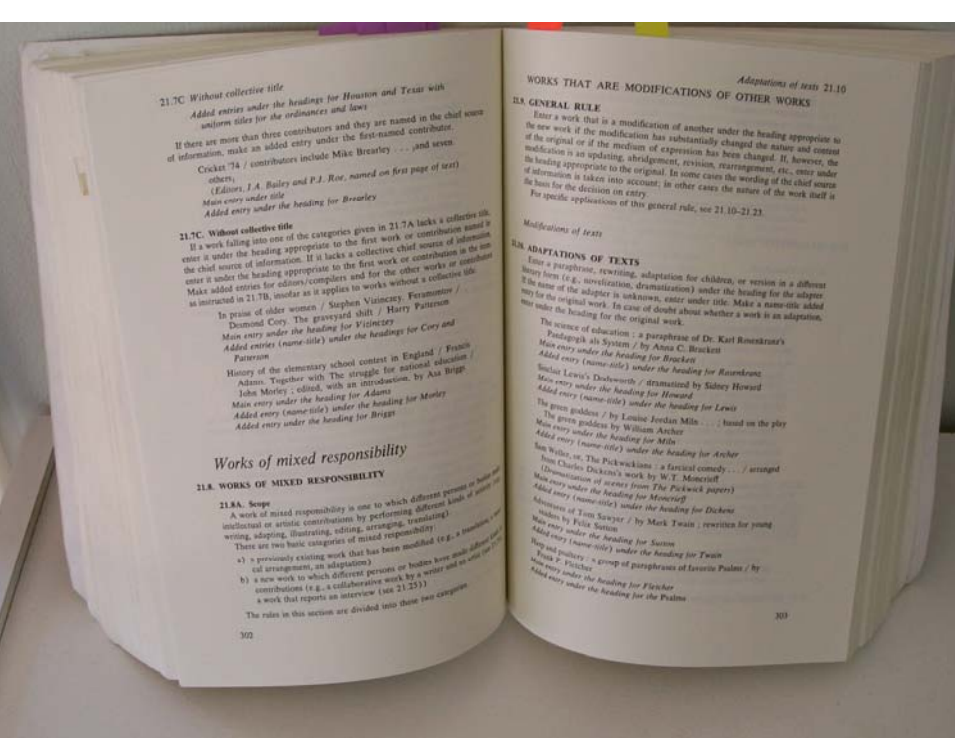# R&D: RDA in RDF

## or:
## Can Resource Description become Rigorous Data?

**Karen Coyle**
**Code4Lib 2008**

Annotated version, March, 2008

This is AACR2*. AACR2 is over 600 pages long, and it is the set of cataloging rules that librarians use when cataloging a book… or a film, or a piece of music, or anything else that will be cataloged.

The cataloging rules are long and complex because the things they catalog are more complex than most of us can imagine. For example, these rules include the proper way to transcribe the names of Thai royalty; and when you have a book written by a spirit speaking through a medium, which one of those is the author? (Note, the answer to that last one changed between AACR1 and AACR2.)

*Anglo-American Cataloging Rules, 2nd Edition

**3.5B2.** If there is more than one map, plan, etc., on a sheet, specify the number of maps, etc.

6 maps on 1 sheet

AACR2 consists of instructions and examples. Here is a short instruction and the example of what the resulting bit of cataloging might look like.

Note that the cataloging is expressed as text. AACR2 was first issued in 1978 when card catalogs were still the predominant form of catalog. A catalog card, of course, is a text document.

```
025.32        Gorman, Michael
Gorman              The concise AACR2, 1998 revision / prepared by Michael
              Gorman. – Chicago : American Library Assoc., 1999.
                     168 p. ; 23 cm.

                     Includes index
                     ISBN 0-8389-3494-3

                     1. Anglo-American cataloguing rules.  2. Descriptive
              cataloging—Rules. I. Anglo-American cataloguing rules.
              II. Title. III. Title: Concise Anglo-American cataloguing
              Rules, 1998 revision.
```

This is a sample AACR2 catalog entry for AACR2.

