ALA Annual 2017

Reported by Tracey Snyder (Cornell University), Chair, MLA Cataloging and Metadata Committee

PCC At Large

Judith Cannan of the PCC Secretariat announced that a new PCC Directory is in the works and that implementation is due to begin in mid-September. Implementation will be phased because there are about 900 member institutions. Members will receive emails by November 1 with instructions for updating profiles and adding contact persons for the PCC programs. Members who haven't made any contributions for a year will be contacted about this. There may be some changes in how statistics are gathered; stay tuned for details. For funnels, such as the NACO-Music Project, there will be a profile in the directory, with member institutions listed, and the funnel coordinator will compile statistics reported by the contributors.

Judith spoke about the results of the PCC survey on strategic planning. There were many comments about communication, e.g. the need for a better website and a central place to access reports and documentation. Judith solicited input on topics such as Linked Data and identity management for inclusion in a forthcoming strategic plan document, and some attendees stressed the importance of moving forward in both areas. It was suggested that PCC could have a role in facilitating the interaction between MARC and Linked Data and in leveraging identity management work being done through ISNI, ORCID, etc.

PCC Participants Meeting

This program consisted of three presentations on a theme of subject access in a Linked Data environment. Nancy Fallgren from the National Library of Medicine spoke about MeSH (Medical Subject Headings) controlling of valid pre-coordinated headings and the work NLM has done to create MeSH RDF and to add $0 with MeSH URIs to NLM bibliographic records. Assigning URIs for infinite combinations of pre-coordinated strings is a significant challenge, and the solution being implemented is to assign URIs to pairs of MeSH terms. For example, the MeSH term "Diabetes Mellitus" and the subdivisions "genetics" and "immunology" each have a URI, and the pairs "Diabetes Mellitus $x genetics" and "Diabetes Mellitus $x immunology" each have a URI. Nancy posed some questions as food for thought: At what point do URIs for pre-coordinated strings become too precise to be useful for finding resources? At what point does assigning URIs to all pre-coordinated strings become excessive in terms of cost/benefit?

Jodi Williamschen from NDMSO (Network Development and MARC Standards Office) at the Library of Congress spoke about LCSH (Library of Congress Subject Headings) in a Linked Data environment, posing similar questions: In a large and ever-growing vocabulary such as LCSH, is it feasible to assign URIs to every possible combination of terms? If you split out time period (which is notated different ways depending on context, e.g. 21st century, 1914-1918, etc.) from the main topic, is meaning lost?
Naun Chew from Harvard (formerly from Cornell) spoke about the potential he sees in FAST (Faceted Application of Subject Terminology) and Cornell's work with OCLC on FAST implementation. Cornell uses FAST terms to populate subject facets in a Blacklight-powered public catalog, and select catalogers use FAST for subject analysis of certain sets of materials. While LCSH has clear strengths, including its longevity and degree of support, FAST's strengths include ease of use by catalogers, compatibility with legacy data, relationship to LCSH, suitability for non-alphabetical browse-based discovery, and Linked Data readiness. Naun pointed out the need for an explicit business model and governance plan and for retrospective conversion of LCSH to FAST.

Linked Library Data Interest Group

The topic of this program was the applicability and practicality of BIBFRAME. Brian Rennick from Brigham Young spoke about a case study comparing the effectiveness of conversion of MARC records to BIBFRAME (via the Library Link Network service from Zepheira and SirsiDynix) with that of search engine optimization (via an initiative launched around the same time to expose BYU's catalog records). Server logs indicate that more catalog referrals are resulting from SEO than from the Library Link Network service, although both methods are increasing the number of catalog hits coming from Google. Catalog records for a portion of BYU's collection are appearing on the second page (or later) of Google results, or on the first page if "BYU" is included in the search. It is not clear why more catalog records are not appearing in Google search results. Future plans include adding schema.org tags, enriching records with Linked Data from other sources, ensuring adherence to Google best practices, and more closely evaluating the benefits of the still-new Library Link Network service.

Philip Schreur from Stanford spoke about one of Stanford's projects in the Linked Data for Production (LD4P) Mellon grant, namely the conversion of a copy cataloging workflow to a new workflow rooted in Linked Data (known as Tracer Bullet #1). The new workflow, which is intended to mimic and improve upon the existing workflow without adding extra staff time or money, begins with ingestion of MARC records supplied by Casalini Libri. These are converted to Linked Data (BIBFRAME 2.0), which is then made discoverable in SearchWorks (Stanford’s Blacklight-powered public catalog). Philip acknowledged that the separation of BIBFRAME and MARC from each other within the same discovery environment is not ideal, but it is a way to demonstrate the Linked Data work that has been done. Next steps will explore the reconciliation of data upon conversion (considering data of varying quality coming from different sources) and the sharing of RDF-based original cataloging (Tracer Bullet #2).

Tiziana Possemato from Casalini Libri spoke about the continued development of the SHARE-VDE (Virtual Discovery Environment). (See share-vde.org.) In a context where libraries have different cataloging practices and Linked Data protocols are emerging, SHARE-VDE is a prototype of a new RDF-based discovery environment with benefits envisioned for librarians and end-users. This project is driven by the library community, with 16 participating institutions, and is separated into three phases. In the first phase (October 2016-January 2017), bibliographic
and authority records were converted to BIBFRAME 1.0 and enriched with URIs. In the second phase (March-September 2017), each library's catalog is being converted to BIBFRAME 2.0, and more URIs are being added. The prototype interface demonstrates the reconciliation and clustering of data from various sources in various forms related to the same entity. Search Vivaldi to see an example.

Cataloging Norms Interest Group

The theme of this program was "The Changing Metadata Arena and Its Practitioners." Jennifer Liss from Indiana University was the featured presenter, and she spoke about the work of the PCC Task Group on Identity Management (as contrasted with the more expensive, labor-intensive, and exclusive practice of authority control), especially as it relates to the work of the PCC Task Group on URIs in MARC. PCC is exploring a NACO Lite option, which could be a means for other cultural heritage institutions to contribute information without it having to conform to RDA or MARC, and organizing an ISNI pilot in which testers from PCC institutions will experiment with using the ISNI system as an alternative to NACO (beginning this summer). A shift to identity management, where we mint URIs as identifiers for entities rather than create unique character strings, would mean not having to worry about bibliographic file maintenance for changed headings or the presence of an author's middle name on a dissertation (and absence of the middle name on later resources). Slides from Jennifer's presentation and the other presentations in this program are available at: http://connect.ala.org/node/268276.

CaMMS Forum

This year's forum was called "Power That Is Moral: Creating a Cataloging Code of Ethics." Elizabeth Shoemaker of Emory University presented examples of moral/ethical dilemmas faced by catalogers (delegating the cataloging of explicit or offensive materials, fielding requests to reclassify materials, etc.), introduced the cataloging code of ethics in development (against a backdrop of the existing ALA Code of Ethics and the ALCTS supplementary guidelines), and offered a rationale for creating such a code — a code helps place dilemmas in context and guide actions. Dr. Hope Olson, Professor Emerita from University of Wisconsin – Milwaukee, spoke about various ethics theories and how they might inform a cataloging code of ethics. Slides from both presentations as well as notes from the open discussion that followed are available at: http://connect.ala.org/node/265990.