

Cataloging & Metadata Interest Group October 2002 Minutes

Member's Council Delegates attending one or both meetings: Carol Pitts Diedrichs (Chair), Ed Weissman (Vice Chair), Rosann Bazirjian, Karen Boehning; Huanwen Cheng, Greg Cotton, Robert Ellett, Eleanor Frierson, Cliff Glaviano, Betsy Hine, Barbara Kriigel, Susan Phillips, Kris Senecal (EC Liaison)

OCLC Staff attending one or both meetings: Glenn Patton (OCLC Liaison), Cynthia Whitacre (Recorder), Judy Barnes, Brenda Block, Sharon Bosarge, Robert Bremer, Robin Buser, Ellen Caplan, Bill Carney, Kay Covert, Becky Dean, Maureen Finn, Linda Gabel, Chris Grabenstatter, Rich Greene, Gary Houk, Lynn Kellar, Fred Lauber, Morgan McCune, Joan Mitchell, Bob Pearson, Jim Simms, Eliza Sproat, Anna Sylvester, Bob Van Volkenburg, David Whitehair, Marty Withrow, Julie Whitley, Lois Yoakum

Others attending: Brad Baker (Board Liaison), Judith Dzierba (Health Science OCLC UG), Betsy Friesen (Connexion UG)

Monday, October 21, 2002, 9:00-11:00 AM

Connexion:

Following welcome and introductions from the Chair, Marty Withrow began an update on OCLC Connexion. He showed progress to date, including an online features spreadsheet, available at: <http://www.oclc.org/prod/cataloging/features.xls>. This spreadsheet compares features of Passport, CatME, Connexion browser interface, and the Connexion Windows client, and outlines when new features will be available in Connexion.

Jim Simms then demonstrated the upcoming November enhancements to be added to the browser interface. Questions/Answers/Comments included:

- What's in Function drop-down for each field? Functions for that field
- Are there short cut keys for these functions? No, but cut & paste can be used. There are a limited number of shortcuts available with the browser.
- It is confusing when clicking tabs to go to another database or area and the screen doesn't change. This is on the list to change, but OCLC is unsure where to take someone who clicks. Suggestions included a different color or a blank screen.
- With keyword searching, one can only enter one term for WorldCat but multiple terms for the Resource Catalog (RC). Why? It has to do with current indexing structures. Adjacency keyword searching will be available when WorldCat moves to Oracle. That will eliminate the inconsistencies.
- Preferences for look, including number of tabs shown in a session, is controlled by the profile and authorization.

David Whitehair then showed progress to date on the Windows Client.

Questions/answers/comments included:

- Dublin Core view will not be available in the first 3 releases. It may or may not be added in the future. It is available in the browser interface.

- There is no timeframe set for the non-English interfaces for Connexion or for non-Roman cataloging moving to Connexion except “future”.
- What is the main reason for offering both the browser and Windows client? Cataloger demand, productivity issues, offline local files which cannot be done via the browser. However, the browser allows cataloging via any machine. Will the next generation of catalogers want a Windows client?
- Has OCLC surveyed librarians about cataloging on local systems vs. on OCLC? Yes, the answer seems to be whatever system has better editing. The economics of online vs. offline also comes into play.
- OCLC is striving for more consistency between interfaces. Some displays in Connexion mirror displays in the ILL Web product and some mirror FirstSearch.

Chris Grabenstatter then talked about Connexion browser enhancements scheduled for February, including authority control improvements, constant data improvements, diacritics changes, and Fixed Field display enhancements. After February, plans included importation of files, searching/editing enhancements, NACO national review file, Cataloging Agent authorizations, and lots of little improvements.

Questions/Answers/Comments included:

- Why doesn't OCLC allow display of delimiter a (≠a) as an option, to make it more consistent with other systems? This was added to the list for future consideration.
- When will libraries be able to import PromptCat files? Tentative May 2003 date.
- OCLC has announced January 19 as end of support date for Netscape 4.x. End of support in this case means users will no longer be able to logon to Connexion using Netscape 4.x after this date.

All delegates are encouraged to submit questions or comments from themselves or library staff at anytime to Chris, Jim, or David. Last year's conference calls were very useful and may be repeated this year.

FRBR:

Rich Greene discussed a development project to “FRBR-ize” WorldCat.

Note that Functional Requirements for Bibliographic Records (FRBR) can be found at: <http://www.ifla.org/VII/s13/frbr/frbr.htm>.

WorldCat as we know it now is cataloged at the “manifestation” level. All development efforts to date concern the “work” level. Most uniform titles are “work” titles. Relating work titles to names is very difficult. OCLC is looking at “agents” which include all personal, corporate, and conference names. An AACR2 option which LC doesn't follow assigns roles (editor, illustrator, etc.) to agents. Catalogers are, however, making these decisions about the agent's role when doing original cataloging.

One internal FRBR prototype is called FISH (FRBR Interface Simulation Habilitation). This prototype contains a small sample of records. 281 J.K. Rowling records were extracted from WorldCat and grouped into 20 work records. The work Harry Potter & the Philosopher's Stone has 83 manifestation records associated with it. The use of a uniform title enabled all of these records to be pulled together. A second example, J.R.R. Tolkien, included 1838 records collapsed to 202 works. This may collapse further with more study.