It is a highly structured, rules-based text, but it is still a text.

(Examples of other structured texts are limericks and the address written on an envelope.)

| | |
|---|---|
| **Author** | Toole, Betty A. (Betty Alexandra) |
| **Title** | Ada, the enchantress of numbers : prophet of the computer age, a pathway to the 21st century / Betty Alexandra Toole. |
| **Edition** | Pbk. ed., rev. & abridged. |
| **Availability** | All items |
| **Location** | Hayden Library - Stacks | QA29.L72.T66 1998 |
| | |
| **Published** | Mill Valley, Calif. : Strawberry Press ; Sausalito, CA : Orders to Critical Connection, c1998. |
| **Description** | xvii, 323 p. : ill. ; 20 cm. |
| **Format** | Book |
| **Note** | Letters to and from the Countess of Lovelace, 1824-1852. |
| **Bibliography** | Includes bibliographical references (p. 309-313) and index. |
| **Subject** | Lovelace, Ada King, Countess of, 1815-1852 — Correspondence. |
| | Mathematicians — Great Britain — Correspondence. |
| | Calculators. |
| | Computers. |
| **Other Title** | Prophet of the computer age, a pathway to the 21st century. |
| **ISBN** | 0912647183 (pbk.) : |

You can take that structured text and mark it up using MARC21, put it in a database and display it on a screen. It is still primarily a group of text strings that are intended to be read by human beings. In fact, I tend to think of the MARC21 format as a mark-up language, defining the structure of text, but far from a data processing record.

# AACR3

Recognizing that these are different times, that all catalog data will be processed by computers; that there is a whole World Wide Web that is the primary information space for our users, the Joint Steering Committee on AACR began work in 2004 (or so) on the next generation catalog rules.

# RDA = Resource Description and Access

It was soon determined that the new cataloging rules would be so vastly different from the ones that came before, that they no longer fit into the AACR tradition. The rules were renamed to RDA.

The principle participants, however, still reflect an Anglo-American nature.

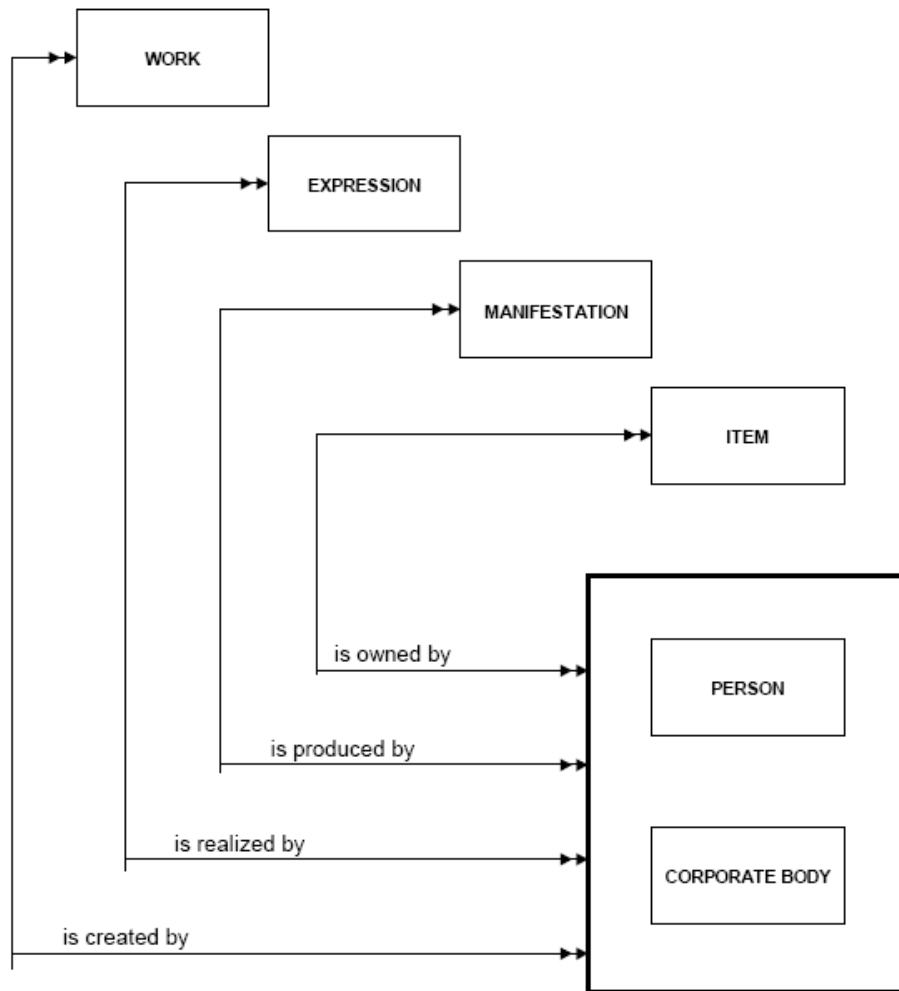A project of the Joint Steering Committee for RDA

