Metadata Working Group
Minutes
August 15, 2011

Present: K. Ananthan (co-chair), I. Beard (via videoconference), M. De Fino, R. Marker, J. Otto, L. Sun, M.B. Weber (co-chair and recorder), C. Zimmerman

Agenda:
1. SOLR/Lucene
2. RUcore Portal IDs
3. RUcore project list

1. SOLR/Lucene- Members discussed the advantages SOLR/Lucene search software will offer and MDWG's contributions to prepare for the migration from Amberfish. Advantages include faceted search, highlighting search terms in results lists, and Unicode compliance. SOLR/Lucene is highly configurable and provides immediate indexing, eliminating the need for overnight indexing. NJVid is currently using SOLR/Lucene.

M. De Fino and J. Otto will serve as liaisons to the RUcore User Services Working Group chaired by R. Marker and L. Langschied to determine the best configuration of elements for searching and sorting results. User Services' feedback is needed to determine what works best.

SOLR/Lucene will be implemented in a separate release which will be an indexing and display release. R. Marker will prepare specifications for the configuration and submit them to J. Triggs.

The timeline for deliverables and target dates was discussed. SOLR/Lucene was installed on the development serving on July 29. A prototype will be installed on the development server using the search portal code and minimal indexing schema on August 19. The prototype will be available on the development server using vetted indexing schema on September 15. The code is expected to be completed by October 15. Testing and bug fixes will take place October 15-November. The target date for the code to be delivered to production is December 15.

SOLR/Lucene offers the possibility of a single index or multiple indexes (the current search software provides a single index). It was decided that we should see how the new search software plays out before the need for multiple indexes is determined.

The Sirsi index for the RUL library catalog is done annually. The same will need to be done for SOLR/Lucene. The rebuild applies to both additions and deletions.

Use of SOLR/Lucene could be extended to federated searching.

2. RUcore Portal IDs- MDWG was charged by G. Agnew to assign portal IDs to RUcore collections following the implementation of the three RUcore portals: research, scholarship and

digital collections (see the July 2010 MDWG minutes:
http://rucore.libraries.rutgers.edu/collab/ref/min_mdwg_20100719.pdf).

There is a Fedora table to register each object in a portal. There is a table for each of the new proposed portals. Each table has list of IDs for a given portal. Collection IDs are provided in these tables. The lists are currently not turned on.

This will require an inventory of ingested collections, determination of in which portal a collection will reside, and determination of whether there is overlap between portals. C. Zimmerman will work with R. Marker on this inventory. For example, ETDs will reside in the scholarship portal while an Equine Sciences collection could reside in the research portal. We may find that entire collections belong in more than one portal. There may also be cases when individual pieces of a collection will belong in more than one portal. There currently are four major projects in the research portal.

The WMS lacks a feature to assign/select a portal ID. This functionality is needed and could be added in DLR Edit.

A list of what collections are currently in which portal will be obtained from Chad Mills. The collections will be reviewed to determine if they are in the correct portal or if it's appropriate for them to reside in more than one portal. R. Marker and K. Ananthan prepared a reorganization of the portals. MDWG will report back to CISC regarding the assessment of the collections and portals. This won't require a separate release. It only needs to be turned on and no further re-indexing will be required.

3. RUcore Project List- M.B. Weber distributed a list of 126 RUcore projects prepared by Peter Konin for MDWG members to review. This is the first step toward getting a handle on our projects. P. Konin has written an overview of requirements for a registration/notification system. The plan is to do this once and it should be self-functioning in the future. Tibor Purger has charged a group to investigate project management tools. P. Konin's overview isn't intended as such. How this list will work with a project management tool isn't clear.

K. Ananthan is a member of the group charged by T. Purger and reported that they looked into web-based read-only collaborative tools. The other members of the group are Tracey Meyer, Isaiah Beard, and John Brennan. They looked at Microsoft Project, Team Work, and SmartSheet. There's a considerable learning curve with Microsoft Project; the other tools are more straightforward and have less of a learning curve. The intent is to integrate this tool into our email system and hook into LDAP to avoid requiring another password, URL, ID, etc. Subscription versus a local solution was also discussed. Cost analysis and reporting features are also desirable. The group's suggestions have been presented to T. Purger; a decision hasn't yet been made.

4. Round Robin
M. De Fino, J. Otto, K. Ananthan, and C. Zimmerman met with Chad Mills to discuss context metadata. The decision on how to move forward is now up to C. Mills. C. Zimmerman gave C.

Mills metadata for the Henry Rutgers project prior to the kickoff of the Milton Exhibit so that he could work with resources that had been ingested in the WMS test system.

L. Sun will send MDWG members a link to the China Boom project which is currently in the WMS test system.  She's seeking our feedback.