Report on the American Library Association’s Committee on Cataloging: Description and Access, ALA Midwinter Conference, Boston, Massachusetts, USA, 2016 January 9 and 11

Submitted to the Standing Committee of the IFLA Cataloguing Section by the IFLA Cataloguing Section Liaison to ALA CC:DA

The American Library Association’s Committee on Cataloging: Description and Access (CC:DA) met at the ALA Midwinter Conference in Boston, Massachusetts, USA, on Saturday, 2016 January 9, 1:00-5:30 P.M.; and Monday 2016 January 11, 8:30-11:30 A.M. The full agenda of the meeting is at http://alcts.ala.org/ccdablog/?cat=33.

CC:DA Chair Ms. Dominique Bourassa (Yale University) reported on motions and other actions taken by the committee between July and December 2015 (http://alcts.ala.org/ccdablog/wp-content/uploads/2016/01/CCDAChair2015-201602.pdf).

Library of Congress Representative Mr. David Reser reported on activities and news from LC (http://alcts.ala.org/ccdablog/wp-content/uploads/2016/01/LC-2016-01.pdf), including some of these highlights:


- Organizational realignment at LC allows hiring from outside of the library for the first time in years, including roughly thirty vacancies within Acquisitions and Bibliographic Access (ABA) alone.

- There will be additional RDA cleanups following the Authorities Phase 3B changes, the timing of which is yet to be determined. Phase 3B will recode AACR2 authority records as RDA when the 1XX fields contain no elements contrary to RDA and will enhance authority records with additional data elements such as ISNIs in field 024.
Ms. Bourassa noted that many cartographic headings currently look correct but actually need to be changed and should be marked with the addition of 667 fields until they can be evaluated. Mr. Reser said that this is not currently planned but made note of it.


- Document numbering will switch from the designation “JSC” to “RSC.”
- Instead of referring to “constituencies, the RSC will refer to “communities.”
- The “Working Principle” ([http://www.rda-rsc.org/node/229](http://www.rda-rsc.org/node/229)), the moratorium on RDA changes in areas that are likely to be affected by the FRBR-LRM and/or the Consolidated ISBD, has been extended through 2016. This was an unintended result of ALA/43 ([http://www.rda-rsc.org/6JSC/ALA/43](http://www.rda-rsc.org/6JSC/ALA/43)), which intended to revamp RDA Appendix K, Relationship Designators: Relationships Between Persons, Families, and Corporate Bodies. The scope of RDs will need to be re-evaluated. Any RD proposals in process during 2015 can continue, but any new RDs not previously considered are now on hold. Communities should continue to discuss and even propose RDs, but they will not be implemented until the moratorium ends.
- Recording the fuller form of name in RDA 9.5.1.1 will use the term “diminutive” rather than “nickname,” with the former to be defined.
- The RSC Places Working Group will consider the ramifications of the new “Place” entity in FRBR-LRM.
- Further work is needed on clarifying the distinction in RDA Chapter 2 between “record” and “transcribe.” The latter applies only to Manifestation elements and can be used only for self-describing resources.
- Gender will be retained as an RDA element, but the existing vocabulary in RDA 9.7.1.3 will be deprecated, allowing the creation of RDA-compatible vocabularies by individual communities.
- The report of the RSC Fictitious Working Group was rejected because it conflicts with the definitions of “person” in the FRBR-LRM and in FRBRoo. The effect of this is to reverse the trend of treating fictitious entities as persons/agents. The RSC acknowledges that the issue is contentious, but that it concerns the role of persons as creators, not about subject entities. As RDA is revised in light of FRBR-LRM, the role of the Nomen entity will need to address fictitious entities, possibly having statements of responsibility in a Manifestation associated with a specific Nomen by the use of the appropriate Relationship Designators.
- Ms. Glennan further reported on the outcome and next steps regarding 6JSC/ALA/40, “Revision to RDA 3.1.4, Resources Consisting of More than One Carrier Type and RDA 3.4.1.3, Recording Extent.” The proposal revealed many more problems under the surface. A new task group will
be formed because of the change of focus, with CC:DA and the Canadian CCC working together on a discussion paper to be presented in 2016.