- Library of Congress
- British Library
- Collections and Archives Canada
- Australian National Library

This is the RDA logo. Does anyone else see a tetris game in here?

# Goals of RDA

# FRBR



RDA has some very interesting and laudable goals. To begin with, RDA uses (or attempts to use) the Functional Requirements for Bibliographic Data as its framework. I say "attempts to use" because the is the first implementation of FRBR as cataloging rules, and in a sense is a test of FRBR as a model.

# Goals of RDA

# SIMPLIFY

Another goal of RDA was to simplify the cataloging rules. Simplification and modularization will encourage communities to use RDA who before might have found it to be too rigid or too library-centric.

# Goals of RDA

## "Other" communities



The main "other communities" near the library space are museums and archives. This need to bring them together in a single data format is especially acute for institutions that have multiple roles: museums that also have libraries, libraries that include an archive, etc. They need to create compatible data, but the different functions often have very different metadata needs at some level.

# Goals of RDA

Dis -- play

DISPLAY

Play, Dis

♋♓♦□●♋⬀

Another goal of RDA, as compared to AACR1 & 2, is to be less prescriptive about display. In fact, the AACRs are very much about *presentation* of information, not its storage or manipulation. This is a symptom of the fact that those cataloging rules were directions for the creation of text displays.

# Goals of RDA

## Machine processing

RDA also has a goal of being mindful of the fact that today's bibliographic data will exist in a computerized, networked world. All of the elements created by catalogers will be processed by computers; some will be presented to humans to be read.

Here are two pages from an RDA draft.

RDA consists of:
- 10 sections
- with 37 chapters
- and 13 appendices

To date we have seen only about ½ of the RDA text (and it is today very much a text). Counting up the available drafts, I found over 800 pages of text, not including any of the appendices.

There are numerous problems evidenced here. One is that such a lengthy text is unlikely to be a simplification of the previous rules, and there have been many complaints from the community that RDA is much, much too complex.

| | Exceptions: |
|---|---|
| | **a) Cartographic resources** |
| 3.4.0.3.1a.1 | Record the extent of a cartographic resource by giving the number of maps, etc., indicating, as appropriate, the number of sheets or segments, as instructed under 3.4.1. |
| | 1 map |
| | 3 diagrams |
| | 1 view in 3 segments |
| | 6 maps on 1 sheet |

Another problem is that RDA continues to instruct its adherents to create text strings. Although in many cases the actual instruction has changed (and catalogers consider some of these changes to be significant), the resulting output is still: "6 maps on 1 sheet"

A string like "6 maps on 1 sheet" is fine for a person, but if you want to do machine processing on the data, a text string just doesn't cut it.

