NISO/DCMI Webinar

The RDA Vocabularies: Implementation, Extension, and Mapping


May 2007
RDA Data Model meeting

2007 recommendations

• The meeting participants agreed that RDA and DCMI should work together to build on the existing work of both communities.

• Specific activities be undertaken:
  – development of an RDA Element Vocabulary
  – development of an RDA DC Application Profile based on FRBR and FRAD
  – disclosure of RDA Value Vocabularies using RDF/RDFS/SKOS

http://www.bl.uk/bibliographic/meeting.html

2007 expected benefits

• The benefits of this activity will be that:
  – the library community gets a metadata standard that is compatible with the Web Architecture and that is fully interoperable with other Semantic Web initiatives
  – the DCMI community gets a libraries application profile firmly based on the DCAM and FRBR (which will be a high profile exemplar for others to follow)
  – the Semantic Web community get a significant pool of well thought-out metadata terms to re-use
  – there is wider uptake of RDA
October 2011

• “A Bibliographic Framework for the Digital Age” (Library of Congress)
  – “The new bibliographic framework project will be focused on the Web environment, Linked Data principles and mechanisms, and the Resource Description Framework (RDF) as a basic data model.”
  – “Many of the libraries taking part in the test indicated that they had little confidence RDA changes would yield significant benefits without a change to the underlying MARC carrier.”

The RDA Vocabularies

What They Are, How They Work

What We’re Doing Today

• Sorting out some important definitions
• A structural tour of the RDA vocabularies, including:
  – Rationale and decision making
  – General and FRBR-constrained elements
  – Potential for extension
• Putting it all together
• Language versions of RDA Vocabularies
• Important related work

Two Models of ‘the World’

• XML assumes a closed world (domain), usually defined by a schema:
  – "We know all of the data describing this resource. The single description must be a valid document according to our schema. The data must be valid."
  – XML’s document model provides a neat equivalence to a metadata ‘record’
• RDF assumes an open world:
  – "There’s an infinite amount of unknown data describing this resource yet to be discovered. It will come from an infinite number of providers. There will be an infinite number of descriptions. Those descriptions must be consistent."
  – RDF’s statement-oriented data model has no notion of record (rather, statements can be aggregated for a fuller description of a resource)
In Both, We Talk About Statements

- In XML these are ‘attribute/value pairs’
  - Ex.: Author = Twain, Mark
  - Where, in the context of a bounded record structure we know what we’re describing (there’s an identifier of some kind in the record)
- In the Resource Description Framework (RDF, the language of the Semantic Web), these are independent statements:
  - Ex.: The author of Huckleberry Finn is Mark Twain, or
  - Subject = Huckleberry Finn
  - Predicate = hasAuthor
  - Object = Mark Twain

The General Strategy

- We used the Semantic Web as our “mental model”
  - Wanted to create a “bridge” between XML and RDF to support innovation in the library community as a whole, not just those at the cutting edge or the trailing edge
  - We registered the FRBR entities as classes in a FRBR in RDA vocabulary, to enable specific relationships between RDA properties and FRBR
- IFLA has followed suit using the Open Metadata Registry to add the ‘official’ FRBR entities, FRAD, FRSAR and ISBD
  - This provides useful opportunities to relate all the vocabularies to each other, improving interoperability

Structure: Rationale & Decisions

- Property and value vocabularies are registered in the Open Metadata Registry (formerly the NSDL Registry): http://metadataregistry.org/rdabrowse.htm
- Used RDF Schema (RDFS), Simple Knowledge Organisation System (SKOS) and Web Ontology Language (OWL)
- Decisions were oriented to favor approaches that can be generalized to make other vocabulary based standards web-friendly, available for use in applications, and easily updated by communities

The Structure, Simplified

- FRBR in RDA Vocabulary declared as classes
- RDA Properties declared as a generalized vocabulary, with no explicit relationship to FRBR entities
- Subproperties for the generalized elements may be explicitly related to FRBR entities (using domain)
  - Label/Name includes (Work) or other class to provide unique name (unless the entity name already appears in the name of the property)
  - Other generalized subproperties usable by others not tied to FRBR
- All this is done according to available best practices for declaring vocabularies on the Web
Why Generalized Properties?

