The Challenges Facing Cataloging and Catalogers

FACING FORWARD
How We’ll Proceed

- What’s this Semantic Web thingy all about, and why do we care?
- Is RDA really going to happen?
  - How is it different from AACR2?
- How will RDA affect catalogers and cataloging?
- How can we prepare for all this?
Whether you want it or not …

A short tutorial on the Semantic Web
What’s This Semantic Web?

- **RDF: Resource Description Framework**
  - Statements about Web resources in the form of subject-predicate-object expressions, called triples
  - E.g. “This presentation” –“has creator” –“Diane Hillmann”

- **RDF Schema**
  - Vocabulary description language of RDF

- **SKOS: Simple Knowledge Organisation System**
  - Expresses the basic structure and content of concept schemes such as thesauri and other types of controlled vocabularies
  - An RDF application

- **OWL (Web Ontology Language)**
  - Explicitly represents the meaning of terms in vocabularies and the relationships between them
Semantic Web Building Blocks

- Each component of an RDF statement (triple) is a “resource”
- RDF is about making machine-processable statements, requiring
  - A machine-processable language for representing RDF statements
    - Extensible Markup Language (XML)
  - A system of machine-processable identifiers for resources (subjects, predicates, objects)
    - Uniform Resource Identifier (URI)
    - For full machine-processing potential, an RDF statement is a set of three URIs
Things Requiring Identification

- Object “This presentation”
  - e.g. its electronic location (URL):
- Predicate “has creator”
  - e.g. http://purl.org/dc/terms/creator
- Object “Diane Hillmann”
  - e.g. URI of entry in Library of Congress Name Authority File (real soon now?)
  - NAF: nr2001015786
- Declaring vocabularies/values in SKOS and OWL provides URIs
  - Without such identifiers, the Web will never become Semantic
Why do we care about this geeky stuff?
The New Alphabet Soup

- RDA: Resource Description and Access
- FRBR: Functional Requirements for Bibliographic Records
  - FRBRoo: Object Oriented FRBR (harmonized with CIDOC)
- FRAD: Functional Requirements for Authority Data
- FRASAR: Functional Requirements for Subject Authority Records
### Standards Upgrade!

<table>
<thead>
<tr>
<th>Type of Standard</th>
<th>Old Standard</th>
<th>New Standard(s)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Bibliographic Model</td>
<td>None</td>
<td>FRBR, FRBRoo</td>
</tr>
<tr>
<td>Metadata Content</td>
<td>RDA</td>
<td>RDA</td>
</tr>
<tr>
<td>Metadata Structure</td>
<td>MARC21 Bibliographic</td>
<td>RDAVocab</td>
</tr>
<tr>
<td>Name Authority</td>
<td>MARC21 Authority</td>
<td>FRAD</td>
</tr>
<tr>
<td>Subject Authority</td>
<td>MARC21 Authority</td>
<td>FRASAR, SKOS</td>
</tr>
<tr>
<td>Encoding</td>
<td>MARC21</td>
<td>XML, XML/RDF</td>
</tr>
</tbody>
</table>
The RDA You’ve Heard About ...

- 4th quarter calendar 2008 – Full draft of RDA is sent out for constituency review
- Late January 2009 – end of constituency review
- 2nd quarter calendar 2009 – RDA content is finalized
- 3rd quarter calendar 2009 – RDA is released
- 3rd and 4th quarters calendar 2009, possibly into 1st quarter calendar 2010 – Testing by national libraries
- 1st and 2nd quarters calendar 2010 – Analysis and evaluation of testing by national libraries
- 3rd-4th quarters calendar 2010 – RDA implementation
What You Might Not Have Heard ...

- JSC has gradually backed away from their original stance that RDA could be expressed easily in MARC21
  - Full integration of FRBR entities has made that problematic
- RDA has been developed explicitly to take advantage of the Semantic Web
- Well supported rumors indicate that LC is considering discontinuing update of MARC21 sometime in 2010
JSC Scenarios

- Scenario 1: separate records for all FRBR entities with linked identifiers
- Scenario 2: composite bibliographic records (with authority records representing each entity)
- Scenario 3: one flat record, with all Group 1 entities on a single record
  - This is the only scenario that MARC can handle
The Rest of the Story

- RDA elements, roles and vocabularies have been provisionally registered
  - The vocabularies and the text will be tied together in the RDA online tool
- Some efforts have begun to consider how MARC21 data can be parsed into FRBR entities and RDA
- Discussions about long term maintenance of both RDA and the vocabularies have yet to occur
- The push is already on for a multi-language RDA Vocabulary
Under the RDA Hood

- A FRBR-based approach to structuring bibliographic data
- More explicitly machine-friendly linkages (preferably with URIs)
- More emphasis on relationships and roles
- Less reliance on cataloger-created notes and text strings (particularly for identification)
RDA & FRBR: Registered!