Chair Ms. Bourassa discharged the Task Force to Investigate the Instructions for Recording Relationships in RDA and the Task Force on Relationship Designators in RDA Appendix K. She is waiting for the RSC to create its group to replace the Task Force on Machine-Actionable Data Elements in RDA Chapter 3. A CC:DA new group is being formed on the FRBR-LRM.

Program for Cooperative Cataloging (PCC) Liaison Ms. Lori Robare (University of Oregon) reported on PCC activities (http://alcts.ala.org/ccdablog/wp-content/uploads/2016/01/PCC-2016-01.docx). The Standards Committee has been working on a proposal for the option of including Creative Commons statements in MARC Bibliographic field 540. The Training Committee’s Series Policy Task Group has nearly completed its work on the series-related LC-PCC Policy Statements and the sections of DCM Z1. The RDA Sample records have been extensively updated. RDA refresher courses are now available at http://www.loc.gov/catworkshop/RDA%20training%20materials/da-refreshers.html. The Standards and Training Committees have been jointly drafting guidelines on Relationship Designators in Authority records.

The American Association of Law Libraries (AALL) will present a proposal regarding names of international courts in RDA 11.2.2.21. There had never been any adequate instructions in either AACR2 or RDA. The group will also investigate a reference to this RDA instruction from RDA 11.2.2.14.11.

Ms. Glennan reported on how the changes in RDA governance affect ALA (http://alcts.ala.org/ccdablog/wp-content/uploads/2016/01/RDAGov-2016-01.pdf). The Committee of Principals is now the RDA Board. There will be more international stakeholders and more cultural heritage organizations involved. There will be more reliance on Working Groups, which will mostly be finite “task and finish” groups, in which ALA members will continue to participate. The future North American structure will include the current voices of ALA, LC, and CCC, will continue to rely on the expertise of CC:DA, SAC, and other such advisory groups, with perhaps a new “lightweight” layer between ALA and the RSC. Ms. Glennan suggested something along the lines of a “North American RDA Committee” (NARDAC), with representatives from ALA, LC, and CCC, but with a North American RSC representative chosen from among its ranks. The extra layer of bureaucracy and the attendant slowing down of an already glacial process are recognized problems. Questions remain about how NARDAC would govern itself, if consensus would continue to be the operating model, and the need for succession planning. For the time being CC:DA’s work will continue as usual, although these changes allow CC:DA to reimagine itself and its work. Creating best practices and application profiles beyond LC and PCC and more prominence for the perspectives of public libraries are possibilities. Choices will have to be made about the levels at which participation and representation will take place, at the CC:DA level, the North American regional level, or the RSC international level.

RDA Steering Committee (RSC) Chair Mr. Gordon Dunsire presented “RDA Data Capture and Storage” (http://www.gordondunsire.com/pubs/pres/RDADataCap.pptx).
• “RDA is package of data elements, guidelines, and instruction for creating library and cultural heritage resource metadata that are well-formed according to international models for user-focused linked data applications.” The Registry provides the infrastructure. The Toolkit provides the instructions.

• RDA offers choices for recording relationships. Playing with the notion of the Buddhist Eightfold Path, Mr. Dunsire posited the following:
  
  o The Fourfold Path of Work, Expression, Manifestation, Item (WEMI) relationships (RDA 24.4): Identifier, Authorized Access Point (AAP, which excludes Manifestation and Item), Structured Description, Unstructured Description.
  
  o The Threefold Path for primary relationships among WEMI (RDA 17.4.2): Identifier, AAP (excludes Manifestation and Item), Structured Description.
  
  o The Twofold Path for relationships among Persons, Families, and Corporate Bodies (PFC) (RDA 29.4): Identifier, AAP.

