This is the final report for ‘CRMS-World’ covering the work accomplished from December 2011 through November 2015. Note that the original grant ran through November 2014; this report covers the entire project including the no-cost extension approved through November 30, 2015. CRMS-World was a second project funded by IMLS to support copyright research on works in the HathiTrust digital library. 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. 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.

Description of Project Partners

As the host of the CRMS-World project, the University of Michigan had the support of eighteen partner institutions, each of whom contributed 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. Duke University
6. Indiana University
7. Johns Hopkins University
8. McGill University
Overview

CRMS-World has been an exemplar of collaborative work across multiple institutions, with over sixty professionals from eighteen partner libraries performing copyright review at a massive scale. Building on the pioneering work of CRMS-US (2008-2011), the CRMS-World project completed copyright determinations on works in HathiTrust that were published in the United States between 1923 and 1963, identifying over 73,000 works as in the public domain. The project also extended and adapted the CRMS review process for rights determination on over 114,000 volumes in HathiTrust published in Australia, Canada, and the United Kingdom, identifying nearly 85,000 more works as in the public domain.

Changes (Personnel, budget, scope, schedule)

Staffing changes
These changes were handled routinely and did not affect overall budget, scope of work, or schedule.

- In April 2012, Richard Adler joined the CRMS project as project manager at the University of Michigan.
- In April 2013, program manager Bobby Glushko took a position as head of scholarly communications and copyright at the University of Toronto Libraries.
- Heather Shoecraft (née Hampton) of the University of Michigan joined the project in February 2013.
- In August 2013, John Wilkin took a position as university librarian and dean of libraries at the University of Illinois, Urbana-Champaign.
- In September 2013, Justin Bonfiglio accepted the position of CRMS copyright specialist at the University of Michigan.
- In May 2014 Michael Furlough became the director of HathiTrust, ensuring the digital library’s continued support for the CRMS project.
Other items

- The CRMS project team was pleased to welcome the California Digital Library as a cost-share partner in May 2012.
- Information about budget adjustments in Summer 2014. *(Formal approvals were not required for these changes; emails were sent as a courtesy. See email exchange, M. Levine and S. Toro, June 17, 2014 through July 7, 2014)*
- Clarification regarding cost share commitments. *(Email exchange, M. Levine and S. Toro, October 31, 2014 to November 19, 2014)*
- No cost extension from November 2014 to November 2015, approved by Sandra Toro in early January 2015. *(Email exchange, S. Toro and M. Levine, January 6, 2015)*
- Budget adjustments in summer 2015, approved by Sandra Toro, that reallocated remaining travel and hosting funds over to payroll. *(Email exchange, S. Toro and M. Levine, August 13, 2015, notification of changes - formal approvals were not required)*

Activities Completed During the Project

We completed all of the goals stated in the original proposal.

**CRMS-World process development.** CRMS-World adapted and expanded the original CRMS review process for determining works published in Australia, Canada, and the United Kingdom. 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.

Development of the interface continued throughout the grant period as we added features, research tools, and other refinements to improve the review process, including:

1. **Pre-population of the author death date field.** Author death dates provided by catalog metadata are central to the rights determination process for works published in Australia, Canada, and the United Kingdom. During the first year of the project, reviewers had to access death dates manually, but in January 2013 we were able to automate this process by altering the interface to pre-populate the author death date field when a death date was available in the catalog record for a given volume. This spurred an immediate improvement in productivity, because
reviewers had only to confirm that death date by finding it in one other trusted source.

2. Integration of the Virtual International Authority File (VIAF) into CRMS-World and CRMS-US interface. VIAF was added as a resource for author death dates that supplied metadata for an additional 9.3% of the volumes under review.

3. An automated error message to alert reviewers who try to log into CRMS from a non-approved IP address. The message also provides information that the information technology team at the reviewer’s university will need to correct the situation. This was one of several measures we introduced during the course of the grant period to improve further the security of the system.

