

ALA Annual Meeting 2011
Report from the ALCTS Metadata Interest Group
and other metadata-related meetings
Jenn Riley, Chair, Metadata Subcommittee, MLA BCC

ALCTS Metadata Interest Group Meeting, June 26, 2011, 8:00-10:00 AM

The MIG meeting at ALA Annual 2011 featured a discussion table format, where participants gathered in small groups to explore areas of mutual interest. Following the small group discussions, all participants reconvened to report back on the topics explored.

The metadata quality control discussion table discovered that there are few shared or standard quality control strategies between institutions. A wide variety of practices exist, and most of these require mostly manual review. Specific problems identified that need community attention include ways to automate quality control processes, authority control for electronic theses and dissertations, making use of metadata from outside sources while still adhering to quality expectations, and encouraging software designers to incorporate quality control processes into tools used for metadata creation and digital object management.

The discussion of material formats focused on workflows that can span different media types, striking an optimum balance between metadata formats specialized for certain types of media and those for more general use. "Non-traditional" media often raise new challenges, such as access rights for a/v material, adhering to the creator's wishes for dissemination of the content, describing faculty papers, dealing with data management plans required by some granting agencies, and ensuring the individual creating metadata for an object understands the materials well enough to do so effectively.

The table devoted to metadata schema selection discovered that most implementers are using DSpace or CONTENTdm for digital object management, but are simultaneously being asked to manage metadata in formats those systems do not natively support, and are facing challenges in integrating metadata from other sources into the formats an institution has chosen for use locally. The group identified a number of factors to consider when selecting schemas for local use: how the collection will be arranged, what information is needed for local management of the materials, what content management system is being used, how much interoperability with other collections is necessary, who will be doing the metadata creation (and what is their background), and what collection curators' needs are.

The discussion of metadata migration identified some key challenges in migrating between metadata formats, including the handling of diacritics, the lack of best practices for migrating holdings data, and the difficulty of using lightweight tools such as Excel as part of a migration workflow. Participants expressed optimism that Linked Data and the work of the Library of Congress Bibliographic Framework Transition Initiative could significantly ease migration challenges.

The table devoted to independent metadata creation and manipulation tools reported on a few initiatives to utilize standalone tools for this purpose, including UVA infrastructure for name authority in

ETD records, and Texas A&M's full ETD submission system named Vireo. It was noted that many tools, protocols, and systems that might be potentially useful as part of a robust metadata creation workflow are not read for primetime, such as ORCID, XMP, and VIAF. Some favorite current metadata manipulation tools include XSLT, MARCEdit, Google Refine, and Recollection from the Library of Congress. The group concluded more concerted effort is needed to promote the development of reusable, modular, and scalable metadata management tools.

Following the roundtable discussion, a brief business meeting was held. New officers for 2011/2012 were elected: Teresa Kenna (University of Montana) as Vice-Chair/Chair Elect, Suzanne Pilsk (Smithsonian) and Jackie Blonigen (University of Missouri) as Co-Secretaries, Valentine Muyumba (Indiana State University) as Publications Chair, and Meghan Finch (Wayne State University) as Program Co-Chair. The Interest Group is still struggling with how to get (admittedly busy) people involved in sharing information. The ALCTS Newsletter is in transition, and the IG's blog does not have frequent volunteers to post. It would be useful to better understand how librarians look for metadata information in order to determine where best to put our efforts. Jenn Riley provided an update from the Music Library Association's BCC Metadata Subcommittee, including work on the music genre/form thesaurus, the survey issued to help plan the online music metadata clearinghouse, and the Linked Data session at the MLA 2012 annual meeting. The IG will explore session proposals for upcoming ALA conferences, including metadata for digital video, describing data sets and library support for data management plans, and training of catalogers to create non-MARC metadata.

Full minutes from this meeting will be posted shortly to <http://connect.ala.org/node/65847>.

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

The BCC Metadata Subcommittee has partnered with the ALCTS PARS Audio Metadata Task Force to create an online resource helping implementers to select appropriate metadata standards for digital audio projects, which was released in 2010. The Task Force did not have a formal meeting at Midwinter, but several representatives did meet informally to discuss the future of the group. As the charge is fulfilled and much awareness has been raised on audio metadata issues through presentations in various formats, the group agreed to recommend its disbanding to its parent committee, the ALCTS PARS Intellectual Access to Preservation Metadata Interest Group.