• The new FRBR-LRM entities of Place and Timespan will need to be accommodated in RDA, as will the super-entity of Collective Agent, comprising both family and corporate body (in opposition to Single Agent/Person). It will have to be determined which of the four techniques will apply to each. Nomen encompasses Identifier, AAP, Variant Access Point (VAP), Structured Description, transcribed title, and so on.

• Structured Description is defined as “A full or partial description of the related resource using the same data that would be recorded in RDA elements for a description of that related resource presented in an order specified by a recognized display standard,” such as ISBD.

• RDA envisioned three database implementation models: (3) Flat-file (not linked, such as the card catalog); (2) Bibliographic and authority records (AAP/Identifier linked, such as in MARC records); (1) Relational or object database (fully linked, but only locally). A fourth scenario has developed since the publication of FRBR and will need to be developed within RDA: (0) Linked Data (fully linked, globally).

• RDA categorizes data elements in one of two ways:
  
  o Recorded Elements:
    
    ▪ Sources: Any (authoritative, recognized, etc.).
    
    ▪ Tasks: All (Find, Identify, Select Obtain, Explore).
    
    ▪ Entities: All.
  
  o Transcribed Elements:
    
    ▪ Sources: Manifestation (Item in hand).
    
    ▪ Tasks: Identify.
- Entities: Manifestation.

Therefore, the transcribed elements seem to be a logical subset of the recorded elements.

- Mr. Dunsire then went through a critique of “transcription” through Optical Character Recognition (OCR), pointing out OCR’s many shortcomings, particularly with early printing and the many quirks of typography involved. OCR would need to be “trained” to account for such conventions as the use of small capital letters for lowercase and the “long-S” character, which appears as the letter “F” to OCR. For the Identify user task, transcription from digital images would be the quickest and easiest, but either the user must be made aware of some of these quirky transcription rules or OCR would need to adjust. Crowdsourcing and other combinations of human and machine techniques would help. “Recording” for the user tasks generally excludes typographical errors, deliberate errors, and fictitious entities. Some recorded data support the FISOE user tasks, but RDA needs to accommodate these data better.

- Returning to the Buddhist Eightfold Path, Mr. Dunsire suggested an undetermined N-Fold Path:
  - 1. Unstructured string:
    - 1) Exact transcription (OCR or born-digital).
    - 2) Transcription using RDA guidelines.
    - 3) Data recorded from another source.
  - 2. Structured string of delimited sub-values:
    - 1) Access point.
    - 2) Structured description.
  - 3. Structured string:
    - 1) Identifier.
  - 4. URI of entity, including Noman [linked data path/“path with no name”]:
    - 1) URI/URL of digital image.

All these paths are available for describing related entities. The same paths may describe the entity in focus.

- Upcoming development within RDA will need to focus on the various methods of recording data, including general guidance on techniques (the aforementioned Fourfold Path), general instruction sets for specific entities and element categories (such as attributes and relationships), and specific instructions for specific elements. The RDA Registry also needs further development for the sake of Linked Data.

- The new entities introduced by the FRBR-LRM will also require new high-level relationship elements and new cross-entity Relationship Designators. “Res,” formerly “Thema.” is the uppermost new entity. The outdated element set views in the RDA Toolkit need to be replaced.
by entity views that will focus on each RDA entity and its elements. This will act as a ready-reference tool for all of the elements and their related instructions. This will be the goal of the ongoing RDA Toolkit reorganization: appendices and tabs; Vocabulary Encoding Schemes; RDA Reference (entities, elements, terms, and beyond); the glossary (perhaps incorporating MulDiCat), translations; Policy Statements and application profiles; Entity views; RDs. The international, cultural heritage, and Linked Data communities have needs that will need to be considered. Work is needed on the primary (WEMI) and secondary (PFC) entities, how much structure and detail are appropriate on descriptions, Nomen control (and its relationship to traditional authority control), and Relationship Designators.

