

Catalogers Group
Minutes
July 6, 2006

Present: Nora Avetyan, Sharon Benamou, Valerie Bross (recorder), Chamy Kincy, Andy Kohler, Luiz Mendes, Caroline Miller, Nancy Norris, Louise Ratliff, John Riemer (facilitator), Sharon Shafer, Kristen Wilson

Announcements

1. Email workshop: Next Thursday
2. Watch for: Endeca demonstration, coming soon

Discussion

Issues that Arose at ALA New Orleans (continued from last week)

1. The "New" catalog
2. OCLC/RLG merger

Additional topic

3. Calhoun report

1. The "New" catalog

a. ALA Session: ALCTS Open Forum on the New Catalog (Louise Ratliff):

- Two presenters demonstrated strategies for enriching the catalog search experience for library resources.
- First, a librarian from Howard County Library demonstrated TLC's AquaBrowser Library, which has 3 columns: (1) a visual display of words (called a tag cloud); (2) an OPAC-like list; (3) a list of facets for search refinement. Catalog records are exported nightly to a separate database. The records are converted to XML, indexed, and linked to book reviews, cover images, etc. The AquaBrowser interface includes user search helps such as "did you mean to type ..." and RSS feeds.
- Next, a Kristin Antelman from North Carolina State University demonstrated Endeca. NCSU has a SIRSI catalog; records are exported to the Endeca database for enhanced discovery tools. For example, users can enter an empty search to pull up the entire catalog. The relevance ranking algorithm is better in Endeca than in SIRSI—a test showed 42% relevance in the retrieved items in SIRSI; this jumped to 68% in Endeca. Serials holdings posed a particular problem; to resolve it, if someone retrieves a record for a serial through Endeca, the system links back to the SIRSI catalog to show holdings.
- Finally, Jennifer Bowen described a Mellon grant to develop a planning document outlining how to implement an open source ILS. More information on that is available at: <http://extensiblecatalog.info>

b. Discussion:

- Considerations (Andy Kohler): Both Endeca and AquaBrowser entail creating a separate database and maintaining separate indexes, in addition to those that support the ILS. As we have learned from ERDb, this means that one needs to consider record maintenance issues, synchronization issues, re-indexing issues (Endeca requires re-indexing of the entire database each night). Also, one needs to think about the balance between real-time functions (such as circulation data) and enhanced discovery. In addition, there is a question of scalability: Melvyl is several times larger than any of the catalogs using Endeca or AquaBrowser. Nevertheless, both of these products demonstrate interesting new strategies for catalogs.
- BSTF Follow-up (John Riemer): The SOPAG BSTF group will be meeting with Bernie Hurley, as well as with reps from HOTS (Heads of Tech Services) and HOPS (Heads of Public Services). Goal: To outline options for next steps, in preparation for the ULS meeting in September.

<http://aquabrowser.hclibrary.org/>

The screenshot shows the Howard County Library website. At the top, there are navigation links for 'Library Home', 'Traditional catalog', 'Cancer information', and 'Flu Resources'. Below this is a search bar with 'deforestation' entered. The search results are sorted by 'relevance' and show five items:

- 1 Deforestation /** Caleb Owens. Briefly describes how the world's forests are being destroyed some of the causes and consequences of this destruction and some ways of preventing it. Subject: deforestation, forest conservation. Call No: J333.750. Year: 1999. Found: Deforestation (2).
- 2 Archangel /** Paul Watkins. Subject: deforestation, forest protection, ... Call No: F Wat. Year: 1995. Found: Deforestation (1).
- 3 Alternatives to deforestation: steps toward sustainable use of the Amazon rain forest** Anthony B. Anderson. Subject: deforestation, forest management, ... Call No: 333.75A. Year: 1990. Found: Deforestation (2).
- 4 The challenge of extinction /** Dorothy Hinshaw Pagant. Focuses on the scientific, technological and social impacts of plant and animal extinction. Series: Environmental issues series. Subject: extinction (biology), extinct animals, ... Call No: J574.529P. Year: 1991. Found: Deforestation (1).
- 5 The fate of the forest: developers, destroyers and defenders of the Amazon** Susanna B. Hecht, Cockburn, Alexander. Subject: deforestation, forest conservation, ...