Other ALA Annual 2011 Sessions of Interest to MLA Members in the Area of Metadata

The ALCTS Metadata Interest Group sponsored the program "Sound Bytes: Audio Metadata Standards in Slightly More than Six Seconds," offered Saturday, June 25, at 10:30 AM. The first speaker, George Blood of George Blood, L.P. (formerly Safe Sound Archive), set the stage by discussing the various taxonomies for classifying metadata (descriptive, technical, preservation, administrative) and models for storing metadata: externally using a standard like Dublin Core or AES57, or embedded in a portion of the

file such as the <bext> or <axml> chunk (both European Broadcasting Union standards). A file and its associated metadata are stored together in a Submission Information Package (SIP). Blood cautioned against inventing new standards for metadata “square pegs”; rather, existing standards should be used for the elements that fit, supplemented with locally-defined fields for those that do not.

The next speaker was Mike Casey of Indiana University, who presented information about two emerging standards from the Audio Engineering Society (AES). The first, AES57 (*Audio object structures for preservation and restoration*), is a draft standard for structural and administrative metadata applicable to analog and digital audio files, whether source or derivative. The closing date for the three-month comment period was June 18, and once concerns are addressed, AES57 will be adopted. AES-X098C (*Administrative metadata for audio objects - Process history schema*) is a draft standard for documenting the transformative processes applied to audio objects—the “who, what, when, where, and why” of digitization—and it is still under development. Casey concluded his talk with a demonstration of IU’s software application Audio Technical Metadata Collector (ATMC), which is used by the Archives of Traditional Music.

The final speaker was Jane Johnson Otto of Rutgers University. Otto discussed efforts to integrate AES-X098B (*Administrative and structural metadata for audio objects*) into their digital object workflow management system, OpenWMS. While it was sometimes difficult to adapt an audio standard to fit the wide variety of objects hosted in the RUCore community repository, there are many benefits: AES is comprehensive and granular, meaning the metadata can be repurposed; both data elements and vocabularies are included; it accommodates both digital and analog formats, including those with physical carriers and those that exist as streams of bits; it has a rigorous delineation of metadata types that make it compatible with METS; it is expressed as XML and supports segmenting and long-term storage. For additional information, Otto referred program attendees to her article “A Sound Strategy for Preservation: Adapting Audio Engineering Society Technical Metadata for Use in Multimedia Repositories,” published in *Cataloging & Classification Quarterly*, volume 48, issue 5, 2010.

The MLA BCC Metadata Subcommittee /PARS Audio Preservation Metadata Task Force document “Metadata Standards and Guidelines Relevant to Digital Audio “ was distributed.

[Note: Summary from Sound Bytes session contributed by Deb Kulczak, MLA BCC Metadata Subcommittee member.]

Linked Data was an overwhelmingly strong theme of the ALA Annual 2011 conference, with a large number of sessions either devoted to or including this topic. A few (though certainly not all) are highlighted here. The CCS Executive Committee Forum featured the theme “Turning Catalogers into Semantic Web Engineers, or, Out of the Catalog Drawer and onto the Internet Highway.” Speakers at this session were Karen Coyle, Gordon Dunsire, and Ed Jones. This session was a look at how libraries are currently producing Linked Data, and how well our efforts fit into overall Semantic Web technologies. Presentation slides are available from <http://connect.ala.org/node/136967>. “Linked In: Library Data and the Semantic Web,” featured speakers Ross Singer and Eric Hellman. This session provided a more critical analysis of libraries’ Linked Data efforts, looking at them from a perspective of what would be

most useful to the wider community rather than starting from an assumption that Linked Data is simply a new encoding of existing library data, and demonstrating the value of actually making links between multiple data sources instead of simply making one's own data available. Hellman's presentation slides are available from http://www.gluejar.com/public_files/why_bother_clean.pdf. The LITA Linked Library Data Interest Group held its inaugural meeting at this conference, a discussion session designed to allow the community to self-organize around Linked Data implementation topics. Ideas discussed included outreach and education, working demos, going beyond bibliographic and authority data, making ties outside of the library community, a focus on software tools, speed-dating to match existing data with those who have specific needs, creating people profiles, developing use cases and functional requirements, and providing training materials. The Interest Group has created a listserv (<http://lists.ala.org/sympa/info/lita-lld>) and is using space within the LITA wiki to document its work and engage the community (<http://wikis.ala.org/lita/index.php/Linkddata>).

An additional list of sessions from ALA Annual 2011 of interest to metadata specialists may be found at <http://www.alcts.ala.org/metadatablog/2011/06/ala-annual-2011-best-bets-for-metadata-librarians-and-call-for-bloggers/>.