RDA, LINKED DATA, & THE 3R PROJECT

Kathy Glennan
Head, Original & Special Collections Cataloging, University of Maryland
ALA Representative to NARDAC
Chair-Elect, RDA Steering Committee
More Than Just Cataloging Rules

“RDA is a package of data elements, guidelines, and instructions for creating library and cultural heritage resource metadata that are well-formed according to international models for user-focused linked data applications.”

-- 2015 Statement from Committee of Principals for RDA
Genesis of RDA Text

- Based largely on AACR2
  - AACR3 draft, Dec. 2004
- Also tied to the FRBR and FRAD models
- U.S. National Libraries Test Plan for RDA (completed 2011)
- Day 1 for PCC RDA Authority records: March 31, 2013
Genesis of RDA as Linked Data

- Work started in 2007
  - DCMI/RDA Task Group formed
    - Participants from the JSC, DCMI, and W3C Semantic Web Deployment Working Group
  - First RDA vocabularies published in the Open Metadata Registry in 2011
  - RDA Registry launched in 2014
Different “flavors” of RDA

- RDA is becoming more of a data dictionary
- Different pieces for different types of users
  - RDA Reference ➔ Developers
  - RDA Vocabularies ➔ Developers
  - RDA Registry ➔ Application developers
  - RDA Toolkit ➔ Catalogers
  - RIMMF ➔ Trainers
RDA Reference

- All RDA Elements, their definitions and any related scope notes, and all value vocabulary terms and definitions
- Stored in the Open Metadata Registry (OMR) in RDF linked data format
- Primary source of RDA Toolkit content
- Includes translations
OMR Example - Researcher

Element Sets: RDA Work properties
Elements: has researcher

<table>
<thead>
<tr>
<th>Metadata +</th>
</tr>
</thead>
<tbody>
<tr>
<td>Detail</td>
</tr>
<tr>
<td>Canonical URI:</td>
</tr>
<tr>
<td>Label:</td>
</tr>
<tr>
<td>Name:</td>
</tr>
<tr>
<td>Lexical Alias:</td>
</tr>
<tr>
<td>Description:</td>
</tr>
<tr>
<td>Comment:</td>
</tr>
<tr>
<td>Type:</td>
</tr>
<tr>
<td>Parent:</td>
</tr>
<tr>
<td>Domain:</td>
</tr>
</tbody>
</table>
### OMR Example – Translations

<table>
<thead>
<tr>
<th>Profile property</th>
<th>Language</th>
<th>Object</th>
</tr>
</thead>
<tbody>
<tr>
<td>description</td>
<td>Catalan</td>
<td>Relaciona una obra amb un agent que fa recerca per donar suport a la creació d'una obra.</td>
</tr>
<tr>
<td>description</td>
<td>Danish</td>
<td>Relaterer et værk til en agent, der udfører forskning til støtte for skabelsen af et værk.</td>
</tr>
<tr>
<td>description</td>
<td>German</td>
<td>Setzt ein Werk mit einem Akteur, der Forschung betreibt, um die Schaffung eines Werks zu unterstützen, in Beziehung.</td>
</tr>
<tr>
<td>description</td>
<td>Greek</td>
<td>Συνδέει ένα έργο με έναν φορέα, ο οποίος διεξάγει έρευνα υποστηρικτικά της δημιουργίας ενός έργου.</td>
</tr>
<tr>
<td>description</td>
<td>English</td>
<td>Relates a work to an agent who does research in support of the creation of a work.</td>
</tr>
<tr>
<td>description</td>
<td>Finnish</td>
<td>Liittää teoksen toimijaan, joka tukee teoksen syntymistä tutkimuksellaan.</td>
</tr>
<tr>
<td>description</td>
<td>French</td>
<td>Met une œuvre en relation avec un agent qui effectue des recherches pour aider à la création d'une œuvre.</td>
</tr>
<tr>
<td>description</td>
<td>Vietnamese</td>
<td>Liên kết tác phẩm với tác nhân tiến hành nghiên cứu để hỗ trợ sáng tạo tác phẩm.</td>
</tr>
<tr>
<td>domain</td>
<td></td>
<td><a href="http://rdaregistry.info/Elements/c/C10001">http://rdaregistry.info/Elements/c/C10001</a></td>
</tr>
<tr>
<td>has element type</td>
<td></td>
<td>2</td>
</tr>
<tr>
<td>hasSubproperty</td>
<td></td>
<td><a href="http://rdaregistry.info/Elements/w/datatype/P10293">http://rdaregistry.info/Elements/w/datatype/P10293</a></td>
</tr>
</tbody>
</table>
RDA Vocabularies

