

**BCC2010/MD/3**

**Report from the ALCTS Metadata Interest Group  
and other metadata-related meetings  
Jenn Riley, Chair, Metadata Subcommittee, MLA BCC**

**ALCTS Metadata Interest Group** (Formerly Networked Resources and Metadata Interest Group) Annual Meeting, June 27, 8:00-10:00 AM

The ALCTS Metadata Interest Group (MIG) held its regular business meeting, and invited two external speakers to discuss metadata efforts that expand into new areas. The first speaker was Rebecca Guenther, who discussed the <http://id.loc.gov> service, which exposes LC-maintained controlled vocabularies as Linked Data. The service provides access to several different types of vocabularies, including pick lists, code lists, and thesauri. It was developed for a variety of reasons: facilitating the development and maintenance of LC vocabularies, making the vocabularies themselves openly available to both human and programmatic users, providing comprehensive information about the controlled terms in these vocabularies, allowing LC to experiment with Semantic Web technologies and Linked Data, and exposing LC vocabularies to wider communities. LCSH was posted in the service in April 2009, and the TGM, MARC Relator Codes, and other vocabularies were added in April 2010. Bulk download of these vocabularies is also available in several serializations. Next steps include adding new vocabularies such as language codes, country codes, various PREMIS vocabularies, and LC name authorities; and an OWL representation of MADS to allow the specification of subject subdivisions, which the current SKOS encoding cannot do.

The second speaker was Sara Russell Gonzalez from the University of Florida on the collaborative VIVO project. VIVO has several goals: facilitating libraries' involvement in the research process, helping researchers to find collaborators and stay abreast of new developments, and providing data on researcher activities to educational and grant administrators. VIVO itself is an open source Semantic Web application to build researcher profiles. These profiles are built automatically by a variety of data sources, and can be updated manually by researchers themselves. VIVO operates on a library support model, using existing library skills and infrastructure to feed data to the application. For representing researcher information, a core VIVO ontology has been developed, building on FOAF and BIBO.

Following these presentations, a business meeting was held. MIG officers serving for the coming year are: Chair—Jennifer O'Brien Roper, University of Virginia; Vice Chair—Michael Dulock, University of Colorado at Boulder; Program Co-Chairs—Rhonda Marker, Rutgers University and Amanda Harlan, Baylor University; Publications Co-Chairs—Kevin Clair, Pennsylvania State University Libraries and Shawn Averkamp, University of Alabama; Secretary—Mary Aycock, University of Missouri—Columbia; Blog Coordinator—Kristin Martin, University of Illinois at Chicago; CC:DA Liaison—Nathan Putnam, George Mason University; LITA Liaison—Susan Cheney, Saint Joseph's University.

The MIG also sponsored the session “Converging Metadata Standards in Cultural Institutions: Apples & Oranges,” held June 26, 2010, at 8 AM. This session covered metadata initiatives that cross library, archive, and museum boundaries. The first presentation was by Danielle Plummer, Coordinator for the Texas Heritage Online (<http://www.tsl.state.tx.us/tho/>) Project. This project provides training for staff members at institutions across Texas, from library special collections, archives and museums, to contribute metadata to a statewide digital resource delivery system. Training was done mainly in 2-day in person modules, with an instructor team made up of a professional trainer from Amigos and a practitioner. Course modules were developed for digital project planning and management basics, digital library systems and applications, metadata standards and crosswalks, controlled vocabulary and thesaurus design, and digital preservation management. Work has begun to migrate the training content to an online instructional medium.

The second speaker at the Apples & Oranges session was Ching-Hsien Wang, Chief Information Officer at the Smithsonian Institution. Ching-Hsien described a “one-stop discovery system” that provides access to records for content throughout the various Smithsonian museums and libraries. The system provides access to 4.6 million records, 445,000 images, and derives from 40 different data sources. It features both a keyword search and a controlled-vocabulary driven faceted browse. The system is built on the premise that the source data is designed according to the

professional traditions of the unit that created it, and that these traditions need to be respected. A long processing queue normalizes metadata, and unfortunately it does not appear that the many scripts written to perform these tasks can be easily reused by other institutions.

