Minutes of the Cataloguing and Authorities Users' Group Meeting
March 1, 2010, Robarts Library Room 4049

Present: Alastair Boyd *(Robarts),* Anne Dondertman *(Thomas Fisher),* David Eden *(New College),* Barbara Geiger *(St. Augstine's),* Anna Golodnitsky *(Knox),* Monica Hypher *(CIRHR),* Connie Lewin *(St. Michaels),* Wendy Li *(Hong-Kong),* Kate MacDonald *(Trinity),* James Mason (*Music),* Sian Meikle *(ITS),* Stephen Qiao *(East Asian),* H. Rashid *(Law),* Mary Reynolds *(Regis),* Mary Ruscillo *(Robarts),* Elisa Sze *(Inforum),* Anna Slawek *(Robarts),* Carmen Socknat *(Victoria College),* Sherry Smugler *(Government Documents),* Eva Spevak *(Robarts),* Stephanie Swift *(OISE),* Pui-Ying Wong *(HSICT – Bloorview)*

1. **Minutes and business arising from November 23ʳᵈ meeting**

   Minutes approved as written. Alastair Boyd (*Robarts*) added a clarification regarding training for Sirsi's "Director's Station": Though originally stated that training would commence "early in the New Year", this is still pending. However, it will take place; ITS will keep us informed.

2. **Report from ITS**

   Sian Meikle (*ITS*) reporting.
   1. *EIR reporting (Update).* There is a concern about how links between print items and their electronic equivalents are being maintained in Catalogue Records. The status is that "a little more glue" is needed to complete the new automated process for notifying ITS of catalogue records which need to be added to EIR. When the system is up and running the MARC 856 field linking to the online resource will be replaced by a link to EIR. Reporting is done through a web form which will be on the staff intranet labeled "Send e-resources".

      Several questions concerning the EIR process were discussed. Some discussion focused around adding 856 fields for local purposes, and the question of when and why records should be reported to EIR. Records that link directly to resources rather than to EIR need to have their 856 links maintained in the catalogue; also, licensed resources cannot be accessed by off-campus catalogue users through such records. So, reporting to EIR is important, since link-checking and maintenance are then handled by ITS, and the resources are accessible to any authorized U of T user, even off campus.

      Question: will it be necessary to go back and delete 856 fields for direct resource links after resources have been reported to EIR? *Answer*: this will not be necessary as a script will be able to flip these 856 fields over to the EIR URL when the system is ready.

      There was some discussion about the work load that the reporting of resources to EIR could create, particularly in Government Documents where in any case there are often links to general landing pages not specific items. Sian suggested that it is still good to report these, even if it is just a general landing page as this is already a situation that they deal with, but a link to the most specific information available is appropriate.

      The example of Annual reports was mentioned. Occasionally there is one record with multiple holdings. Alastair Boyd (*Robarts*) said that these items should be treated as a serial and a link to a general page is appropriate.

      Links to items such as online Tables of Contents don't get reported to EIR as they don't get catalogued on their own and don't exactly represent the item being catalogued.

      Finding Aids was another example discussed. Sian suggested that these would be good to report, even if the catalogue record links only to a more general landing page as it still can

help with access and maintenance of the link. Yet, it was acknowledged that at times it may be difficult to decide when to, or not to, report to EIR and that there is more discretion for local items. An Index to a collection, for example, could be included in an 856, but would only be reported to EIR if it is catalogued separately.

Sian clarified the situation by stating that the EIR system is about a one to one title relationship for the same item; one a physical copy in the catalogue and one available electronically. Sian also stated that EIR is not just for licensed print resources but can also be for free internet resources that get catalogued.

If an E Resource is purchased, and its record is loaded into Sirsi before a print version, it is easiest to add another record for the print version, then request that holdings and links etc. from the e-book record be transferred. Send email to Marlene or Alastair and they will merge the records.

2. *Sirsi news:* ITS is still *g*etting the Browse search function ready. To see the test catalogue go to: (http://search2.library.utoronto.ca/UTLBrowse/search.jsp) . Some problems persist to do with UTF 8 non-Roman script searches.

Endeca's interface is in Java (programming language), and making modifications is intensive on programmers' time. So, some things like library location have been pulled out into separate tables rather than being hard-coded, so they can be updated more quickly.

Work is continuing on the Mobile Interface and Streaming Media Server.

There were questions and concerns about priorities in ITS. There was some feeing that basic things like the proper display of public notes, serials holdings and On Order records were constantly being delayed for work on the browse function, media server, and now perhaps the mobile device interface. Sian assured us that all requests will be worked on but it will take time. There is one programmer to do a lot of work. A list of issues is being maintained and all items will be addressed. Sian will provide this list so that we can confirm that issues relevant to us are represented.