ALA Publishing reported that during Fiscal Year 2015 (ending in August), peak RDA Toolkit subscriptions were 3100, but are now down to 2800. The renewal rate is down from about 90% to 81%, but the number of users is up to about three per subscription. Revenues are just short of projections. Some 752 print units have sold, eight electronic books, and 329 print Spanish translations, which bodes well for the new governance model. There will be no print edition in 2016, but they are looking toward a schedule of every other year. *RDA Essentials* is coming in April 2016, but will have revisions quickly because it is current only to April 2015. There were three RDA releases during FY 2015. DCRM will probably be added to the Toolkit in 2017. Currently international sales total about 35%, with about 55% of users outside the United States.

MARC Advisory Committee Liaison Mr. John Myers (Union College) reported (http://alcts.ala.org/ccdablog/wp-content/uploads/2016/01/MAC2016-1-prelim.pdf) on the results of the MAC meetings. My own report as OCLC Representative to MAC follows:

MARC Advisory Committee (MAC). Saturday, 2016 January 9, 8:30-10:00 a.m.; and Sunday, 2016 January 10, 3:00-5:30 p.m. OCLC Representative.

The MAC agenda is available at http://www.loc.gov/marc/mac/mw2016_age.html. Here are the summaries of each of the two proposals and sixteen discussion papers plus the outcomes:


Summary: This paper proposes defining new values for some 007 field positions in the MARC 21 Bibliographic Format to better accommodate digital reproductions of sound recordings. Changes to 007c/01 to better accommodate certain types of commonly used storage devices are also included.

Outcome: The definition of “sound recording” for 007/00 suggested by the British Library (BL) was accepted: “A storage medium containing recorded sound or a representation of a musical composition for which sound can be mechanically reproduced, such as a piano roll.” The definition of 007/03 code “n” for “not applicable” was accepted as: “Speed is not applicable to *remote* digital sound recordings because it pertains to calculations specific to physical
aspects of carriers.” The definition of 007/10 code “n” for “not applicable” was accepted as: “Kind of material is not applicable to remote digital sound recordings because it pertains to characteristics specific to physical aspects of carriers.”

MARC Proposal No. 2016-02: Defining Subfield $r and Subfield $t, and Redefining Subfield $e in Field 382 of the MARC 21 Bibliographic and Authority Formats (http://www.loc.gov/marc/mac/2016/2016-02.html).

Summary: This paper looks at the need to establish two new subfields within MARC Field 382 (Medium of Performance), one to describe the total number of ensembles and the other to describe the number of solo instruments performing with the ensembles. Also suggests a clarification of the field’s existing subfield $e (Number of ensembles).

Outcome: Accepted without changes.


Summary: This discussion paper presents the need for subfields $3 (Materials specified) and $5 (Institution to which field applies) in Field 382 (Medium of Performance) of the MARC 21 Bibliographic Format.

Outcome: The appropriateness of subfield $5 sparked the most discussion. This will come back as a proposal.


Summary: This paper presents suggestions for clarifying four code values in Field 008/20 (Format of Music) in the MARC 21 Bibliographic Format, in order to bring the code values in line with RDA and clarify their use.

Outcome: It was suggested that “piano score” may deserve its own code. This will come back as a proposal.


Summary: This paper examines the need to distinguish music and videorecording distributor numbers from music and videorecording publisher numbers recorded in MARC field 028 (Publisher Number) and suggests defining a new first indicator 6 for Distributor number to accomplish this. The paper also
suggests minor clarifying changes to MARC field 037 (Source of Acquisition) regarding the numbers to be recorded there.

Outcome: Both defining a new First Indicator for field 028 and refining the definition of 037 were well received. Guidance about how to determine the distinction would be an issue for best practices. This will return as a proposal.


