Evolution of the RDA Toolkit and Its Impact on Catalogers

Kathy Glennan
Chair-Elect RDA Steering Committee
Head, Original & Special Collections Cataloging, University of Maryland
Transformation in Cataloging

• Conceptual models
  ▪ IFLA Functional Requirements models (FRBR, FRAD, FRSAD) replaced by IFLA Library Reference Model (IFLA LRM, 2017)

• Descriptive/content standards
  ▪ RDA: Resource Description and Access (integrating resource, 2010-)
  ▪ ISBD (consolidated edition, 2011; complete revision process started in 2018)

• Encoding standards
  ▪ MARC 21 (ongoing development)
    ○ Significant additions made in response to RDA, 2008-2011)

• Data models
  ▪ BIBFRAME (ongoing development)

• Introduction of LCGFT, LCMPT, LCDGT, FAST
IFLA LRM

• Framework for the analysis of non-administrative metadata relating to library resources
  ▪ High-level, enhanced entity-relationship model

• Consolidation and update of existing IFLA conceptual models
  ▪ FRAD: Functional Requirements for Authority Data (2008, amended 2013)
  ▪ FRSAD: Functional Requirements for Subject Authority Data (2010)

  (CIDOC = International Council of Museums Committee on Documentation)
  ▪ And its extensions
    ○ FRBRoo (2009; version 2.4, 2015) – object oriented version of FRBR
    ○ PRESSoo (2013) – Describes serials and continuing resources
Applying LRM

• Cataloging code implementations
  ▪ Must respect the basic structure of the entities and relationships among them
  ▪ May add expansions within the basic framework (specialized attributes, specific subclasses, refinements of relationships, etc.)
  ▪ May focus only on WEMI attributes and relationships, omitting the others
  ▪ May use either literals (strings) or URIs/IRIs (things) to represent attributes
    ○ In the latter case, these really become relationships

• RDA (beta Toolkit) is the first implementation of LRM
  ▪ With expansions
IFLA LRM – Elements

• User tasks – 5 total
  ▪ Find, Identify, Select, Obtain, Explore
  ▪ The functional scope of the model: everything supports at least one of these tasks

• Entities – 11 total
  ▪ Key objects of interest to users – define the framework of the model
  ▪ Abstract class of conceptual objects
  ▪ Serve as domains and ranges of the relationships

• Attributes – 37 total
  ▪ Data that characterize specific instances of an entity
  ▪ Recorded when considered relevant
  ▪ LRM listing not exhaustive

• Relationships – 36 total
  ▪ Connect entities to each other and provide context for them
  ▪ Relationships between WEMI are the core of the model
Scope of RDA

- RDA: Resource Description and Access is a package of data elements, guidelines, and instructions for creating library and cultural heritage metadata that are well-formed according to international models.
- Focused on “things” of bibliographic interest.
- A refinement of IFLA LRM:
  - Elements fulfill one or more LRM user tasks
    - Other information is out of scope
      - Administrative metadata
      - Rights metadata
      - Subject access beyond RDA entities
3R Project: RDA Toolkit Restructure and Redesign

• Goals
  ▪ Implement IFLA LRM entities
    o *New:* Agent, Collective agent, Nomen, Place, Timespan
    o *Revised:* Person
  ▪ Retain current RDA elements unless they conflict with LRM
    o Redundant elements will be removed after sometime after 3R
      ▪ Unstructured description of an element (e.g., *type of recording*) preferred over old “details of” approach (e.g., *details of type of recording*)
  ▪ Generalize the instructions
  ▪ Offer more flexibility in choosing how to record a piece of information
  ▪ Develop a new approach to relationship designators
3R Project: RDA Toolkit Restructure and Redesign

• Goals, continued
  ▪ Promote even greater international adoption of RDA
    ○ Original Toolkit available in 8 languages
    ○ RDA vocabularies (in whole or in part) available in 18 languages now
  ▪ Better position RDA for use in linked data applications
  ▪ Improve the Toolkit interface and address long-standing problems
  ▪ Restructure the underlying data
    ○ Provide efficient and reliable work processes and tools for RDA editors and translators
3R Project Milestones

• 2016: Some wording changes start
  ▪ “agent” replaced “person, family, or corporate body”
  ▪ Wording in relationship definitions standardized