4. Development of the CRMS Sandbox. In order to give trainees a chance to practice, we created a static “sandbox” instance of the review interface. The sandbox is a clone of the production interface but totally separate, so any mistake a trainee makes has absolutely no impact on daily CRMS production. This offers new reviewers the opportunity to become accustomed to the tools they will be using. Hands-on practice in the sandbox makes it easier to visualize and internalize the decision-making steps involved in copyright review.

Training. The first CRMS grant relied on face-to-face training, but the greater number of partners during the second grant period, and their wide geographic distribution, made online training a more practical solution. The first group of reviewers were trained during a three-day summit in Ann Arbor, but the tools prepared for that meeting—including decision trees and a wiki—were then used to build a suite of tools and resources to support online training. By late 2012 we developed handouts, videos, and the ability to host online sessions using voice-over-IP technology. Trainees had access to a "sandbox" which was an exact reproduction of the actual CRMS interface populated with previously judged works for the trainees to review. The sandbox allowed them to submit reviews just as they would in the live interface without interfering in the actual review process in any way. Online training sessions allowed reviewers to "share screens" with experts who could offer real-time recommendations and feedback while the reviewer worked in the sandbox interface.

Regular conference calls with all reviewers allowed us to present brief training lectures and provide updates on new project developments. Feedback from these calls and from other contact with reviewers were used to continuously improve project resources, especially the CRMS wiki. During the final year of the grant, the project team used Qualtrics survey software to create two standardized tests to confirm each trainee’s knowledge of the research process.

The "terminal date." At the outset of this project, we structured our assumptions for CRMS-World on a "life plus 70"year framework applicable today under US and UK law.
However, Canada and Australia follow a "life plus 50" year framework for the volumes under consideration in CRMS-World. This led us in 2013 to investigate the potential consequences of moving the bibliographic cut off date ('terminal date') forward 20 years in "life plus 50" jurisdictions.

Some explanation: When dealing with "life plus 70 years" jurisdictions in CRMS-World, we assumed an author would live, on average, for 70 years after a date of publication, and to this we added the “plus 70” span, for a total of 140 years. This calculation led the project team in 2011 to set the terminal date--the earliest year of publication that would be in-scope for CRMS-World--at 1871. (Because this date increments each year by one, the terminal date in 2013 was 1873.)

While this date was appropriate for the United Kingdom, it seemed overly conservative for Australia and Canada, which follow a "life plus 50 years" rule. That twenty year difference suggested the terminal date in 2013 for these countries should have been 1893. In order to confirm the validity of that later date, we reviewed the rights determinations made in CRMS-World to see how many Australian and Canadian works published prior to 1900 had been determined to be in copyright. After sharing our results with our advisory working group, we moved the "terminal date" for these two "life plus 50 years" jurisdictions to 1893. These dates are incremented forward by one year each January.

Note: Changing the terminal date for these countries also meant changing the corresponding end date for CRMS-World review. Therefore, while works under review from the United Kingdom remained within the span 1873-1943, the scope of CRMS-World for Australian and Canadian works shifted to 1893-1963.

Inheritance. We implemented "rights inheritance" programming for volumes that are duplicates of those that have undergone review. If there are multiple copies of the same book in CRMS-US, the determination for one book will flow to the others.

Monitoring and improving performance. We responsively improved the CRMS-World interface to include more effective search tools, additional rights codes, and more sophisticated ways of gathering data. For example, if the reviewers found a given resource more useful than the others, we moved it up in our list of search tools so it would be more visible and easier to access.

The CRMS-World Advisory Working Group. Each September the members of the advisory working group assembled in Ann Arbor. These meetings were excellent opportunities to present and assess the accomplishments and practices of CRMS, and to solicit feedback from our panel of experts.

Analyzing time commitments and reviewer effort. Over thirty reviewers contributed to CRMS-World, and their levels of commitment to CRMS varied from 5% to 50% of their
work week. The research method used in CRMS requires a detail-oriented focus over an extended period of time, and we were curious to see whether the amount of time a reviewer committed to CRMS each week had an effect on productivity. For each reviewer, we sampled statistics to see if there was a proportional relationship between the amount of time a reviewer committed to CRMS each week and that reviewer’s productivity. A study of reviewer statistics from November 2012 to May 2013 led us to conclude that a 25% - 33% weekly time commitment is most likely to result in optimal productivity. This confirmed the importance of seeking formal commitments from reviewers and established a preferred minimum and maximum amount of time that reviewers should devote each week to the project.