Summary: This paper discusses the definition of subfield $0 (Authority record control number or standard number) in linking entry fields 760, 762, 765, 767, 770, 772, 773, 774, 775, 776, 777, 780, 785, 786, 787 in the MARC 21 Bibliographic Format.

Outcome: Subfield $0 properly holds the identifier for the entity in the field. Subfield $4 already deals with relationships, so perhaps broadening its definition to include the URI for the relationship makes more sense than forcing subfield $0 to do double duty to identify both the entity and its relationship. The British Library should be brought into the PCC URL Task Force to prevent a clash of proposals. This will be reworked and return as a proposal from the Task Force and the BL.


Summary: This paper discusses expanding the scope of subfield $w (Bibliographic record control number) in the MARC 21 Bibliographic and Authority Formats.

Outcome: This will return as a proposal.


Summary: This paper proposes to add $2 (Source of term) and $0 (Authority record control number or standard number) to field 753 (System Details Access to Computer Files). Subfield $2 will allow the vocabulary used for the terminology in subfields $a (Make and model of machine) and $c (Operating system) to be documented and the subfield $0 (Authority record control number or standard number) would allow the URI of the vocabulary term to be entered.
Outcome: Enabling better bibliographic control and identification of computer game platforms and their operating systems, the GAme MEtadata and Citation Project GAMECIP) is developing a controlled vocabulary for the platforms. The vocabulary has been assigned MARC Code “gcipplatform”. The 753 would be repeated when there are multiple platforms or operating systems. The discussion paper was allowed to be considered as a proposal and was accepted unanimously.


Summary: This paper proposes broadening the usage of field 257 (Country of Producing Entity) to include autonomous regions so that regions with strong film cultures can be used in this field. This will involve changing the name of the field and changing the field definition and scope.

Outcome: Historically, the field had been used to record the name of the country or countries from which the financing for a film had come, in keeping with film industry traditions. This slight broadening to include autonomous regions is intended to cover, in particular, the four autonomous regions that have had significant film industries: Hong Kong, Palestine, Puerto Rico, and Taiwan. OLAC would establish best practices for using the name of the region at time of creation. In addition to the field definition, the definition of subfield $a would also need to be adjusted. This will come back as a proposal.


Summary: This paper proposes removing the sentence “Dates contained in subfield $k may not be coded elsewhere in the formats” currently in subfield $k (Beginning or single date created) in field 046 (Special Coded Dates) of the Bibliographic format.

Outcome: Sentiment was strong for removing the seeming restriction in subfield $k. I pointed out that the sentence slated for removal was highly ambiguous in its meaning, saying either that dates appearing in subfield $k are not allowed to appear elsewhere in the record or that it might or might not happen that such dates appear elsewhere in the record. This will come back as a proposal, with the field definition and other subfield definitions being similarly clarified if necessary.

Summary: This discussion paper presents options for coding of named events used as subject access points in the MARC Authority and Bibliographic formats.

Outcome: OCLC’s Mr. Robert Bremer presented this OCLC proposal, prompted by discussions with OCLC Research. The two options presented were: (1) Redefine the X11 fields to explicitly include all types of named events, and, (2) Define a new series of X47 fields for the coding of named events. The Deutsche Nationalbibliothek (DNB) suggested a third option of using the existing X50 fields, because entities in the X11 fields would mix events with agency (such as conferences) and events without agency (such as wars). A straw poll suggested a strong preference for Option 2, including subfield coding to accommodate places and dates, for consistency with the X11 fields. The British Library suggested the deprecation of X11 First Indicator, but that would require a separate proposal. This paper will return as a proposal.


Summary: This paper proposes defining field 347 (Digital File Characteristics) (R) for the MARC 21 Holdings Format to contain copy specific technical specification relating to the digital encoding of text, image, audio, video, and other types of data in the resource.