• April 2017: Original RDA Toolkit content frozen

• June 13, 2018: Beta Toolkit publicly released
  ▪ Undergoing continuous improvement
    ○ Based in part on feedback from users
  ▪ Certain changes announced as “releases”
    ○ The most recent of these was on Oct. 9, 2018
3R Project Next Steps

• Stabilize the English text (anticipated April 2019)
  ▪ Identify which elements have “final” wording and which will undergo development in the “post 3R” period
    ○ For the former, commit to no substantive rewording that would lead to new translations
  ▪ Critical step before work can begin in earnest for translators and policy statement creators

• Add policy statements and translations to beta site

• Remove the “beta” status – the new Toolkit becomes official
  ▪ Requires unanimous votes from RSC, RDA Board, and the copyright holders
  ▪ One year countdown clock starts on original Toolkit
  ▪ Late 2019?
RDA – Entities

• Entities – 13 total
  ▪ Same scope as LRM
  ▪ *Corporate Body* and *Family* retained as subclasses of LRM’s *Collective Agent*
  ▪ *RDA entity* is highest entity in model, rather than LRM’s *Res*
RDA – Elements

• Specific aspects, characteristics, attributes, or relationships used to describe an entity

• Each element has its own page in RDA, with the following structure
  ▪ Definition and Scope
  ▪ Element Reference – includes
    o Canonical IRI
    o Domain and Range
    o Alternate labels
    o MARC 21 and Dublin Core mappings
  ▪ Prerecording
  ▪ Recording
  ▪ Related Elements
RDA – Elements

• Elements – over 1700

<table>
<thead>
<tr>
<th>Element</th>
<th>Count</th>
<th>Element</th>
<th>Count</th>
</tr>
</thead>
<tbody>
<tr>
<td>Work</td>
<td>388</td>
<td>Agent</td>
<td>175</td>
</tr>
<tr>
<td>Expression</td>
<td>291</td>
<td>Person</td>
<td>85</td>
</tr>
<tr>
<td>Manifestation</td>
<td>282</td>
<td>Collective Agent</td>
<td>34</td>
</tr>
<tr>
<td>Item</td>
<td>70</td>
<td>Corporate Body</td>
<td>84</td>
</tr>
<tr>
<td>Place</td>
<td>45</td>
<td>Family</td>
<td>46</td>
</tr>
<tr>
<td>Timespan</td>
<td>54</td>
<td>Nomen</td>
<td>169</td>
</tr>
<tr>
<td>RDA Entity</td>
<td>27</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
RDA – Attributes, Relationships

• Attributes
  ▪ Same scope as LRM

• Relationships
  ▪ Same scope as LRM
  ▪ Represented as elements – no longer “relationship designators”
  ▪ Includes inverses
  ▪ Relationship matrix – 13 x 13 categories
    ○ Browse all entities and relationships in a two-level hierarchy (domain & range) that places all of the RDA relationships in context
RDA – Guidance Chapters

- Contain general information, such as
  - Background statements
  - Guidance that applies to more than one element
  - Explanations of concepts and specific terminology
- Some chapters built on text in the original Toolkit; others are brand new
- Designed for various audiences
  - Catalogers
  - Administrators
  - Application developers
Changes with 3R Project

• Organization
  ▪ Data focused, not workflow driven
    o Instead, choose your own adventure!

• Consistent presentation of four recording methods
  ▪ Unstructured, Structured, Identifier, IRI
  ▪ Indication when one or more of them are not applicable
  ▪ Choice of which one(s) to use up to cataloging agencies

• Access points have become elements
  ▪ Three tiers
    o Access point for [manifestation, person, place, etc.]
    o Authorized access point for [expression, agent, timespan, etc.]
    o Variant access point for [work, corporate body, etc.]
Changes with 3R Project

• Concepts introduced or refined
  ▪ Aggregate and diachronic works
    ▪ Significant impact on serials
    ▪ Instructions under development, with advice from serials experts
  ▪ Manifestation statements
  ▪ Representative expression elements
  ▪ Data provenance
    ▪ Includes sources of information
  ▪ Application profiles
  ▪ Nomen and nomen string

• New terminology and phrasing
  ▪ Including how to present options and exceptions
Changes with 3R Project

• Removal of analytical, hierarchical, and comprehensive descriptions
  ▪ In favor of overall guidance about describing the work in focus

• Elimination of instruction numbers
  ▪ Plan to create “citation numbers” instead
  ▪ These will refer to specific sections in an element chapter

• No more core elements
  ▪ Application profiles and/or policy statements will address this
Recording Methods
Recording Methods (4-fold path)

- Four ways to capture data
- Included in every element page
- Flexible – supports:
  - Different implementation scenarios
  - Different approaches (internationalization)
  - Multiple choices within a given cataloging agency
  - Can use more than one, if applicable and desired
Which Method?

