Overview

The Impact of CRMS. On June 18, 2013, HathiTrust became a formal partner of the Digital Public Library of America (DPLA). Sharing the metadata for 3.5 million public domain volumes will make HathiTrust the single largest “content hub” in the DPLA, ensuring a wider audience for these freely available works. The partnership leverages the strong support IMLS has shown for CRMS by also helping to cultivate the DPLA as it enters a critical period of high-profile promotion and expansion.

CRMS-US. We expect the CRMS-US team to complete their review of over 290,000 volumes in HathiTrust shortly before the close of calendar year 2013. With that milestone reached, we will transition the US reviewers and experts to the CRMS-World project where their contributions will give World a timely boost during its final phase.

CRMS-World. The CRMS-World team made steady progress during the past six months. We refined the review process, made several improvements to the interface (including adding new research resources), and updated our training material accordingly.

All major requirements and milestones for the grant are well underway:

- Establishment and ongoing improvement of the CRMS-World review workflow and interface
- Corresponding improvement of support and training material
- Consistent progress in reviewing American, Australian, Canadian, and British works.
- Pilot explorations for broadening the scope of CRMS to include Spanish, German, and other non-English works.
- Providing opportunities for internships and student engagement.

Some useful statistics:

<table>
<thead>
<tr>
<th></th>
<th>CRMS-US</th>
<th>CRMS-World</th>
</tr>
</thead>
<tbody>
<tr>
<td>Volumes remaining to review</td>
<td>~24,000</td>
<td>~133,000</td>
</tr>
<tr>
<td>Actual determinations made</td>
<td>~267,000</td>
<td>~59,000</td>
</tr>
<tr>
<td>Percentage of determinations: PD</td>
<td>~142,000 (~53%)</td>
<td>~32,000 (~54%)</td>
</tr>
<tr>
<td>Percentage of determinations: IC</td>
<td>~47,000 (~18%)</td>
<td>~7,000 (~12%)</td>
</tr>
<tr>
<td>Percentage of determinations: UND</td>
<td>~78,000 (~29%)</td>
<td>~20,000 (34%)</td>
</tr>
</tbody>
</table>

Project Activities, December 2012 – June 2013

The CRMS-World Review Process, Workflow, and Interface. The CRMS-World interface and its underlying logic were thoroughly tested over the past year, and we have continued to refine the review process and add new functionality to the interface. Our reviewers’ experience has shown that the most time-consuming part of the review process is the identification of author death dates. Thus, we have added features to help streamline that step:

Pre-population of the author death date field with values from Mirlyn. Mirlyn, the University of Michigan Library’s online catalog, is one of the trusted resources available to reviewers within the CRMS-World interface. The catalog’s metadata includes author death dates, but reviewers have had to manually select Mirlyn as a resource in order to access this information. In January 2013 we were able to automate this process by altering the interface to pre-populate the author death date field with metadata from Mirlyn when a death date was available 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. When we first implemented this feature, Mirlyn was able to provide author death dates for 15,152 out of 26,851 volumes (56.4%).

Integration of the Virtual International Authority File (VIAF) into CRMS-World and CRMS-US interface. Soon after providing pre-populated author death dates from Mirlyn, we made an additional improvement to the CRMS-World interface by integrating author death dates from the Virtual International Authority File. VIAF was able to supply author death dates for another 9.3% of the volumes under review, providing a further improvement to reviewers’ productivity. We also added VIAF as a trusted resource to the CRMS-US interface, where it now helps CRMS-US reviewers to identify non-US authors whose work may be subject to GATT restoration. (Works subject to GATT restoration are in copyright in the US, even if they are in the public domain in their countries of origin.)

Moving the "terminal date." We noted in our previous report that we had begun a preliminary study of the "terminal date" question. 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 to investigate the potential consequences of moving the bibliographic cut off date ("terminal date") forward 20 years in "life plus 50" jurisdictions. We looked at data generated from twelve months of CRMS-World reviews to see how many works were identified as in copyright. Interestingly, no works were identified as in copyright before 1900 from Australia or Canada. After sharing this data with our advisory working group, we moved the date for bibliographic determinations forward by 20 years, to 1872, in the "life plus 50" jurisdictions of Australia and Canada. This "terminal date" will roll forward by one year each January. We have limited this shift only to a 20-year increase, rather than moving the terminal date to 1900, in an abundance of caution. The number of books that will be opened as a result of this change may not be significant, but it may encourage participation in HathiTrust by libraries in life plus 50 countries.

Contributions to Name Authority Records. Some of the reviewers in CRMS-World are now working with the UM-based project team to improve the trusted resources used for researching author death dates in CRMS. When reviewers identify author death dates that were previously unknown, they can submit that information to the UM Library’s Technical Services Department where selected staff members are being certified to update Library of Congress Name Authority Records. Since January 2013, nearly 400 author death dates have been submitted to the Authorities
Unit. These updates are currently under review by the Library of Congress to ensure they meet Resource Description and Access and Name Authority Cooperative program (RDA NACO) standards. Upon approval they will be used to revise Name Authority Records.