I should mention here that RDA is not being designed as a print document. It will be an online service, with the sections and references hyperlinked. This is the explanation for some of the repetition of instructions and definitions throughout the text. The draft itself is an MS Word document with a combination of formatting and print clues to the text's structure, such as bold fonts and textual numbering. Moving from this text to an online service is a serious challenge.

**Audio carriers**
audio cartridge
audio cylinder[1]
audio disc
audio film reel[2]
audio roll[3]
audiocassette
audiotape reel

**Computer carriers**
computer card
computer chip cartridge
computer disc
computer disc cartridge
computer tape cartridge
computer tape cassette
computer tape reel
online resource[4]

**Microform carriers**
aperture card
microfiche
microfiche cassette
microfilm cartridge
microfilm cassette
microfilm reel
microfilm slip
microopaque

**Microscopic carriers**
microscope slide

On the previous page we saw the example "6 maps on 1 sheet." It may not be clear from the example, but some terms in that statement are based on controlled vocabularies.

There are about 55 separate controlled vocabularies embedded in RDA. This means that the lists are part of the text of the RDA document, which makes it difficult to provide support for these terms in any systems based on the cataloging rules. Each system must create and keep its own version of the list, and updates to the vocabulary lists must be done redundantly in hundreds or thousands of systems.

**Audio carriers**
audio cartridge
audio cylinder[1]
audio disc
audio film reel[2]
audio roll[3]
audiocassette
audiotape reel

**Computer carriers**
computer card
computer chip cartridge
computer disc
computer disc cartridge
computer tape cartridge
computer tape cassette
computer tape reel
online resource[4]

**Microform carriers**
aperture card
microfiche
microfiche cassette
microfilm cartridge
microfilm cassette
microfilm reel
microfilm slip
microopaque

**Microscopic carriers**
microscope slide

○ h – Microform
– Specific material design
licates the special class of
○ a – Aperture card
○ b – Microfilm cartridge
○ c – Microfilm cassette
○ d – Microfilm reel
○ e – Microfiche
○ f – Microfiche cassette
○ g – Microopaque
○ u – Unspecified
○ z – Other

MARC

Similar lists in the MARC21 format have the same problem. Updates are announced in email, and system developers must manually update their versions. For lists that are embedded in the MARC21 standard documentation, it takes a standards update request, often a 2-year process, to get a new value approved for a list, and more time before the new value is added to systems.

By the way, RDA and the MARC21 standard have some lists in common, but more that they do not share.

— **O P I N I O N** —

# Resource Description and Access (RDA)

## Cataloging Rules for the 20th Century

Karen Coyle
kcoyle.net
<kcoyle@kcoyle.net>

Diane Hillmann
Cornell University
<dih1@cornell.edu>

Not one to suffer in silence, I teamed up with my "partner in crime," Diane Hillmann, to write an article for D-Lib Magazine about RDA. Our subtitle appears to have been too subtle: "Cataloging Rules for the 20$^{th}$ Century" was actually a scathing comment on the deficiencies of RDA as a modern cataloging code.

In this article, we proposed that a modern cataloging code must not be text-based but must be based on a solid framework that can be supported in a computing environment. Since then, I have become even more interested in how we might make use of the principles of the semantic web to make library data more visible and useful in the networked world we live in.