- RDA Elements:
  - http://metadataregistry.org/schema/show/id/1.html

- RDA Roles:

- RDA Vocabulary: Base Material
  - http://metadataregistry.org/vocabulary/show/id/35.html

- FRBR Relationships (Sandbox version)
  - http://sandbox.metadataregistry.org/vocabulary/show/id/90.html
Who’s Doing This?

- DCMI/RDA Task Group
  - See: [http://dublincore.org/dcmirdataskgroup/](http://dublincore.org/dcmirdataskgroup/)
  - Set up during the London meeting between JSC and DCMI
  - Gordon Dunsire and Diane Hillmann, co-chairs
  - Karen Coyle, consultant

- IFLA Classification and Indexing Section
  - Gordon Dunsire, Centre for Digital Library Research, University of Strathclyde
From the Cataloger Scenarios

Walking through a concrete example ...
A Cataloger Scenario

Jane Cataloger is assigned to work on a gift collection. Her first selection is a Latvian translation of Kurt Vonnegut’s "Bluebeard: a novel." She searches the library database for the original work, and finds:

* **Author:** Kurt Vonnegut
* **Work Title:** Bluebeard: a novel
* **Form of Work:** Novel
* **Original Language:** English
with links to the following expression information:

*Language of Expression: English
*Content Type: Text

and one manifestation:

*Edition: 1st trade edition
*Place of Production: New York
*Publisher’s Name: Delacorte Press
*Date of Production: 1987
*Number of Units: 300 pages
*Resource Identifier: [ISBN]0385295901
Jane begins her description by linking to the existing Work entity. She then creates an expression description:

*Language of Expression: Latvian
*Translator: Arvida Grigulis

She creates an authority record for the translator since none yet existed. She continues by creating a fuller description for the new manifestation, linking to the authority record for the Latvian publisher (what luck, it already existed!).

*Title: [in Latvian]
*Place of Production: Riga
*Publisher’s Name: Liesma
*Date of Production: 1997
A Cataloger Scenario: Updating

Jane Cataloger is assigned to work on a gift collection. Her first selection is a Latvian translation of Kurt Vonnegut's "Bluebeard: a novel." She searches the library database for the original work, and finds:

*Author: Kurt Vonnegut
*Work Title: Bluebeard: a novel
*Form of Work: Novel
*Original Language: English
A Cataloger Scenario: Updated

Jane Cataloger is assigned to work on a gift collection. Her first selection is a Latvian translation of Kurt Vonnegut's "Bluebeard: a novel." She searches the library database for the original work, and finds:

*Author: http://lcnaf.info/79062641
*Work Title: Bluebeard: a novel
*Form of Work: http://RDVocab.info/genre/1008
*Original Language: http://marclang.info/eng
with links to the following expression information:

*Language of Expression: English
*Content Type: Text

and one manifestation:

*Edition: 1st trade edition
*Place of Production: New York
*Publisher’s Name: Delacorte Press
*Date of Production: 1987
*Number of Units: 300 pages
*Resource Identifier: [ISBN]0385295901
with links to the following expression information:

*Language of Expression: http://marclang.info/eng

*Content Type: http://purl.org/dc/dcmitype/Text

and one manifestation:

*Edition: 1st trade edition

*Place of Production: http://www.getty.edu/tgn/7007567

*Publisher’s Name: http://onixpub.info/2039987

*Date of Production: 1987

*Number of Units: 300 pages

Jane begins her description by linking to the existing Work entity. She then creates an expression description:

*Language of Expression: Latvian
*Translator: Arvida Grigulis

She creates an authority record for the translator since none yet existed. She continues by creating a fuller description for the new manifestation, linking to the authority record for the Latvian publisher (what luck, it already existed!).

*Title: [in Latvian]
*Place of Production: Riga
*Publisher’s Name: Liesma
*Date of Production: 1997
Jane begins her description by linking to the existing Work entity. She then creates an expression description:

*Language of Expression: [link]
*Translator: [link]

She creates an authority record for the translator since none yet existed. She continues by creating a fuller description for the new manifestation, linking to the authority record for the Latvian publisher (what luck, it already existed!).

*Title: [in Latvian]
*Place of Production: [link]
*Publisher’s Name: [link]
*Date of Production: 1997
A Dublin Core View of the World

A Dublin Core View of the World

DCMI Abstract Model: http://dublincore.org/documents/abstract-model/
Anatomy of a Statement

Place of Production: New York
Anatomy of a Statement

Place of Production: http://www.getty.edu/tgn/7007567
A Related Description

Research

Research Home ▶ Conducting Research ▶ Thesaurus of Geographic Names ▶ Full Record Display

Getty Thesaurus of Geographic Names® Online
Full Record Display

Click the icon to view the hierarchy.

ID: 7007567

New York (inhabited place)

Coordinates:
Lat: 40 42 00 N degrees minutes  Lat: 40.7000 decimal degrees
Long: 074 00 00 W degrees minutes  Long: -74.0000 decimal degrees

Note: Located at mouth of Hudson River; visited by Europeans 16th-17th century. Settled by the Dutch; seized by the British in 1664. Was a cosmopolitan center and the nation's largest city by the 18th century. Today is a huge metropolis composed of five boroughs corresponding to five counties.