- RDA Reference exported to a GitHub repository
- Published as a cumulative snapshot in discrete, releases with version control
- Allows applications to update and synchronize
RDA Registry

- Provides links to download the individual element sets and value vocabularies in the current release of RDA Vocabularies
- Additional documentation and tools for RDA developers
  - Technical Guide
  - Examples of RDA linked data
  - Machine-actionable maps for transforming RDA data for use by non-RDA applications
- With each new release, sends updated data to RDA Vocabulary Server
RDA Vocabulary Server

- Provides download management and linked data de-referencing services for machine applications

- Services available for
  - Language specific extractions
  - Individual Elements and Concept/Values extractions which provide RDF representations of specific IRIs
RDA Toolkit

- An integrated, browser-based, online product that allows users to interact with a collection of cataloging-related documents and resources including *RDA: Resource Description and Access*

- The only part of “RDA” that is not freely available
RIMMF3

- A visualization tool
- A cataloging training tool
- A prototype for what an RDA-based cataloging interface might look like
- Uses the RDA Vocabularies data to keep up-to-date
  - Now includes support for some translations
RDA Reference Data Workflow
RDA Implementation Scenarios

- Flat file (card catalogs)
- Linked authorized access points (e.g., OCLC Connexion)
- Relational or object database (fully linked at local level)
- Linked data (fully linked at global level)
What Drives RDA Development?

- International standards & translation needs
- Description needs for cultural heritage organizations
- 3R Project
- Linked data and the Semantic Web
- User communities, including ALA, PCC, & NARDAC
RDA – Connected to International Standards

- IFLA Library Reference Model (LRM), 2017
  - A consolidation and update of the 3 earlier functional requirements models (FRBR, FRAD, FRSAD)

- CIDOC Conceptual Reference Model (CRM), 2006
  - Developed for use by museums and cultural heritage institutions
  - FRBRoo is an object oriented extension of the CRM
3R Project Recap

- New version of the RDA Toolkit coming: June 2018
  - Not RDA 2.0 – instead, a new Expression
- Goals include:
  - Implement new LRM entities (Agent, Timespan, etc.)
  - Improve the interface and address long-standing problems
  - Generalize the instructions where possible
  - Offer more flexibility in choosing how to record a piece of information
New Concepts Coming to the Toolkit

- Recording methods
- Manifestation statement
- Rethinking serial works
  - Diachronic works
- Redefinition of person
- Relationships preferred over attributes
Recording Methods

- Entity chapters will present up to four different recording methods for capturing data
- Can use more than one
  - Unstructured description [on Manifestation]: Denver, Colorado
  - Structured description [LC/NACO NAF]: Denver (Colo.)
  - Identifier [Getty Thesaurus of Geographic Names]: 7013545
  - IRI [Wikidata]: http://www.wikidata.org/entity/Q16554
Manifestation Statement

- How a resource represents itself
- Transcribed from the *Manifestation*
  - A more literal transcription than we’ve been doing
    - Capitalization, punctuation, etc. retained
  - Recording method = Unstructured description
- Fulfills the “identify” user task
- Also supports machine transcription derived from digitized and born-digital manifestations
Reconceptualizing Serials via LRM

- Serial work = sequentially issued sequence of aggregate manifestations
  - The plan for the aggregating work, not the results
- Combination of both aggregate and whole/part works
  - Each issue is an aggregate of articles
  - Each issue has a whole/part relationship to the serial work
- Dynamic
  - Description not limited to the past; includes assumptions about the future
Reconceptualizing Serials via LRM