On the right side, there are filters for 'Select Location', 'Format', 'Author', 'Subject', and 'Series'. A 'Barcode' field and a 'Login' button are also visible.

<http://aqua.queenslibrary.org/>

The screenshot shows the Queens Library website. At the top, there are navigation links for 'Home', 'Books', 'Articles', 'DVD/Video', 'Music', 'Websites', 'Software', 'Audiobooks', and 'Events'. Below this is a search bar with 'deforestation' entered. The search results are sorted by 'relevance' and show four items:

- 1. Deforestation** social dynamics in watersheds and mountain ecosystems. J. Jurek, David C. Pitt. Subject: deforestation. Found: Deforestation (2). Date: 1988. Publisher: Routledge. Call Number: 333.75 D.
- 2. Degradation of the land** Earth at risk. Mark Chernigoin. Examines the problem of land degradation through erosion and deforestation discussing how these processes occur and what effects they have. Subject: soil degradation, soil conservation, ... Found: Deforestation (2). Date: 1991. Publisher: Chelsea House Publishers. Call Number: 333.7612 C.
- 3. Deforestation in the postwar Philippines** University of Chicago geography research paper ; David M. Hunter. Subject: deforestation. Found: Deforestation (2). Date: 1992. Publisher: University of Chicago Press. Call Number: 333.7613 K.
- 4. The rainforest** Earth at risk. Laura Tangley. Subject: rain forest ecology, rain forests, ... Found: Deforestation (2). Date: 1992. Publisher: Chelsea House Publishers.

On the right side, there are filters for 'Select a Branch', 'Format', 'Author', 'Subject', 'Language', and 'Series'. An 'External Sources' section is also visible.

<http://www2.lib.ncsu.edu/catalog/>

The screenshot shows the NCSU Libraries website. At the top, there are navigation links for 'Search the Collection', 'Browse Subjects', 'Services', 'Library Information', 'Community', and 'News & Events'. Below this is a search bar with 'deforestation' entered. The search results show 342 matching items. The first few results are:

- 1. The economics of deforestation in the Amazon : dispelling the myths** Author: Campari, Joë S. Published: 2005. Format: Book. SD418.3 .A53 C26 2005. Stacks. Available.
- 2. The dynamics of deforestation and economic growth in the Brazilian Amazon** Published: 2002. Format: Book. SD418.3 .A6 D96 2002. Stacks. Available.
- 3. Deforestation, environment, and sustainable development : a comparative analysis** Published: 2001. Format: Book. SD418 .C43 2001. Stacks. Available.
- 4. The Causes of tropical deforestation : the economic and statistical analysis of factors giving rise to the loss of the tropical forests** Published: 1994. Format: Book. SD418.3 .T76 C38 1994. Stacks. Available.

On the left side, there are filters for 'Subject: Topic', 'Subject: Genre', 'Format', and 'Library'. A 'Sort By: Relevance' dropdown is also visible.

Additional background information:

Size of institutional holdings, from ARL statistics of 2004: NCSU 3,389,517; UCLA 7,988,925; Queens Borough Public Library 9,691,126. Melvyl (from its FAQ) "over 25,000,000 records."