**Publicly provided author death date information:** Author death dates are sometimes supplied by members of the public when they submit inquiries to HathiTrust. To support this effort, the HathiTrust User Support Group worked with the CRMS team to create an online list of preferred sources for author death dates. The list is available here: [http://www.hathitrust.org/help_copyright](http://www.hathitrust.org/help_copyright)

**New Experts.** Two new expert reviewers completed their training in May 2013. Jennifer Block of Princeton and Heather Hampton of UM have provided valuable help in adjudicating conflict reviews, reducing the backlog in our "conflict queue," and removing the "bottleneck" in review processing that we noted in our previous report last December.

**Improvements to the Expert Reviewer interface.** Another response to the "bottleneck" in expert review processing was an improvement to the expert interface. Previously experts had to cut-and-paste data from one interface to another in order to adjudicate reviews. Now an expert can select a review, and the system will import its values into corresponding fields in the expert interface. This saves time, and it has eliminated what had been a very tedious step in the expert review process.

**Working With Our Partners**

**Analyzing time commitments and reviewer effort.** 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 the accuracy of that reviewer's work (determined by how often a reviewer's judgments were later changed by an expert). 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 a higher quality of reviews. This confirmed the importance of seeking formal commitments from reviewers and established a preferred minimum of time that reviewers should devote each week to the project in order to ensure their proficiency in the review process.

**Conference calls.** In February and May 2013 we conducted conference calls with the CRMS-World reviewers. The calls were an opportunity to touch base, clarify copyright policy, and solicit feedback from our partners. Response has been positive, and we plan to continue the calls on a quarterly basis. We noted an increase in productivity immediately following each call, which made a case for reminding even committed partners about the importance of a project and the value of the work they contributed to it. A project like CRMS requires a sustained, long-term view and a detail oriented focus over an extended period of time. Helping the reviewers understand that they are "part of the team" is critical for morale and the final success of the effort.

**Training material.** We finalized the development of a "sandboxed" training interface and revised our standards for remote training. The training interface provides a hands-on experience for trainees that is entirely separate from the actual CRMS interface. Trainees can submit reviews just as they would in the "production interface" without interfering with the actual review process in any way. The hands-on experience is supplemented by screencast videos which provide detailed instructions and examples drawn from common situations. Trainees may contact the project team at UM via email to solicit advice. They can also request "shared screen" sessions via Skype in which an expert provides real-time recommendations and feedback while the reviewer works with the
interface. So far, five reviewers have been trained using online materials; the sandbox interface; and email and phone contact with the CRMS staff at UM.

Publicity and Outreach

Coalition for Networked Information (CNI). Melissa Levine and Richard Adler presented "Rights, Research, Results: The Copyright Review Management System Projects" at the CNI Spring Membership Meeting 2013 in San Antonio, TX, on April 5, 2013. Their presentation offered an overview of the goals and accomplishments of CRMS; provided a brief explanation of the CRMS review process; and invited a discussion of possible next steps for the CRMS projects and what lasting impact they might have on the future of copyright, fair use, and the public domain.

Internships and Student Projects. We have engaged students in several projects related to the CRMS effort:

- **Claire Tatro**, a graduate student in the University of Michigan School of Information, is completing a 180-hour Practical Engagement internship with us this summer. Tatro is writing a series of blog posts to showcase interesting works that had been identified as in the public domain by CRMS and which may now be read in HathiTrust. The blog posts note the support provided to the project by IMLS. The series is called “HathiTrust Summer Selections” and can be found on the MPublishing blog at: [www.publishing.umich.edu](http://www.publishing.umich.edu) (Summer 2013)
- **Dean Atiya**, a third year student at the University of Michigan Law School, completed an internship with the CRMS Project last semester. His research on copyright law in Spanish-speaking countries has been a valuable resource for our effort to expand the scope of CRMS to include non-English works. (January – May 2013)
- **A team of students** enrolled in the University of Michigan's School of Information selected CRMS-World for a usability project in fulfillment of the course SI 622: Evaluation of Systems and Services. Their study of the CRMS-World interface included heuristic evaluations, user surveys, reviewer interviews, and an assessment of "competitor" products. (January – May 2013)

Explorations and Pilots

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 (1868-1939)*. 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 eighty" 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

---

2 Training resources are available at the following URL: [http://www.lib.umich.edu/imls-national-leadership-grant-crms-world/presentations-and-resources](http://www.lib.umich.edu/imls-national-leadership-grant-crms-world/presentations-and-resources)

3 Nerea Llamas, *Head of international studies and subject specialist for Latin American, Caribbean and Iberian Studies*, and Dean Atiya (UM Law School, Class of 2014)
commitments, though we remain in communication. The CRMS team at UM will work on a Spanish CRMS interface with the goal of crafting a workflow for Spanish books.