Diane Vizine-Goetz demonstrated a research project employing FRBR known as FictionFinder. Approximately 2 million records for printed fiction were extracted from

WorldCat and organized by FRBR works. These records grouped into 1,287,178 clusters of works. The quality of data in the records is crucial for successful clustering.

The top 10 clusters include works like Robinson Crusoe and Alice's Adventures in Wonderland. The Tolkien and Rowling clusterings were also available in FictionFinder.

Questions/Answers/Comments:

- Where does expression fit in? This is vague and controversial. A good example of an expression is a translation of a work in a particular language. Each translation is a different expression. Sometimes a different format is an expression and sometime it is a new work. For example, where does a radio play from a novel fit?
- What does this mean for OCLC catalogers and users? Too soon to know.
- What was the goal of IFLA with FRBR? IFLA realized there was not a conceptual model of what catalog records do. FRBR provides such a conceptual model. Online catalogs and the need to catalog electronic resources (as opposed to a piece in hand) also prompted the development of FRBR. When moving from cards to online catalogs, we lost the supporting interrelated structure of how catalogs were organized. FRBR brings some of that organization into the online world.
- ISTC (International Standard Text Code) will become a work/expression identifier and may help in collocating works to realize FRBR. OCLC will be the agency administering ISTC, which is coming out in January. Work is underway to accommodate this code within MARC.
- Local system vendors other than VTLS don't seem to be openly working toward use of the FRBR model. In house developers at Univ. of Texas are interested in learning.

Tuesday, October 22, 2002, 8:30-9:45 AM

Glenn Patton noted the vendor report which had been distributed Monday, and highlighted the information on local system vendors (LSVAP). David Whitehair asked anyone interested in participating in a field test for any phase of the Windows Client to contact him.

What is Quality?:

Lynn Kellar and Glenn Patton reviewed the handout on this topic that had been distributed prior to the meeting. As OCLC plans to re-implement WorldCat in a new database structure, how will quality be defined? What records will be considered duplicates? What will catalogers or other users want in brief record displays to indicate a quality record? Relevance ranking is useful for reference. Is it also useful for cataloging? Resource sharing users want to select records with the most holdings. How will introduction of formats other than MARC affect quality of records in WorldCat? How will FRBR work? Will different versions of the same record co-exist?

Questions/Comments:

- There is an obvious need to maintain cataloging productivity.
- Have the summary displays identify which records are vendor records.
- Flexibility is key, so that users can define their preferred view(s) for searching and retrieval (DLC first, or DC only, etc.). User needs may differ daily.

- Show source of records in brief display
- Show which records have subject headings and classification in preferred schemes
- Will profile be tied to authorization? Will it be customizable online by the user when logging on? Can administrator be given ability to profile staff authorizations?
- Show which records have linked authorities
- If a record has been entered in one scheme, such as DC, can it be retrieved in another scheme, such as MARC? Qualified yes, depending on format. It may not always be possible to map between metadata formats. If a record is entered in a format, it can always be output in the same format.
- An indicator of fullness of records in brief displays would be useful. But, how do we define fullness? Different users may have different definitions.
- Will there be any record merging in the new system? Yes, but exactly what will merge needs to be defined. The definition of “duplicate” needs to be re-defined.
- Lynn Kellar discussed a concept referred to as the “common data format” in which all records will potentially be stored. It is XML based and will allow MARC in XML, EAD in XML, etc. This will allow manipulation of data, and it will evolve over time. OCLC has discussed this with LC, RLG, and some local system vendors. Display would extract views out of the XML common data.
- Different format views of the same record (DC and MARC for example) are not duplicates, but need to be linked
- With linking products and programs, will everything need to go into the local OPAC? We may just want to point to data stored elsewhere.
- ONIX allows lots of non-MARC types of data and pointers
- Within records that have contributions from many sources, could the display show which pieces came from which sources? (There was disagreement over the usefulness of this.)
- OCLC intends to link as many headings as possible to authority records, to allow automatic updating to occur.
- The Virtual International Authority file project was mentioned, along with work on mapping RAK and AACR2 headings to each other.
- Have there been member surveys on quality that could help guide decisions? Yes, but not focused to appropriate audience perhaps.
- Metadata created by author’s themselves early in the publication process is an intriguing idea. Faculty are being asked to do this, as are graduate students for theses and dissertations.
- What won’t we allow in WorldCat? Will there be anything excluded? For example, will things of only local value with local rights, such as local TV station digital video inventory be excluded? What is appropriate in a local group catalog view may not be appropriate for display to a wider audience. Are there guidelines on what we ought to be submitting to WorldCat? Perhaps we ought to consider developing these.

Action needed: The CMIG would like to look at sample records in many types of metadata schemes to help in making comparisons, such as ONIX, DC, EAD, MARC, etc. This might be a good learning tool to share once developed.

Planning for future meetings:

Carol Diedrichs asked for suggested topics for future meetings. Those suggested were:

- Continuing development of Connexion
- FRBR
- Regional groups, consortia, including for-profit entities, and their impact on OCLC and WorldCat
- Rights & resolutions services
- Currency surveys of WorldCat content
- WorldCat indexing (perhaps a conference call topic)
- Display enhancements from the cataloging perspective
- Open WorldCat