

Catalogers Group  
March 10, 2005

Present: G. Aivazian, S. Benamou, V. Bross, B. Feinberg, C. Horning, C. Kincy, J. Lopear, J. Matthiesen, L. Mendes, C. Miller, H. Phan, L. Ratliff (recorder), J. Riemer, A. Riggio, R. Villena  
Guest: Hannah Walker (LIT)

Valerie Bross moderated a panel of presentations about a cataloging productivity program called Web Cataloging Assistant (WCA), developed by the Library of Congress Business and Economics Cataloging Team, and demonstrated how it could be used to automate the production of MARC records for electronic resources in the California Digital Library's eScholarship Repository. Valerie acknowledged Luiz and Raphael for the ideas and technical work which resulted in the program as it was demonstrated today.

(See **below** for an article about WCA from the *LC CATALOGING NEWSLINE*.)

1. Angela Riggio began with an overview of her E-Resources and Metadata Cataloging section, which supports the e-resources (ER) cataloging workflow. Her staff check for new ERs that are added to the ERDb by collection development librarians, search for cataloging copy, and then route the resources to the appropriate cataloging staff:
  - a. The E-Resources Cataloging & Metadata Section catalogs integrating resources;
  - b. Serials are routed to the Serials Cataloging Section;
  - c. Monographs with copy go to the Copy Cataloging Section;
  - d. Original monographs go to the Monographic Cataloging Section.
2. For resources with no copy, before being sent on for full cataloging, Encoding Level 3 records are added to the UCLA catalog, consisting mainly of author, title and URL. This ensures their timely addition to Voyager. These records will later be overlaid by full records. The goal of the E-Resources Cataloging & Metadata Section is to catalog every resource needed by the selectors.
3. The Web Cataloging Assistant (WCA) tool can create a Level 3 MARC record automatically for e-resources in the eScholarship Repository.

Next, Valerie demonstrated the WCA.

1. The WCA assists in productivity by automating steps. It extracts html data from an e-resource (or: extracts data from an html page) and creates a bibliographic record directly in Voyager.
2. She played a short video produced by LC's BEAT in which Gabriel Horchler describes the functions of the WCA <http://www.loc.gov/catdir/beat/webcatvideo.html>. BEAT uses WCA to catalog selected high-research-value e-resources that are free and are part of selected monographic series. At this time they have processed about 300 series, with about 40,000 individual monographs.
3. The WCA, developed by David Williamson, creates a MARC record in Voyager, and transcribes the descriptive bibliographic data. All the cataloger needs to add are subjects, classification, and other authors besides the first one.
4. Note: a separate WCA profile needs to be created for each separate monographic series. WCA requires an HTML display in a uniform layout for each monograph in the series.
5. For more information, see LC's Bibliographic Enrichment Advisory Team website. <http://www.loc.gov/catdir/beat/webcat.html>
6. Locally, Raphael Villena customized WCA for UCLA use. To run the UCLA-customized WCA:
  - a. Create a folder-- C:\Programs\Web Cataloging Assistant
  - b. Start up Voyager maximized in the background and run the program, located in:  
I:\Serials\Web Cataloging Assistant\escholarship.2.exe
  - c. Start the .exe program for the series.
  - d. Fill in cataloger's initials and URL.

That's all!

7. Connexion also provides a builder for e-resource bibliographic records, but the results are uneven because it depends on the HTML source code in the record.

WCA eliminates a lot of copying and pasting. So where do we go from here? People could monitor the UCLA portion of the eScholarship Repository, but this would take too much time. It would be good to receive email notices of changes to a web page, to know when a paper has been added; or the submission process could generate a notification to be emailed.

Hannah Walker from Library Information Technology gave an overview of the eScholarship Repository, including how papers are submitted, the organization of it, and reasons for inconsistencies in the data.

See <http://repositories.cdlib.org/escholarship/>

There are no guidelines for campus organization naming conventions for the submitting units. They set up their profiles, which cannot be subsequently changed.

For behind-the-scenes information, see the article "Administering a Working Paper Site Using the bepress System" <http://repositories.cdlib.org/wpguide.pdf> It tells what data fields they can add, for example: author, affiliation, title, abstract, subject words. The required format is RTF or PDF (Word documents are converted to PDF upon submission.) The default order for lists of the papers is reverse date, then label (a unique label, assigned by the submitter).

There are email alerts, but you need to enter search terms. There is no way to do it except by setting up a search for each individual campus site. There is no feedback mechanism.

There is a tracking mechanism for updated versions of papers; the updated web version will carry the original date. The digital version doesn't always match the paper version that we have. A digital cover page is created at the time the paper is submitted (a "built" page), so the actual paper begins with page 2.

The question was asked, is there any way to work with campus units to get ISSNs for their series? Hannah indicated that is a CDL question. Cindy Shelton is our representative to the Office of Scholarly Communication at CDL. There are no advisory groups to help coordinate the submission process.

The CDL is registered in the OAIster project, and sends information for harvesting by the OAIster server. See <http://www.oaister.org> at the Univ. of Michigan.

---

From:

LC CATALOGING NEWSLINE  
Online Newsletter of Bibliographic Access  
Library of Congress

Volume 13, no. 5      ISSN 1066-8829      March 2005

WEBCAST NOW AVAILABLE DESCRIBING "WEB CATALOGING ASSISTANT"

The Bibliographic Enrichment Advisory Team (BEAT) is pleased to announce production of a Webcast to describe a new cataloging tool, the "Web Cataloging Assistant" (Web Cat Assistant), in use for the creation of full MARC 21 records for monographs published within selected series available on the Web. This software allows a cataloger to look at the abstract page for a particular monograph, copy the URL for that page into the PC clipboard, and

paste it into the application. The tool then retrieves the Web page itself and extracts bibliographic data from that page. It also copies available abstracts and keywords provided by the author(s) to the record. The cataloger may add subject headings and name entries. Web Cat Assistant then creates a MARC record that is loaded into the Library's cataloging client, and the record is distributed by the Cataloging Distribution Service. An application has been developed in conjunction with Web Cat Assistant that automatically generates an email message whenever the Web pages of a particular online series have changed, alerting the cataloger that a new title has probably appeared. David Williamson, cataloging automation specialist, is the creator of Web Cat Assistant.

At this time, Web Cat Assistant is being applied to the cataloging of the current and retrospective monograph titles issued as part of some twenty selected, high research value series in the social sciences. It will gradually be extended to other social science series and, in the near future, to scientific series. It has also proven to be a useful tool for staff at the Library who are authorized to work from home, as Web Cat Assistant can be easily used off-site.

A six and one-half minute video production describing the project and tool is now available on the BEAT Web site (<<http://www.loc.gov/catdir/beat/>> [Feb. 2004]) by clicking on the link under "What's new." Gabriel Horchler, the narrator in the video, is leader of the Business and Economics Cataloging Team, Social Sciences Cataloging Division.