# Dublin Core Singapore Framework
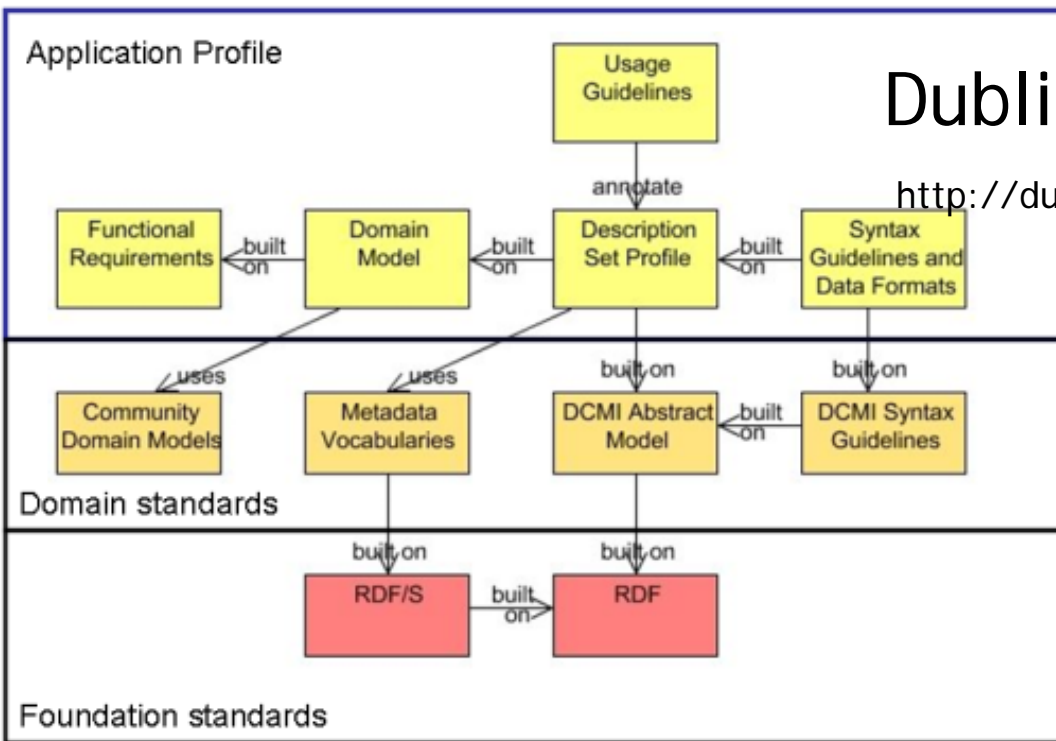
This multi-layered model is similar to the one that I have in mind, although mine is less detailed, and when I tried to draw it, it was embarrassingly silly looking. So I'll use this Dublin Core design to illustrate the issues I wish to convey.

Essentially, to have functional, coherent, machine-actionable metadata, you need, first, to have a foundation based in IT standards. Above that, you have to clearly define your domain (eg. "metadata for library and archival resources"). Your domain needs clear functional requirements (FRBR has at least some of this). You also need to have formally defined vocabularies (we'll see more of that in a bit). In most cases, you also want to be able to make use of application profiles. These allow you to define different variations of your data for different users in your domain. So general libraries and specialized libraries could have many elements in common, but each could also extend or reduce the vocabulary set that they use based on their needs.

```
Application Profile

                                    Usage
                                  Guidelines          RDA

                                     annotate

   Functional     built     Domain     built    Description    built      Syntax
  Requirements     on       Model       on     Set Profile      on     Guidelines and
                                                                        Data Formats

            uses              uses          built on           built on

   Community              Metadata       DCMI Abstract  built    DCMI Syntax
  Domain Models          Vocabularies      Model         on      Guidelines

Domain standards

                     built on           built on

                        RDF/S    built     RDF
                                  on

Foundation standards
```

RDA is the top box on this diagram, the "usage guidelines." It provides instructions on how to assign the values used by a community. FRBR may cover the functional requirements and the domain model, in yellow. The "Metadata Vocabularies" in the orange box are implicit in the RDA document, but are not formally defined in the sense intended here. Most of the needed structure, however, is missing from our metadata standard.

