerial considerations pertaining to rights research.