

**ALCTS Technical Services Directors of Large Research Libraries Group**  
**Cornell University Library**  
**June 2010 ALA Annual Round Robin**

**Large-Scale Digitization**

Central Library Operations (CLO) Batch Processing staff have completed Voyager record creation for the 62,693 public domain titles digitized by Kirtas and held by the Internet Archive. Free, unimpeded access to these titles – representing over 71,000 Cornell print volumes – is now available through both Voyager and OCLC WorldCat. We have also begun to send bibliographic metadata to OCLC for those Cornell titles digitized by Google and made available in Google Books. Catalog records for these titles will soon appear in WorldCat. Access to all full-text and “limited preview” Google Book versions of Cornell print holdings, when they exist, will be available through Cornell’s local catalog (Voyager) by means of an application programming interface, or API.

Jim LeBlanc has been appointed to the Google Metadata Working Group, which is investigating ways to improve the quality of metadata associated with Google Books. Jim and CLO staffer, Laurie Stevens, recently participated in a review of the efficacy of Google’s revamped duplicate detection program, which is now operating with 93%-95% accuracy. Given the nature of the metadata with which Google is working, and based on our own experience with this kind of catalog data matching, this fully automated duplicate detection rate is actually quite respectable.

**OpenURL Research**

Adam Chandler, CLO’s E-Resources and Database Management Research Librarian, is collaborating on a Mellon-funded grant to build a Classical Works Knowledge Base (CWKB) – a relational database and specialized link resolver software – to facilitate linking from citations of ancient texts to the online versions of those texts (<http://www.news.cornell.edu/stories/June10/ClassicsBase.html>). The database will ultimately cover all Latin and Greek authors from Homer to Bede, from approximately the eighth century B.C. to the mid-eighth century A.D. In conjunction with his work on the CWKB, Adam has written a program that parses OpenURLs, counts the elements that are most likely to be needed for successful linkage (title, ISSN, author, date, and so forth), and indicates whether these elements are present or absent in the OpenURLs (<http://openurlquality.niso.org/>).

**ARTstor Shared Shelf update**

The Metadata Services Unit has been participating with ARTstor in the development of a shared institutional image management system called “Shared Shelf”. As part of the Shared Shelf cataloging Interface sub-committee meetings, Metadata Librarians Glen Wiley and Liz Muller have been actively contributing to discussion of short- and long-term objectives for the new visual resources cataloging tool. Discussions have focused on cataloging interfaces, complex object cataloging issues, supporting multiple authorities, use case scenarios, data values and labels. These discussions are in preparation for the summer 2010 beta version release of the Shared Shelf product and test migration of three Cornell image collections. Glen and Liz will continue working with ARTstor on open access solutions to existing Cornell image collections.

**Retrospective Barcoding**

CLO recently finished the retrospective barcoding of the Cornell circulating print collection: a project that began over twenty years ago with the implementation of our first integrated library system. With barcoding complete, all this material can now be easily charged and tracked through our Voyager circulation module. Moreover, all items chosen for transfer to the library’s offsite storage facility can now be accessioned in the Library Annex inventory control system – a process that requires a barcode and item record for each piece transferred.

### **Patron-driven purchase plan**

As you will recall from the last Round Robin update, in August 2009 Cornell selected Coutts MyiLibrary to support a small, patron-driven plan for ebooks from 10 publishers in the subjects of engineering, computer science, physics, astronomy, chemistry, and mathematics. MARC records for new (c2009+) ebooks meeting the profile have been loaded into the catalog and if the ebooks are used by patrons twice, they are purchased for our permanent collections. The corresponding print editions have been blocked from our print approval plan. Since the beginning of this new service, we've experienced disappointing and continued delays in the availability of the electronic editions after the print editions have been published and distributed. From follow-up investigations, we've learned that this is typical for several publishers (Oxford UP, Princeton UP, ...) in the Cornell profile. In the near future, we will supplement the patron-driven ebook plan with a patron-driven print plan for the same profile. If an e-version is not available within 60 days of the availability of the print version, a record for the print version will be added to the catalog for discovery by our users. When they select these titles, rush orders will be executed to expedite delivery of the print edition. We'll be sure to provide an update of the outcome in our report next summer.

### **POOF!: bringing some magic to acquisitions**

Cornell and Columbia libraries (2CUL) are collaborating on a new web tool to facilitate the workflow for firm orders that fall outside of the WorldCat Selection stream (Cornell is a WorldCat Selection library). The Pre-order online form (POOF!) is still in the development phase, though we expect to have a working proto-type by year's end. We anticipate that POOF! will be used by library patrons making purchase requests, subject specialists submitting order requests for either or both Cornell and Columbia collections, acquisitions staff keying order requests, and specialized materials vendors who could submit basic metadata directly to a subject specialist for selection. Once the metadata exist in a standard form, poof! the orders get submitted through our existing tools as part of a batch process.

### **More about 2CUL**

Columbia and Cornell are also working on procedures for cross-cataloging each other's collections (starting with Turkish books), procedures to support a shared selector for Slavic materials, and exploring automated workflows for e-resource ordering.