**Assessment of CRMS process by Durationator/Limited Times.** In 2012, in order to have an independent analysis of our review process, we provided three sample sets of volumes reviewed in the CRMS-World process to Professor Elizabeth Townsend Gard of Tulane University Law School. Her Durationator team ([http://www.durationator.com/](http://www.durationator.com/)) conducted an independent copyright status analysis on these records and then compared their results to the determinations made by CRMS-World. The findings of the Tulane team matched 100% of the public domain determinations by CRMS-World.

In 2013 we tested the results of the CRMS-World review process again. We entered a contractual agreement with Limited Times, LLC, whose founding had allowed Professor Townsend Gard to enhance the Durationator in many respects. Our contract had Limited Times make independent determinations on a sample set of 1,000 volumes previously reviewed by CRMS-World. A comparison of the two sets of determinations allowed us to confirm the consistency and reliability of the CRMS-World review process while also suggesting potential paths for further improving it.

**Re-review of select volumes using VIAF.** When the CRMS project began in 2008, VIAF was not yet a service provided by the Online Computer Library Center (OCLC). When OCLC made VIAF available in 2012, it was added to the CRMS interface as a trusted resource, but the reviewers had already made a large number of rights determinations by that point. In the summer of 2013 we decided to compare these determinations with the nationality metadata provided by VIAF. This generated a report with slightly over 9,500 volumes in which VIAF indicated the possibility that an author identified by CRMS as American might, in fact, be of foreign nationality (in which case, their works could be subject to closing due to copyright restoration). While we remained confident about the CRMS review process, in an abundance of caution we decided to conduct a re-review of all of these works to ensure that the review process remained reliable and comprehensive.

**Pilot projects with German and Spanish works.**

- **Humboldt University, Berlin.** In 2012 Michael Seadle, director of Humboldt’s Institut für Bibliotheks- und Informationswissenschaft, approached us to discuss the
possibility of a German-language CRMS. During the spring semester of 2013, Joyce Ray, his colleague and a guest instructor at Humboldt, taught a project seminar for German and Danish iSchool students based on the CRMS method and supporting materials. The class incorporated works from HathiTrust in a study of copyright law in Germany. Melissa Levine, PI for the CRMS-World project, visited Humboldt in late June 2013 and discussed the CRMS projects.

• **Biblioteca Universidad Complutense de Madrid.** In February 2013, we initiated a pilot discussion with the library at the Universidad Complutense de Madrid to identify legal requirements and research tools for reviewing books published in Spain. We began with the idea of working on a specific project tied to a particular research interest. Our counterparts at the Biblioteca proposed working with The Literaturas Españolas y Europeas: del Texto al Hipertexto (LEETHI) Research Group’s project, *Mnemosine: The Digital Library of Rare and Forgotten Spanish Texts (1866-1939).* The LEETHI Research Group provided a sample set of 216 authors to investigate Spanish copyright law and bibliographic resources. The research conducted by the UM team has helped to establish a framework for making rights determinations about Spanish works. Spain has an unusually long copyright duration of “life plus 80 years” for authors who died prior to October 1987. The Biblioteca informed us in May that they could not formally collaborate further on this pilot given other commitments, though we remained in communication.

• Following discussions with Complutense, the CRMS project team began designing a CRMS interface for determining works in Spanish published in Spain. We completed the legal research necessary for making copyright determinations on works published in Spain between 1871 and 1936, and we identified bibliographic resources useful for research of Spanish works. The algorithms underlying the new interface follow a "life plus 80 years" copyright regime, and authors identified by the *Universidad Complutense de Madrid* allowed us to create a pilot pool of approximately 800 candidates (including works by authors who died before 1934). Once the modified CRMS interface was complete in 2014 we completed the pilot project with a testing team composed of a select group of reviewers with previous experience in CRMS-US and CRMS-World. Works in non-English languages are a promising area for the future of copyright determination, and we expect this experiment with Spanish would be equally applicable to works in other languages.