- We think of the ‘generalized’ RDA properties as the *real* RDA vocabulary
  - The ‘bounded’ properties should be seen as the first pass at an Application Profile
  - Extensions can be built more usefully from the generalized properties
  - Mapping will be cleaner using the generalized properties (since most properties mapped to or mapped from will not be based on FRBR)
  - Generalized properties are much more acceptable to non-library implementers (not often using FRBR)
More Complex Relationships

- There are multiple techniques used in RDA to make the connection between FRBR entities and RDA properties.
- We tried to reconcile those in the RDA vocabularies.
- Some properties are related to more than one FRBR entity.
  - Relationships in Appendix J actually include the name of the FRBR entity in the name and have separate definitions (we re-used this strategy for the FRBR-bounded properties).
  - Other properties and sub-properties appear multiple times in the text and ERDs, with the same definitions and no indication that they might be repeated elsewhere (we consolidated these).

The Not-So-Simple Case: One Property—more than One FRBR Entity
Roles, Entities or Relationships?

- In 2005, the DC Usage Board worked with LC to build a formal representation of the MARC Relators so that these terms could be used with DC
  - This work provided a template for the registration of the role terms in RDA (in Appendix I) and, by extension, the other RDA relationships
    - Roles, relationships and attributes are registered as properties
    - Roles are not classes (Author, Illustrator, etc.), but relationships between FRBR Groups 1-2

Aggregated Statements

- RDA sets up Publication, Distribution, Manufacture and Production statements very much the way they have been done since catalog card days:
  - Assumed aggregation of Place, Name and Date are obvious leftovers from catalog cards, and are not necessary to enable indexing or display of those elements together if libraries want to do that
  - We viewed those aggregations as ‘Syntax Encoding Schemes’ (as defined in the DCAM) and built in ways to accommodate them within the bounded properties
  - Those using the generalized properties (outside libraries, usually) need not be constrained by these traditional aggregations of properties
Two full hierarchies, related at every level

General properties at top level, FRBR properties more specific in usage

What Does This Structure Buy Us?

- Release from the tyranny of records
  - Potential for use with a variety of encodings
  - Ability to maintain statements at a more granular level
- Opportunity to re-think how libraries build, maintain and share data
- Potential for sharing data beyond the library silo (in both directions!)
- A challenge to old notions (ours and others) of what library data can do and should be doing
  - As library users migrate to the web and away from library catalogs, libraries need to follow them there
Time for Questions?

Putting it All Together

- How do we take advantage of the RDA Vocabulary infrastructure?
  - Relationships between properties are built-in, allowing ‘bottom up’ aggregation of data
  - Tools to enable easier data creation (with machine assistance) are critical
- Our thinking about how we build and share data needs to reflect the modern environment of data use and the Web
  - We can avoid the baggage of past practice and retain the best of our legacy data approaches

Statements on the Floor?

Is This Really Chaos?
Or Just an Aggregation
In the Making?
Extension

• The inclusion of generalized properties provides a path for extension of RDA into specialized library communities and non-library communities
  – They may have a different notion of how FRBR ‘aggregates’
    • For example, a colorized version of a film may be viewed as a separate work
  – They may not wish to use FRBR at all
  – They may have additional properties to include, that have a relationship to the RDA properties
    • This allows other communities to use this data without problems
What We’ve Learned

• We wrote about the decisions we made for RDA in DLib Magazine:
  http://dlib.org/dlib/january10/hillmann/01hillmann.html
• We will continue to disclose what we’ve learned and work on building best practices documentation in this environment
• We need more research and innovation to help us all move forward!

Multi-Lingual RDA

• Registered elements and vocabularies can include multiple languages
• The Open Metadata Registry has been working with the Deutsche Nationalbibliothek to demonstrate this with German translations in the OMR for specific vocabularies
  – Other language versions being discussed with national libraries and language groups by ALA Publishing in conjunction with translations of RDA guidance text
  – OMR enhancements planned to make building of language versions easier
First, Read the Definitions!

dc:creator: “An entity primarily responsible for making the resource.”

rdvocab:creator: “A person, family, or corporate body responsible for the creation of a work.”

Are they the same?

Crosswalking Relationships ...

... use only one predicate

Mapping Relationships ...

... can use any available predicate
Links and Contact Info

- DCMI/RDA Task Group Wiki: http://dublincore.org/dcmirdatastaskgroup/ (moving soon)
- RDA Vocabularies: http://metadataregistry.org/rdabrowse.htm

- Diane: metadata.maven@gmail.com

Flickr images thanks to: Trekk Såtho (TT), Craftology