NCSU on reindexing:

"The entire process, including re-indexing the entire database, takes approximately 7 hours. Future changes in the architecture of the technical backend should cut this re-indexing time in half." <http://www.lib.ncsu.edu/endeca/technology.html>

NCSU on circ status:

"Circulation status can be up to 23:59 out of date, depending on when the activity takes place. NCSU is looking into more permanent solutions to the display of this information but has determined that the benefits of displaying the information in the hitlist outweighs the limited potential of it not being completely accurate." <http://www.lib.ncsu.edu/endeca/faqs.html>

"AquaBrowser Library's search engine holds an index of the full title catalog. When the user enters a query, all relevant fields of the MARC record are searched instantaneously." <http://www.medialab.nl/ABLWhitePaper2006.pdf> (page 6)

"There are two points of integration with an ILS. In the back-end AquaBrowser Library retrieves MARC extracts from the ILS to create its indexes. In the user interface, the user is guided towards the native OPAC display of a title when he selects one, providing him with the functionality the OPAC already offers." <http://www.medialab.nl/ABLWhitePaper2006.pdf> (page 9)

Aquabrowser's customers are almost all public libraries and small colleges. However, Queens is a very large public library. [How many bibliographic records Queens has, comprising its 9 million volumes, is not readily available.] <http://www.medialab.nl/index.asp?page=customers/overview>

2. OCLC/RLG merger

- a. Standing Committee on Training, CONSER/BIBCO meeting (Caroline Miller): Glenn Patton (in a side-conversation) explained that the two organizations merged administratively on July 1st. For the time being, RLG will maintain offices in Mountain View, California. Twelve migration teams are looking at various issues, including: (1) a Digital/Cultural resources team; (b) a Eureka services team (renegotiating licenses with databases); (c) a Resource Sharing team; and (d) a Technical Services Mapping and Functionality team. OCLC will be posting a list of tasks and timelines on the Web site.
- b. Copy Cataloging meeting (Luiz Mendes): Rich Greene discussed OCLC's meetings with the 95 RLG clients to determine concerns. Clients have expressed concern about: (a) preserving the individual institutional records (as opposed to OCLC's master record structure); (b) Z39.50 access to authority records with vernacular; and (c) access to superseded authority records (limited archive available in OCLC, though not visible; more complete archive available in RLIN—and available to customers).

The merger will not halt OCLC's development of new and enhanced services to existing Connexion customers. Robert Bremer talked about features of the new multiple terminologies services. This service uses existing Windows features to allow catalogers to search thesauri and bring the valid terms into an OCLC record, with the appropriate tags. In addition, OCLC's new Harvester creates preliminary records from HTML, pdf, and Word files. But the Harvester is still far from perfect--catalogers will need to review the resulting record for accuracy (e.g., of dates). OCLC is also working to make other non-MARC schemas available; watch for more news this fall.

- c. OCLC President's Luncheon (John Riemer): UCLA had been considering joining RLG because of the opportunities in standards development. After the OCLC/RLG merger, a RLG-Programs unit will continue to exist, reporting to Lorcan Dempsey; Gary Strong is on the Program Council:

http://www.oclc.org/services/brochures/12203_RLG_Programs_1.0.pdf. Participation in the group should strengthen UCLA's voice.

Quote to remember (Kristen Wilson): At the Webwise conference, Murtha Baca quipped: "Standards are like toothbrushes. Everybody thinks they're a good idea, yet nobody wants to use anybody else's."

3. Calhoun report: *The Changing Nature of the Catalog and Its Integration with Other Discovery Tools*. Final Report. Mar. 17, 2006. Prepared for the Library of Congress by Karen Calhoun. <http://www.loc.gov/catdir/calhoun-report-final.pdf>
 - a. Series: p. 18, 4.2.11: The report encourages cost/benefit analysis to determine the need for controlled series
 - b. Extending the use of metadata: p. 14: In the "Extend" trapezoid, note "Deploy existing catalog data in new ways." This recommendation implies usefulness of metadata currently available in library catalogs
 - c. Simplification: p. 13: One challenge to innovation, according to Calhoun, is "Resistance to simplifying cataloging." This is also reflected in 4.1.1 "Simplify catalog records..." Does this conflict with the desire to make more use of the data in records?
 - d. Support for decision-making: p. 13: Calhoun cites as a challenge to innovation, "Limited availability of data to support management decisions." Conclusion: We need better reporting capability to provide that data.
 - e. RDA: p. 17: Calhoun advises monitoring RDA. But will RDA simplify cataloging?