**The State Government Documents Project.** In 2013 Kris Kasianovitz of Stanford University, Sabah Eltareb of the California Research Bureau, and Bernadette Bartlett of the Library of Michigan collaborated with CRMS on a study of copyright law in relation to the digitization and open dissemination of state government publications. Unlike publications of the federal government—which are typically public domain—a state government may choose to assert copyright over its own publications. This possibility makes rights research a necessary part of any library project seeking to make state government documents freely available.
available. Kasianovitz and Bartlett’s project strives to educate and encourage state
governments to adopt more open copyright regimes.

To complement their effort, the CRMS team designed a new research process to address the
kinds of rights issues associated with state government documents. As a test of the new
process, we imported approximately 61,000 government documents from HathiTrust into
the CRMS-US interface and assigned three reviewers to use a modified decision tree to
determine their rights status. Reviews of state government documents began during 2014
and continued into 2015 and the third CRMS grant period. Currently, we have noted that
approximately 70% of the volumes in this candidate pool have entered the public domain.

**DPLA Content Hub.** On June 18, 2013, HathiTrust joined the Digital Public Library of
America (DPLA) as a formal partner and immediately became their largest "content hub,"
ensuring a wide audience for the 3.5 million public domain works in its collection. The
partnership leveraged the strong support IMLS has shown for CRMS by also helping to
cultivate the DPLA as it entered a critical period of high-profile promotion and expansion.
Melissa Levine worked with DPLA on their cooperation with Europeana
rightsstatements.org effort to develop cohesive rights metadata for DPLA and Europeana as
aggregators. The resulting rights statements were in part influenced by CRMS and rights
statements used by HathiTrust.

**Contributions to Name Authority records.** Five CRMS-World reviewers were given an
opportunity to help the U-M Library’s Technical Services Department in their effort to
update Library of Congress Name Authority records. When the reviewers identified author
death dates that were previously unknown, they could submit that information to Tech
Services where selected staff members were certified to update the LC records. Since
January 2013, nearly 400 author death dates were submitted to the Authorities Unit for
approval in revising Name Authority Records.

**Project Results**

**Total US determinations, December 2011 - November 2014:** 191,731 volumes
**Public Domain determinations, December 2011 - November 2014:** 98,017 volumes

The CRMS-US effort reached an important milestone during the second grant period: the
successful completion of the CRMS-US candidate pool. In all this included copyright
determinations for 317,784 volumes published in the United States between 1923 and
1963 of which 167,577 were identified as being in the public domain.

**Total World determinations, December 2011 - November 2014:** 163,731 volumes
**Public Domain determinations, December 2011 - November 2014:** 86,117 volumes
What's Next?

We leveraged the accomplishments and lessons learned in this grant in a proposal to the IMLS for third National Leadership Grant (LG 05-14-0041), currently underway with these deliverables:

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

- A primary goal of the third CRMS grant period (2014-2016) is the creation and publication of a CRMS-Toolkit which will be a guide to other institutions wishing to create a CRMS-style copyright determination project of their own. The toolkit will be produced in various formats and will offer practical advice and recommendations as well as lessons learned during the course of the CRMS projects.

- An additional goal of the third grant period is continued effort to determine the copyright of works in HathiTrust published in Australia, Canada, and the United Kingdom between 1871 and 1945.

- Work continues on the copyright review of 61,000 state government documents that are part of our pilot project.

- Finally, the CRMS project team will create a records preservation plan to 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.

- We also working on refining the online interface - exploring ways to make the programming more modular and thus reusable in and beyond the HathiTrust context.

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

Public work product will be archived in Deep Blue, the University of Michigan Library’s institutional repository. This will provide public access and stable DOIs for publicly available work product.

The primary products of CRMS are the copyright determinations as documented in the
rights database (currently not public), the training materials and the interface tool. The CRMS Toolkit, subject of our third grant now in process, will document the CRMS process including training and the interface in a new publication, forthcoming.