- **Unstructured**
  - E. B. White [transcribed from title page]
    - May also be recorded as name of person
- **Structured**
  - White, E. B. (Elwyn Brooks), 1899-1985
    - May also be recorded as authorized access point for person
- **Identifier**
  - 0000 0001 1028 499X [ISNI]
    - May also be recorded as identifier for person
- **IRI**
  - [VIAF]
    - IRI of person
Aggregates
Aggregate Manifestations

• Three types
  ▪ Collection: gathering of independent works
    o Anthology of short stories
    o Compact disc with multiple works
  ▪ Augmentation: one main work with supplemental work(s)
    o Book with introduction, index, etc.
    o Novel with illustrations
  ▪ Parallel: more than one language expression of the same work
    o Multi-lingual government publications
    o Text in the original language with a translation
    o DVD with a choice of spoken and/or subtitle languages

• An aggregate may consist of more than one type
An Aggregate Manifestation...

• Has an aggregating expression
  ▪ The selection of the expressions to be published together
  ▪ Not a whole/part relationship

• Which in turn has an aggregating work
  ▪ The plan for selecting the expressions to be published together
  ▪ Creator is responsible for the plan
    o May or may not have creative responsibility for the intellectual content of the works in the resulting manifestation
  ▪ Realized by only one aggregating expression
    o But an aggregating expression can be embodied by more than one manifestation
    o Any change in content means it’s a new work
LRM Model for Aggregates

Figure 5.7 General Model for Aggregates

- The Works
  - is realized through
  - is embodied in

- Expressions
  - is embodied in

- Aggregating Work
  - is realized through

- Aggregating Expression
  - is embodied in

Aggregate Manifestation
In Real Life...

• Sometimes the cataloging agency cares about more than one part of the aggregate manifestation (but often not)
  ▪ Contributor is employed by your institution
  ▪ Author of introduction and/or chapter is famous
  ▪ Archival or special collections considerations

• If important, describe the related work, etc.

• If not of great importance, recording methods to the rescue!
  ▪ Unstructured description:
    o Include the information when transcribing the title page, or make a note
  ▪ Structured description:
    o Provide an authorized access point (work, person, etc.)
Elements of Style

AE: Expression of the plan ...

AW: Work plan for *The Elements of Style* Fourth Edition

E1: Strunk’s text in English

E2: White’s chapter in English

The Elements of Style Fourth Edition (Longman, 2000)

W1: The Elements of Style

W2: An Approach to Style

realizes

aggregated by

embodies

embodies

realizes
Diachronic works
Diachronic Works

• Planned to be embodied over time
  - As opposed to static works, which are embodied in a single act of publication

• A new way of categorizing works
  - Different mind set from the familiar multipart monograph, serial, and integrating resource
  - Instead, three new concepts
    - Extension requirement: essential vs. inessential
    - Extension mode: integration/replacement vs. accumulation/succession
    - Extension termination: predetermined end vs. ongoing
New Element – Extension Plan

• Categorization that reflects intention to extend content of work
  ▪ Integrating determinate plan
    ○ Example: conference website
  ▪ Integrating indeterminate plan
    ○ Example: ongoing wiki
  ▪ Successive determinate plan
    ○ Example: serialized version of novel
  ▪ Successive indeterminate plan (serial work plan)
    ○ Example: newsletter
  ▪ Static plan
    ○ Example: novel in one volume
Diachronic Works

• Concept applies to more than just serials
  ▪ Includes multipart monographs that are not complete when first published
    o The complete poetical works / Lord Byron ; edited by Jerome J. McGann: 7 volumes, 1980-1993
  ▪ But not all multipart monographs are diachronic
    o The winning of the West / by Theodore Roosevelt: 4 volumes, [2004]
WEMI Relationships (FRBR)

