Overview

The grant project for the Copyright Review Management System (CRMS) is currently proceeding according to schedule. The project is in Phase Two, which is scheduled to run from June 2009 through May 2010. As reported in our previous Interim Performance Report (submitted July 31, 2009), Version 1.0 of the CRMS was released on July 10, 2009. Since that time, staff have reviewed over 8000 volumes (published in the US between 1923-1963) in the HathiTrust Digital Library using the CRMS. Two reviewers now review each volume, and an expert arbitrator makes a final determination for any conflicting reviews. Since the release of Version 1.0, project staff have worked to resolve technical issues, define additional priorities and implement features and enhancements to the System on an ongoing basis. We have put in place a mechanism for sharing rights determinations with other interested parties. We also continue to focus efforts on ways to improve the accuracy of our determinations and to clearly document our guidelines, in order to improve efficiency in determinations work, and in preparation for training reviewers at other CIC institutions starting next year.

Phase Two Goals and Accomplishments to Date for the Reporting Period

1. Staff use CRMS to perform copyright status review work
   Accomplishments: As of the release of version 1.0, staff at the University of Michigan rely solely on the CRMS to perform and manage copyright review work. We have implemented the “double-blind” review process, store all reviews in the CRMS database, and export rights determinations automatically to the HathiTrust rights database on a nightly basis.

2. Project Manager works with Technical Services Staff to monitor issues; Project Manager works with Programmer to resolve system-related issues
   Accomplishments: Project Manager continuously tracks issues and works with Electronic Access (EA) Unit Administrator and Programmer to introduce enhancements to the functionality of the CRMS, which are incorporated into major and minor updates to the System. We most recently released version 1.3.4, and expect to release one more update (1.3.5) before the end of the year. Examples of enhancements include:
   o improved search functionality, including ability to search all reviews by volume, date range; publication date (see historical-reviews-screenshot)
   o added reporting for reviewer statistics (see user-stats-cumulative-screenshot and user-stats-individualreview-screenshot)

1 Note: various screenshots of the interface are included on the CD submitted along with this report
The document includes the following points:

- Added export statistics reporting (see export-stats-screenshot and export-stats-graphs-screenshot).
- Added support for downloading tab delimited export statistics data.
- Added mechanism to indicate validated vs. non-validated reviews so reviewers can identify “incorrect” reviews for learning purposes (see historical-reviews-screenshot – last column = “verdict”).
- Added mechanism to add volumes to the queue for “on demand” reviewing (see add-to-queue-screenshot).
- Significantly enhanced “user help” and “search help” pages.

3. Develop methods to share findings on copyright status (scheduled September 2009-February 2010)
Accomplishments: After deliberation, stakeholders determined that the best course of action for sharing rights determinations at this juncture was to add the information to the tab-delimited HathiTrust metadata files that are provided at http://www.hathitrust.org/hathifiles, beginning with the full metadata file produced on December 1, 2009. Fields added include the copyright determination reason code and the date the rights database entry was last updated. With this data included, the tab-delimited files will become an ongoing accessible source for information on how and when rights determinations are made. In addition, rights status determinations for volumes in HathiTrust are also currently being shared through OAIster, which was recently transferred to OCLC (as they pick up our records via OAI feed).

4. Requirements gathering for version 2.0 (scheduled December 2009)
Accomplishments: The Project Manager, Programmer and EA Unit Administrator have begun to define requirements for version 2.0 of the CRMS. Areas under discussion include:

- User interface improvements, including reviewer and administrative views
- Additional search/query/reporting needs
- Database redesign
- Revisions to review process (to reduce ambiguity thereby reducing number of conflicting reviews)
- Mechanism(s) to accommodate reviewer questions/feedback;

---

2 A description of the files is available at http://www.hathitrust.org/hathifiles_metadata.

3 In the grant proposal, we stated that our intent was also to share determinations with OCLC WorldCat, Bowker Global Books in Print, and OCLC’s Copyright Evidence Registry, however circumstances surrounding each of these have evolved since the time the proposal was submitted. Determinations will be shared with OCLC as part of the OCLC e-content synchronization project, which is delayed by OCLC at this time for reasons unrelated to this project. Sharing determinations with Bowker is pending the outcome of contract negotiations. OCLC has postponed further development and implementation of their OCLC Copyright Evidence Registry.
5. Publicity – website; conference presentations

Accomplishments:
- Expanded grant webpage to a project “site” within the University Library website as of October 2009 (http://www.lib.umich.edu/copyright-review-management-system)
- Project Manager traveled to Washington DC November 2-5, 2009, to discuss the CRMS with various staff from the US Copyright Office
- Project Manager presented on the CRMS as part of the Yale Copyright Lecture Series on December 9, 2009 in New Haven, CT

6. Central Coordinator will monitor productivity and take steps to evaluate improvements in accuracy as described in the evaluation plan and refine processes as needed (dates not specified)

Accomplishments:
- Project manager works closely with EA Unit Head and EA Unit Administrator to monitor productivity and takes steps to reduce conflict rate
- Project Manager and EA Unit Administrator revised and updated the “Use Cases” and “Determinations Guidelines” used for Copyright Review work to help reduce ambiguity
- Project Manager and EA Unit Administrator conducted training with staff in various areas where conflicting reviews were most prevalent
- EA Unit Head conducted one-on-one training with staff on an as-needed basis

7. Begin to work with CIC institutions (Indiana University, University of Minnesota, University of Wisconsin) to engage in collaborative testing in Phase Three

Accomplishments: Project Manager has been in contact with the appropriate parties at each of the three CIC institutions and is in preliminary discussions regarding staffing, scheduling and training activities.