**University of Texas, Austin – Argentine Books:** Building on the interest in Spanish language materials, Georgia Harper, a leading copyright expert at the University of Texas in Austin and a member of the CRMS-World advisory group proposed that she work with us to develop a legal and bibliographic framework for evaluating works from Argentina as part of the Spanish diaspora. Her work has also informed the CRMS Spanish-language project. UT Austin has a significant collection of books from Argentina. This pilot might result in a foundation for reviewing those books as well as an opportunity to test our ability to replicate the ‘CRMS Process’ at another institution with the same level of responsibility and reliability now in place at UM.

*Humboldt University, Berlin.* Michael Seadle, director of Humboldt’s Institut für Bibliothekswissenschaft, Informationssysteme und Informationswissenschaft, approached us to discuss the possibility of a German-language CRMS last year. He also took part in the first meeting of the CRMS-World advisory group held in September 2012. During the spring semester of this year, 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 and discussed the CRMS projects. The course continues through this summer, and we will evaluate the work product at its conclusion. (Joyce may be familiar as a former program manager at IMLS.)

**The Harry Ransom Center’s WATCH File.** In our previous report, we noted that we had initiated conversation with the Harry Ransom Center to identify common interests around copyright and author information. That discussion continues. Richard Adler, the project manager for CRMS, plans to meet with Ransom Center counterparts at the Society of American Archivists meeting in August 2013 in New Orleans and discuss possibilities for sharing author metadata.

**State Documents Proposal.** Also in our previous report, we noted that Kris Kasianovitz at Stanford University and Bernadette Bartlett of the State of Michigan Library, have embarked on a project to demonstrate and address the constraints imposed by federal copyright law that inhibit access to digitized post-1923 state government publications. The state’s right to assert copyright over their publications significantly impacts and impedes knowledge institutions’ ability to digitize historic state publications, capture and archive born digital publications, and freely disseminate them to the public via library catalogs or digital repositories. Currently Kasianovitz and Bartlett are submitting proposals for conferences and seminars to discuss and promote the project. They are keeping us apprised so we can support them where possible. Their effort could be an important complement to our work as we look for ways to involve others in the overarching goal of copyright determinacy.

**Changes**

**John Wilkin to Dean, University of Illinois, Urbana-Champaign.** John Wilkin has been named university librarian and dean of libraries at the University of Illinois at Urbana-Champaign, effective
August 16, 2013. He was responsible for the original conception of CRMS and has been integrally involved with the project from its inception. We are discussing the implications of his departure from the University of Michigan.

**Departure of Program Manager.** On April 30, CRMS Program Manager Bobby Glushko accepted an offer at the University of Toronto Libraries. The CRMS-World project team had several weeks to prepare for this change, and Bobby will continue to be involved with the project as a member of the advisory working group. We have begun the interview process to fill his open position.

**Looking Ahead**

*Assessment of the CRMS-World process:* In order to test the CRMS-World review process, we have made a formal agreement with Elizabeth Townsend Gard of Tulane University and her Durationator Project. The Durationator group will make determinations on a sample set of volumes previously reviewed by CRMS-World. Comparing their results with those of CRMS will provide a means for assessing the consistency and reliability of the CRMS-World review process. Establishing terms has taken slightly longer than anticipated due to the unprecedented nature of the arrangement. While CRMS-US was able to make a testing arrangement with the US Copyright Office, no corresponding agency exists for works published overseas.

*Analysis of “Undetermined” volumes.* As of July 28, 2013, 29% of CRMS-US and 34% of CRMS-World volumes have been categorized as “undetermined,” requiring further research before they can be found in copyright or in the public domain. Several factors may lead to a judgment of “undetermined,” and some of these are likely to be easier to resolve than others. An analysis of undetermined volumes may suggest sub-categories of works within this group that may allow us to re-evaluate them more efficiently. We look forward to exploring this issue further with our advisory working group.

*Opportunities for crowdsourcing.* In the spirit of openness and transparency, the project team has been committed to exploring ways to open the CRMS-World review process to a wider community of researchers. This has been limited by our rigorous security and commitment to protecting copyrighted material. However some parts of the review process are not necessarily dependent on access to restricted content, and these may lend themselves to crowdsourced research efforts. One possibility for a crowdsourced project would be the identification of author death dates, which might help to reduce the time required to research volumes in CRMS as well as leverage the diverse knowledge and skill sets of a large group of researchers.

*Preparation for Advisory Working Group Meeting in September 2013.* Our advisory group has been generous with its time and expertise, and we are looking forward to our next meeting with them in September. The agenda for that meeting will likely include a general update; a review of pilot projects and concerns; a discussion of the problem of inserts; and ideas for the next phase of CRMS and its funding.

---

4 Note that we obtained permission from the IMLS to adjust the grant budget to reallocate $4,000 from travel funds to pay for expenses related to working with Professor Townsend Gard. Email correspondence between Melissa Levine and Chuck Thomas, March 14, 2013, on file.