Outcome: The need for field 347 in Holdings records was widely acknowledged to allow differentiation among different presentation formats for electronic books, PDF versus HTML, different formats for streaming media, and so on. It was suggested that consistency among fields (Holdings 347 and 856, in particular) in the placement of subfields $3 for human readers would also be advantageous. Subfield $8 would also be useful for machine action. This paper will return as a proposal.


Summary: This paper explains why libraries from German speaking countries do not provide punctuation when content designation identifies an element sufficiently. It proposes coding to indicate the absence of punctuation redundant to field and subfield coding via a Leader position.

Outcome: Discussion focused on whether a solution is needed for both Bibliographic and Authority records. Omitting punctuation would apply only to
punctuation at subfield boundaries, not to all punctuation, and would need to consider RDA Appendix E. This will be reworked and returned as a proposal.


Summary: This paper discusses a way that information about matching two records can be expressed in the MARC Authority format.

Outcome: In the context of marking records for an appropriate merge, this idea does amount to temporary data. But in the context of recording information that certain records have been examined and found not to be proper matches, this corresponds to ideas that we at OCLC have discussed for years about Duplicate Detection and Resolution (DDR). This would mean expanding a proposal to cover Bibliographic as well as Authority records and making the field repeatable to cover multiple matches/non-matches. Subfield $5 may also be added for temporary data. In Authorities, “Do not confuse” notes would become redundant. This will come back as a proposal.


Summary: This paper proposes a way of giving a definition in a MARC Authority record.

Outcome: Because Authority field 668 had been defined and made obsolete in the past, field 667 was preferred. The DNB believed that field 680 (General Public Note) was defined too broadly and field 678 (Biographical or Historical Data) was too narrow, but the discussion revealed a preference for the use of 680 (and in some cases, 678) rather than the definition of a new field. This will return as a reworked discussion paper focusing on fields 680 and 678.


Summary: This paper proposes a way of coding which type of entity is described in a given MARC Authority record.

Outcome: The differentiation of types of entity would be in keeping with where the FRBR-Library Reference Model (FRBR-LRM) models are going. The variable field 072 option was generally preferred over a Leader element solution. This will come back as a proposal.

Summary: This paper proposes a way of coding the RDA Media Type and Carrier Type in the MARC Authority format.

Outcome: The DNB has unique resources that are accounted for in authority records. Because authority records represent names for the things (Nomen), not the physical things themselves (Res), there were many objections to this practice from the committee. It may, however, be in keeping with the FRBR-LRM distinction between manifestation singletons and manifestation multiples. It is recognized that the FRBR WEMI model does not work well for unique resources. There was a wider discussion about breakdown of the distinction between bibliographic and authority data, leading to the thought that perhaps something like format integration needs to happen between the bibliographic and authority formats. The DNB will take all this under consideration and return with a revised discussion paper.


Summary: This paper proposes a way of extending Leader position 17(Encoding Level) in the MARC Authority format.

Outcome: The DNB’s proposed new Encoding Levels were thought not be more local than universal. It was thought that perhaps an Encoding Level referring to a new code in field 042 reflecting more detailed authentication might work. That would require making field 042 repeatable and adding subfield $2, but then the subfield $2 would apply only to the newly-defined subfield rather than to entire field. This will return as a revised discussion paper.

CC:DA is now promised consistent microphone coverage. If meetings return to hotels from the convention centers, there will probably be no Wi-Fi, so we will continue to ask for convention center meeting spaces. Ms. Mary Huismann (University of Minnesota) will replace Ms. Tracey Snyder (Cornell University) as the Music Library Association (MLA) representative. The next CC:DA meetings will be on 2016 June 25 and 27 in Orlando, Florida, at ALA Annual.

Respectfully submitted by

Jay Weitz

Senior Consulting Database Specialist

Data Infrastructure and WorldCat Quality Management Division, OCLC

IFLA Cataloguing Section Liaison to ALA CC:DA

2016 January 29