Names:
- New York (preferred, C,V,N) 
- New York City (display, C,V,N)
- City of New York (C,V,N)
- Greater New York (C,V,N) 
- Big Apple (C,V,N) 
- Gotham (C,V,N)
- New Orange (H,V,N)
- New Amsterdam (H,V,N)

Hierarchical Position:
- World (facet)
Description Sets a Key Concept!
Description Set=
“A set of one or more descriptions, each of which describes a single resource.”*

*DCAM Definition
A Description Set “Package”

Work

- Author: [http://lcnaf.info/79062641](http://lcnaf.info/79062641)
- Work Title: Bluebeard. a novel
- Form of Work: [http://RDVocab.info/genre/1008](http://RDVocab.info/genre/1008)
- Original Language: [http://marclang.info/eng](http://marclang.info/eng)

Expression

- Language of Expression: [http://marclang.info/eng](http://marclang.info/eng)
- Content Type: [http://purl.org/dc/dcmitype/Text](http://purl.org/dc/dcmitype/Text)

Manifestation

- Place of Production: New York
- Publisher: [http://onixpub.info/2039987](http://onixpub.info/2039987)
- Date of Production: 1987
- Number of Units: 300 pages
- Resource Identifier: [ISBN]

Identifier: [http://lcnaf.info/79062641](http://lcnaf.info/79062641)
- Name: Vonnegut, Kurt
- BirthDate: Nov. 11, 1922
- DeathDate: Apr. 11, 2007

Identifier: [http://RDVocab.info/genre/1008](http://RDVocab.info/genre/1008)
- PrefLabel: Novel

Identifier: [http://marclang.info/eng](http://marclang.info/eng)
- Name: eng
- PrefLabel: English
- AltLabel: en

Identifier: [http://purl.org/dc/dcmitype/Text](http://purl.org/dc/dcmitype/Text)
- Definition: A resource consisting primarily of words for reading.

Identifier: [http://onixpub.info/2039987](http://onixpub.info/2039987)
- Name: Delacourte Press
- Related: Random House
- Descr.: Delacourte Press is a subsidiary of Random House.
Go ahead, take a breath ...
How Soon Will All This Happen?

- The bad news: This isn’t like 1981, when there was a “start date” and we knew exactly when to change gears
- More bad news: This transition is likely to be a pretty messy one, and last a long time
- The good news: library vendors are starting to wake up and smell the coffee!
What Are the Challenges?

- Coordination with JSC (or it’s successor, given the need to be more inclusive) on long-term maintenance planning
  - Need for lightweight process, where change is not a multi-year marathon

- Continuing development towards a more Semantic web-friendly standard (less transcription, for instance)

- Tool development (including ILS vendors)
Yet More Challenges

- Application profiles that express more than one notion of “Work”
  - Requires strategy for including FRBR entities and relationships explicitly
- Moving the MARC legacy data into RDA
  - OCLC’s silence is worrisome
- Multi-lingual and specialized extensions
  - Non-Anglo-American communities eager to participate
Multi-lingual RDA, 1

- The DCMI Registry approach:
  - Translations of labels, definitions and comments
  - URIs stay the same, as do relationships
  - Responsibility for updating translations rests with translation “owner”; no updating services informing about changes that might need attention

- Disadvantages
  - Translations tend to become outdated over time
  - Communication with translation “owners” is managed loosely by a committee
Multi-lingual RDA, 2:

- Managed vocabulary approach
  - The NSDL Registry allows for language specific versions of all properties except for URI (There are scalability questions, but this is a starting place)
  - Mapping between language versions may be a more scalable approach over time
  - SKOS has put off discussion of mapping until its next version
Keeping up ...
Self-directed Learning

- **Web tutorials:**
  - [http://www.w3schools.com/](http://www.w3schools.com/)

- **Blogs**
  - Get a Bloglines account (free)
  - Start with a few, and expand:
    - Lorcan Dempsey
    - Karen Coyle
    - The FRBR Blog
      - A good synopsis of the FRBR Review Group meetings at IFLA:
        - [http://www.frbr.org/2008/08/11/frbr-review-group-meeting-1](http://www.frbr.org/2008/08/11/frbr-review-group-meeting-1)
Mailing lists

- Migrate from Autocat to some of the lists discussing newer ideas
  - web4lib@webjunction.org
  - metadata-librarians@lists.monarchos.com
  - lita-l@ala.org

- Study groups, anyone?
- Ask questions!
Thanks & Acknowledgements

 Thanks for your attention!
  • Some slides were from Gordon Dunsire’s presentation to IFLA RDA Satellite conference

 Contact for Diane:
  • Email: metadata.maven@gmail.com
  • Website (under construction): http://managemetadatadata.com/