- Work
  - is realized through Expression
  - is embodied in Manifestation
    - is exemplified by Item
WEM Lock – Diachronic Works

• Diachronic plan creates one-to-one links between Work, Expression, and Manifestation

1 and only 1
Work

is realized through realizes
Expression

1 and only 1
is embodied in embodies
Manifestation
Challenges with Serials

• A change in plan = New work
  ▪ Different carrier types
  ▪ Regional editions
  ▪ Translations
  ▪ Cessation of the print version in favor of the online version

• But these different works are still closely related
  ▪ RDA introducing new concept of “work group”
    ▪ A group of two or more works that have a common appellation assigned from a vocabulary encoding scheme [such as an authority file]
    ▪ Will allow collocation of serial works in the same “family”
  ▪ ISSN-L already functions in this way
Challenges with Serials

- In many cases, serials are aggregates of aggregates
  - An issue contains different articles
  - The serial consists of many issues

- RDA uses three high-level relationships from LRM to model the complexity of serials – implemented as elements in RDA
  - Work-Work relationships
    - LRM-R19 precedes [logical, not chronological]
    - LRM-R22 was transformed into [by policy, etc.]
  - Expression-Expression relationship
    - LRM-R25 was aggregated by
Options
Why So Many Options?

• Internationalization
  ▪ No single way to describe and access a resource

• Four different recording methods
  ▪ Two or more are valid for many elements
    ○ Choice depends in part on local cataloging environment

• Nothing is mandatory
  ▪ Except “nomen string” – the combination of signs that forms an appellation associated with an entity
Evolution of Options

Original Toolkit

- **Alternative**: a different approach to what is called for in the preceding instruction
- **Optional addition**: data that supplements what is called for in the preceding instruction
- **Optional omission**: ability to omit specific data called for in the preceding instruction
- **Exception**: takes precedence over the preceding instruction

- Policy statements often provide guidance about which of these to apply

Beta Toolkit

- **Option**
  - Optional guidelines and instructions (covers *optional additions/omissions*)
  - Different choice for the same set of conditions (covers *alternatives*)
  - Different set of conditions (covers *exceptions*)

- **Conditions**
  - Many instructions apply only when one or more conditions are met

- Application profiles or policy statements provide guidance about which to apply
2.10.4.4 Statement of Function

Record words or phrases indicating the function performed by an agent as they appear on the source of information.

**EXAMPLE**

Manufactured and marketed by PolyGram Video, a division of PolyGram Records, Inc.

**Optional Addition [LC-PCC PS]**

If the function of an agent recorded as a manufacturer’s name is not explicit or clear, add a term indicating the function. Indicate that the information was taken from a source outside the manifestation itself (see 2.2.4 RDA).
name of agent of manifestation

Statement of function

OPTION
Include words or phrases indicating the function performed by an agent as they appear on the manifestation.

CONDITION
A function is performed by an agent that is not explicit or clear.

OPTION
Add a term indicating the function.
Indicate that the information was not taken from the manifestation itself.
Application Profiles
Purpose

• Manages the choices – what to record
  ▪ Mandatory elements
  ▪ Recommended elements
  ▪ Repeatability
  ▪ Vocabularies
  ▪ Recording method(s)
  ▪ Which options to apply

• Differs in part from policy statements – if they offer...
  ▪ Guidance about how to record
  ▪ Supplemental information
  ▪ Examples
Application Profiles in the New Toolkit

• Multiple ways to implement
  ▪ Bookmarks and notes (in your institutional subscription)
  ▪ Policy statements (linked to instructions)
  ▪ Workflows and other user documentation (can be shared locally or globally)
  ▪ External documents
For More Information
Helpful Resources

• RDA Beta Toolkit: https://beta.rdatoolkit.org/RDA.Web
• RDA Toolkit website: https://www.rdatoolkit.org/
  ▪ 3R Project website: http://www.rdatoolkit.org/3Rproject
• RDA Steering Committee: http://www.rda-rsc.org/
  ▪ Presentations: http://www.rda-rsc.org/rscpresentations
  ▪ Documents: http://www.rda-rsc.org/documents
  ▪ 3R Project FAQ: http://rda-rsc.org/node/551
• IFLA Library Reference Model
Questions?