- Regard *any* serial as a distinct instance of the *work* entity
  - Each serial work has only one expression and one manifestation, including:
    - Regional “editions”
    - Translations
    - Different formats (e.g., print vs. electronic)

- ISSN: Identifies an individual serial *work* (not a *manifestation*)

- ISSN-L: Identifies a relationship between two serial *works*
Diachronic Works

- Planned to be embodied over time
  - Includes works with a planned end, and those without

- *Work* content changes over time
  - Revision / replacement – Single expression
    - May be done at regular intervals
  - Additions / accumulations – Successive expressions
    - Previous content retained

- Once content is completed, no longer a diachronic work
Aggregating Works

- A plan for bringing together distinct *expressions* into one or more *manifestations*
  - Creator: the *agent* who came up with the plan
  - Relationships: hold true only for a single aggregating *work*
- For diachronic works: cataloging reflects the plan
- For static works: catalog the results
Redefinition of Person

- Restricted to real persons who live or are assumed to have lived
  - Can have separate bibliographic identities
- Out of scope: fictional, literary, and legendary personas
  - These are valid concepts, but are not instances of the Person entity
- In a name-based authority file, this redefinition does not require a change in practice
Implications for Name Authorities

- Do NARs control names (Nomen) or persons?
  - Nomen aspects: Separate NARs for Charles Dodgson and Lewis Carroll
  - Person aspects: Dodgson’s/Carroll’s birth place, birth date, associated institution, etc.

- Potential for confusion when different attribute values are assigned to the same person:
  - NAR for J. K. Rowling: 375 Females ≠ 2 lcdn
gt
  - NAR for Robert Galbraith: 375 Males ≠ 2 lcdn
gt
Recording Methods: Identifier vs. IRI

- **Identifier**
  - Assigned by an agency/community, following a pattern (local/regional)
  - No preference for Nomen vs. Person approach – supports either one

- **IRI/URI**
  - Globally unique
  - In an LRM/RDA context, expect a single IRI for the Person
    - Rowling/Galbraith/Scamander/Whisp collectively
  - Problematic mapping to URI’s for real world object agents in id.loc.gov
    - Rowling, J. K.: [http://id.loc.gov/rwo/agents/n97108433](http://id.loc.gov/rwo/agents/n97108433)
    - Galbraith, Robert: [http://id.loc.gov/rwo/agents/n2013043083](http://id.loc.gov/rwo/agents/n2013043083)
Relationships vs. Attributes

- LRM recasts attributes as relationships where possible
- In RDA,
  - *Timespan* will include relationships from an RDA entity to date of birth, date of publication, etc.
  - *Place* will include relationships from an RDA entity to address, location associated with a conference, etc.
  - *Nomen* will include the appellation relationship between
    - A work and its title
    - A person and his/her rank, honour, or office
    - A corporate body and its name
Modeling Events

- While not an explicit entity, an event can be described through relationships to the following entities, as appropriate:
  - Place
  - Timespan
  - Agent
  - Nomen
New Governance Structure - NARDAC

- North American RDA Committee – new!
  - Two representatives each from ALA, CCC, and LC
  - Should start its work before the end of February
  - One NARDAC rep to RSC
    - Other North American experts may be co-opted members of RSC
- Change in proposal process
  - ALA → NARDAC → RSC
Potential Next Steps for PCC

- Update/revise LC PCC PSs
  - As instructions become more general, more guidance will be needed at the policy level

- Training
  - On LRM?
  - On new RDA Toolkit?
    - ALA Publishing training will cover basics of using the Toolkit, but not application of the instructions
A Toolkit for Today and Tomorrow

“The 3R Project builds on RDA’s strengths to benefit institutions where they are today and where they would like to be tomorrow.”
More Information

- RDA Toolkit/3R Project website
  - http://www.rdataoolkit.org/3Rproject

- RDA Steering Committee – Presentations
  - http://www.rda-rsc.org/rscpresentations
    Includes recent presentations by various RSC members

- IFLA Library Reference Model

- Ask me:
  - kglennan@umd.edu
Questions?