3. Report from the Metadata Librarian

Alastair Boyd reporting on behalf of Marlene van Ballegooie who was away on Sirsi training.

1. *MARC 046 Changes:* all remaining 75,000 records have been fixed and should no longer cause any errors in our monthly Worldcat contributions.

2. *MARC 041 Language Changes:* Back in the fall, the 442,000 language codes were split apart and corrected, but the changes were not yet reflected in Sirsi because we were waiting for some programming time to become available so that we can re-write this data back to the records. Since the fall, Marlene has proceeded to gradually make the changes herself. Currently, 122,000 language codes have been fixed. Also, with new updates to the MARC editing software that she uses, it now looks like it will be much easier to make the changes to the remaining 320,000 records.

3. *Consultation on Digital Collections:* Marlene has also been consulting with a couple of libraries on metadata issues relating to new digitization projects.

   Henri J.M. Nouwen Video Archives – St. Michael's College
   Hong Kong Digitization project – Canada Hong Kong Library

4. *E-Theses:* Marlene is working to match approximately 6,000 electronic theses held in T-

Space with the corresponding records for the print versions in Sirsi. This will allow us to simply attach the URL to the existing Sirsi record rather than load a completely separate record for each thesis held in T-Space. All of the matches were made before Christmas, however we are still waiting for ITS to create a script so that we can load the data.

The project just mentioned deals with a retro file of electronic theses, more recent submissions have not yet been dealt with. There appears to be a problem with the paperwork between SGS and ProQuest. As a result, the electronic theses that have been submitted last year are not yet appearing on the ProQuest website.

We have been debating whether to load locally generated MARC records from the T-Space Dublin Core, based on the student original submission, or wait for the MARC records created by ProQuest. Because of the paperwork problem at SGS, we have not been able to compare the quality of the ProQuest MARC records with our own. We have, however, been able to compare the records from the Queen's and the University of Montreal's D-Space applications and the corresponding ProQuest records. By comparing the D-space version and the Proquest version, we noticedthat ProQuest actually adds quite a bit of information to the record. You get a fuller form of name, more consistency in capitalization in the title, an ISBN, Proquest Document ID and Publication number. They also add more standardized keywords/subject headings and pagination information. With these enhancements, it is Marlene's sense that the ProQuest MARC records are worth waiting for.

5. *XML Parsing and Matching:* Marlene has been working with ITS on a project to link Internet Archive items with no match points to the corresponding records in Sirsi. This work involves writing scripts to parse data from the Internet Archive MARC XML records and then query the Sirsi database via Z39.50 for records that match. The project has been highly successful. Out of the 9,442 items that needed to be matched, only 341 remain.

   The project has been quite valuable in that she can both parse and write XML files using scripting methods. Also, she has also learned to better target her Z39.50 searches by querying specific Use attributes that are defined on the host server.

6. *Vetting Vendor E-book Records:* Marlene continues to vet batches of MARC records supplied by e-book vendors, modifying them as necessary.

7. *Bound Withs* and *Order Status in Endeca:* Marlene has also been working with Bilal on the issue of Bound With records in Endeca. Presently, the 'child' items are not showing up at all in Endeca. What is needed is that when Endeca finds a child item, the call number and barcode of the parent item is displayed, preferably with a note or some indication that the item is bound within another item.

   The relationships between the data in the Sirsi database have been identified and the changes to Endeca will be made when programming time permits.

   Marlene will also be working with Bilal on making order records appear in Endeca. Currently, if a record has '0' items attached, it will not display in Endeca. We would like to change that and make order records appear in Endeca, as it does in the old library catalogue where an item was listed as being "on order" for a particular library.

4. Other Business

There is a bug in Sirsi browse indexes. Hyphens do not file correctly. The hyphen is disregarded but not replaced with a space, so a name such as *Lloyd-Webber* will be indexed as *Lloydwebber* and file accordingly. This problem also extends to Authority Files. Be aware of this and check

for it to avoid creating duplicate authority records. The same problem will likely exist in the Endeca browse.

CIP Records: many of the first records found are based on pre-publication information. If you're not planning on updating a CIP-level record to an "item-in-hand" level, Alastair suggested leaving the Date Cataloged (on Control tab) as NEVER. This will allow it to be replaced automatically by a later batch-loaded full LC record supplied by Coutts, in the event that Robarts or another UofT Coutts customer acquires the same title. Changing NEVER to TODAY will prevent any record from being overlaid during a batch load.

MARC Holdings 852 field call number not showing in Endeca: Will pass this issue on to Sian.