# http://dublincore.org/dcmirdataskgroup

**Dublin Core Metadata Initiative®**
*Making it easier to find information.*

> FrontPage

### DCMI/RDA Task Group Wiki

This Task Group is for collaborative work to enable broader use of the Resource Description and Access (RDA), building on agreements made at a ●meeting held at the British Library April 30/May 1, 2007. Participants in the meeting came from DCMI and other Semantic Web groups, and the RDA development effort. The Task Group is led by Diane Hillmann of Cornell University and Gordon Dunsire of Strathclyde University.

The planning for the London meeting was originally located on the ●DC-Libraries Wiki Some information gathered for the meeting, including documents, agendas, a meeting packet, and other information of historical interest remains at that location.

### Wiki Pages

- Use cases Use Cases
- Analysis task 1 analysisTask1
- Analysis task 2
  - List of in-line vocabularies from RDA RDAVocab
  - List of in-line vocabularies from MARC (for information) MARCVocab

### Charter and Work Plan

In a meeting in London on April 30, 2007 something extraordinary occurred. Representatives of JSC and the Dublin Core Metadata Initiative (some of whom also very active in the W3C semantic web activity) agreed that RDA must have a formal declaration of its vocabulary.

A group was formed to work on this project. Everyone agreed that it was essential to the success of RDA as a modern, web-based metadata standard.

Unfortunately, no one was in a position to fund it, but work is going forward none-the-less.

# RDA Vocabulary Project

- Declare all data elements in RDA in an RDF-based vocabulary (SKOS? OWL?)

- Assign URIs to all declared elements

- Declare all value vocabularies in an RDF-based vocabulary

- Register vocabularies online in an open repository

- Allow the creation of application profiles

The essential goals of the project are to create a machine-actionable registry of RDA vocabulary terms that is well-defined and open. Anyone wishing to make use of the terms would be able to do so.

**NSDL REGISTRY**

Supporting Metadata Interoperability

**Vocabulary: Show detail for RDA Carrier**

| Detail | Concepts | History | Versions | Maintainers |

**Detail**

| | |
|---|---|
| Owner: | Authority List Maintainers Association |
| Name: | RDA Carrier |
| URL: | |
| Note: | Based on list at RDA Chapter 3 rev. : 3.3.0.2.2. |
| Community: | Libraries |
| Status: | Published |
| Language: | English |

**URI**

| | |
|---|---|
| Base Domain: | http://authoritylists.info/uri/ |
| Token: | RDACarr |
| URI: | http://authoritylists.info/uri/RDACarr |

**Users +**

At about this same time I had one of those "ta-da" experiences. It was 2007, around tax time in the US (March or April). I had been reading RDA and had noticed the embedded vocabularies, in particular the "carrier" vocabulary that I showed earlier. This is a list of all of the physical formats for resources, including all of the computer formats.

At the office supply store I ran into something new: tax software being sold on a thumb drive. I looked at the RDA list of carriers, and thumb drives were not there.

So I blogged it: http://kcoyle.blogspot.com/2007/03/theres-always-something-new.html

# NSDLREGISTRY

Supporting Metadata Interoperability

## Vocabulary: Show detail for RDA Carrier

| Detail | Concepts | History | Versions | Maintainers |
|--------|----------|---------|----------|-------------|

### Detail

| | |
|---|---|
| Owner: | Authority List Maintainers Association |
| Name: | RDA Carrier |
| URL: | |
| Note: | Based on list at RDA Chapter 3 rev. : 3.3.0.2.2. |
| Community: | Libraries |
| Status: | Published |
| Language: | English |

### URI

| | |
|---|---|
| Base Domain: | http://authoritylists.info/uri/ |
| Token: | RDACarr |
| URI: | http://authoritylists.info/uri/RDACarr |

### Users +

RDA doesn't provide a way to update its value vocabulary lists. I made the suggestion that these terms should be outside of the text of the rules, and should be managed in an online registry.