The third and final speaker at the Apples & Oranges session was Joyce Chapman, Library Fellow at North Carolina State University. Joyce presented work she performed to study in detail the cost vs. value of various aspects of EAD creation, inspired by the ALCTS Task Force on Cost/Value Assessment of Bibliographic Control (<http://connect.ala.org/node/106017>). She collected data on the time needed to generate abstract, bioghist, scopecontent, subjects, and collection inventory data in EAD, and then interviewed researchers to determine the relative value of these elements in relation to the FRBR Identify user task. While a few researchers valued the bioghist element, it was the lowest ranked option by researchers for identification purposes. Collection inventories were the most expensive but the most valued, and abstract the second-most valued but least expensive. This research has some admitted drawbacks, for example only studying a single user task and the overlap in meaning of scopecontent and abstract in EAD. This methodology could be expanded to analyze other metadata formats and other user tasks.

The Metadata Interest Group maintains a blog at which summaries and links to presentation slides from other ALA meetings and presentations relevant to metadata issues are listed. This blog has recently moved, and now may be found at <http://www.alcts.ala.org/metadatablog/>. Presentation slides from MIG-sponsored sessions and additional information may be found in ALA Connect at <http://connect.ala.org/node/65847>.

### **ALCTS Preservation and Reformatting Section (PARS), Audio Metadata Task Force**

The BCC Metadata Subcommittee in late 2009 and early 2010 partnered with this Task Force to produce an online resource “Metadata Standards and Guidelines Relevant to Digital Audio” ([http://www.ala.org/ala/mgrps/divs/alcts/resources/preserv/audio\\_metadata.pdf](http://www.ala.org/ala/mgrps/divs/alcts/resources/preserv/audio_metadata.pdf)). *With the release of this resource, the Task Force’s charge is largely complete. The Task Force held an informal lunch meeting at the ALA 2010 Annual Conference to discuss wrapping up loose ends. Members believe the audio metadata standards document needs to have ongoing review and revision (either annually or semi-annually), and has asked the PARS Preservation Standards & Practices Committee to assume ownership of this document in to the future for this purpose. Task Force members will also look for opportunities to facilitate knowledge and use of this document, including sponsoring sessions at future ALA conferences.*

### **Other ALA Annual 2010 Sessions of Interest to MLA Members in the Area of Metadata**

The FRBR Interest Group met Friday, June 25, 2010, from 10:30 AM – Noon, with the following presentations:

- Jenn Riley, Indiana University, on work to represent FRBR data in an XML structure
- Yin Zhang, Kent State University, on characteristics of FRBR work sets
- Martha Yee, UCLA Film & Television Archive, on cataloging rules for an interpretation of FRBR involving depression, and modeling bibliographic data in RDF

Presentation slides from the session may be found at [http://presentations.ala.org/index.php?title=Friday%2C\\_June\\_25#FRBR\\_Interest\\_Group\\_Meeting](http://presentations.ala.org/index.php?title=Friday%2C_June_25#FRBR_Interest_Group_Meeting).

The MODS/MADS Editorial Committee sponsored a session through LITA on MODS and MADS: Current Implementations and Future Directions, Sunday June 27, 2010, from 10:30 AM – Noon, with the following presentations:

- Bill Leonard, Library and Archives Canada, on MADS in LAC’s Trusted Digital Repository
- McCallum and Rebecca Guenther, Library of Congress, on LC’s XML Data Store and Metadata for Digital Content Working Group
- Amanda Harlan, Baylor University, on the Texas Digital Library’s Electronic Theses and Dissertations activities and Vireo ETD management system

- Karen Miller, Northwestern University, on mapping from EAD to MODS

More information about most of these sessions may be found at ALA Connect: <http://connect.ala.org>.

---

*Last updated August 17, 2010*