8. Miscellaneous Accomplishments not specifically included in grant proposal

- Staffing: The new permanent CRMS Programmer (who was hired as of July 6, 2009) worked with the Interim Programmer to get oriented to the project and took over all programming duties as of mid-August.
- Legacy Reviews: Project staff assessed needs and determined plan for re-review of oldest copyright status determinations (done in earliest phases of manual

---

4 Reviewers and experts rely heavily on the ability to view the title page and other frontmatter of a book in order to make a reliable determination. Currently we utilize the HathiTrust pageturner to view page images, however, if there are many pages of frontmatter to page through, this can slow the process down quite a bit. Our goal is to find a more efficient solution for viewing the page images of the digitized books in order to decrease the average time per review, without sacrificing accuracy and consistency.
copyright review work - pre 2008) in order to improve accuracy and reliability of these determinations. We also developed routines and schedule for loading legacy reviews (Jan 2008-June 2009) into the CRMS database to improve access to these determinations.

- Copyright Office Visit: As part of the evaluation plan for the project, we agreed to assess our ability to improve reliability by contracting with the US Copyright Office for Services to evaluate our rights determinations. In November 2009, Project Manager traveled to Washington DC to meet with Research staff at the US Copyright Office, in part to clarify next steps required for hiring the Copyright Office to undertake comparison searching.

Preliminary Findings

We are currently grappling with several factors that constrain our productivity (numbers of determinations, and consequently, number of volumes determined to be PD) since introduction of the CRMS and the “double blind” review process. The number of copyright determinations per month has gone down by more than 50% with the introduction of the CRMS, as the process now requires that each volume is reviewed by at least two reviewers, and up to three if the initial reviews conflict. Prior to the introduction of the “double-blind” review process, we were making an average of approximately 3700 determinations per month. With the double blind process in place, and 18.5% of volumes requiring expert resolution, we are averaging about half that number of determinations (1561) per month since introduction of the CRMS in July. This is mainly due to limits on the amount of time qualified experts have to devote to reviewing conflicts on a daily basis. We are therefore limiting reviewers to no more than 50 reviews per day to keep the amount of time required for expert review to a manageable amount. Since we submitted the last Interim Narrative Report, we have taken steps to reduce ambiguity in the review process, for example, by clarifying guidelines for making determinations and working one on one with staff who were having difficulties in specific areas, and we are clearly identifying “incorrect” reviews so reviewers can go back and learn from their mistakes. However we are still in the process of fully analyzing the types of conflicts that are occurring most often in order to determine where and how we can refine our review processes and thereby increase productivity rates.

In the grant proposal, we state that our goal is to open up on average 1700 volumes per month. The average percentage of volumes reviewed that are likely to be in the public domain is currently around 56%. Therefore we essentially need to double our current

---

5 This number is based on work performed during the seven months prior to release of the CRMS (12/08 through 6/09)

6 Expert review is generally more time consuming and requires a higher level of staff. These staff (currently the Project Manager, EA Unit Head and EA Unit Administrator) are limited in the amount of time they can devote to resolving the daily conflicts because the majority of their time is already committed to other managerial responsibilities.

7 We have seen a corresponding increase in the rate of undetermined status as the percentage of public domain determinations has decreased. Historically this rate has been around 7-8%; however, for volumes reviewed using the CRMS, the undetermined rate is close to 17%. We attribute this to the fact that our
review rates (bringing them back up to pre-CRMS levels, with the double review process in place) to achieve our stated goals. We believe that this is not an insurmountable task in the long term, especially as we bring on more reviewers. However in the short term, new reviewers will require even more expert time for guidance, supervision and feedback. Therefore, in order to succeed in reaching the goals we set for ourselves in the grant proposal, we will need to work more vigorously in the short term to reduce conflicts and will need to reduce the amount of time it takes to make a reliable determination.

**Next Steps**

As we continue with Phase Two of the project, we plan to stop releasing major updates to the system as of the end of 2009 and focus the majority of our efforts on more fully defining requirements for version 2.0 of the CRMS. Development and testing will take place starting next year, with a target release date of May 2010. In early 2010 we plan to identify the set of volumes for comparison searches and submit the required information to the Copyright Office to ensure that results are returned to us by May 2010. The Project Manager will continue to work on plans to train and add reviewers from three other institutions (Indiana Minnesota and Wisconsin) over the coming months. We will also explore options for explicit sharing of rights determinations on an “on-demand” basis, possibly via the HathiTrust Digital Library catalog and/or pageturner interface. Finally, we will continue to address issues that limit productivity under the CRMS process, with a goal of reducing conflict rates by further refining our determinations guidelines, and possibly increasing time devoted to expert review by shifting resources among current project staff. We are also considering adding staff to the project as early as next fiscal year. We plan to make adjustments to our reviewing process that can help reduce the amount of time required to review a volume and reliably make a determination of its rights status.