Diane Hillmann, who had been working on the registry of vocabularies in the NSDL Registry project, took up the challenge and created a sample registry entry for the RDA carrier vocabulary.

http://sandbox.metadataregistry.org/vocabulary/show/id/44.html

| | Detail | Concepts | History | Versions | Maintainers |
|---|---|---|---|---|---|

| | Preferred Label | URI | Status | Updated |
|---|---|---|---|---|
| | Aperture card | .../uri/RDACarr/1024 | Published | 2007-04-11 11:45 |
| ✔ | Audio carrier | .../uri/RDACarr/1012 | Published | 2007-04-04 17:21 |
| | Audio cartridge | .../uri/RDACarr/1015 | Published | 2007-04-11 11:26 |
| | Audio cylinder | .../uri/RDACarr/1016 | Published | 2007-04-11 11:26 |
| | Audio disc | .../uri/RDACarr/1017 | Published | 2007-04-11 11:26 |
| | Audio film reel | .../uri/RDACarr/1018 | Published | 2007-04-11 11:26 |
| | Audio roll | .../uri/RDACarr/1019 | Published | 2007-04-11 11:27 |
| | Audiocassette | .../uri/RDACarr/1020 | Published | 2007-04-11 11:27 |
| | Audiotape reel | .../uri/RDACarr/1021 | Published | 2007-04-11 11:27 |
| | Card | .../uri/RDACarr/1048 | Published | 2007-04-15 13:18 |
| | Computer card | .../uri/RDACarr/1002 | Published | 2007-03-24 8:33 |
| ✔ | Computer carrier | .../uri/RDACarr/1011 | Published | 2007-04-04 17:20 |
| | Computer chip cartridge | .../uri/RDACarr/1003 | Published | 2007-03-24 8:33 |
| | Computer disc | .../uri/RDACarr/1004 | Published | 2007-03-24 8:34 |
| | Computer disc cartridge | .../uri/RDACarr/1005 | Published | 2007-03-24 8:34 |
| | Computer tape cartridge | .../uri/RDACarr/1006 | Published | 2007-03-24 8:34 |
| | Computer tape cassette | .../uri/RDACarr/1007 | Published | 2007-03-24 8:34 |
| | Computer tape reel | .../uri/RDACarr/1008 | Published | 2007-03-24 8:35 |

Diane entered all of the RDA Carriers into the registry, and coded the relationships between them (broader term, narrower term, etc.)

A key aspect of the registry concept is that each element has a Uniform Resource Identifier (URI). This means that wherever the identified term is used it always has the same semantic value. This is essential to promote interoperability, and it also supports internationalization, since the linguistic value of the term can change as long as the URI remains the same.

# NSDLREGISTRY

**Supporting Metadata Interoperability**

## Vocabulary: RDA Carrier
## Concepts: USB flash drive

| Detail | Properties | History |
|---|---|---|

### Detail

| | |
|---|---|
| Preferred Label: | USB flash drive |
| Language: | English |
| URI: | http://authoritylists.info/uri/RDACarr/1001 |
| Top Concept?: | |
| Status: | Published |

### Properties

| | | | |
|---|---|---|---|
| preferred label | USB flash drive | | Published |
| history note | Definition from Wikipedia. | English | Published |
| definition | USB flash drives are NAND-type flash memory data storage devices integrated with a USB (universal serial bus) interface. | English | Published |
| has broader | Computer carrier | English | Published |
| alternative label | Thumb drive | English | Published |
| alternative label | Flash drive | English | Published |
| alternative label | Jump drive | English | Published |

List | Get RDF

Then Diane created an entry for the "new" carrier, the "USB flash drive," based loosely on my blog post. Initially set as "provisional," this record showed how a vocabulary list could be extensible, and could also be available to everyone in the interested community online and at the same time. The entry includes information that we do not have today in the RDA nor MARC21 vocabularies, such as definitions, alternate terms, and hierarchical relationships. (Some of this information is available elsewhere, but not by any means at every point where you need to think about the term.)

```xml
- <rdf:RDF>
    <!-- WARNING: This is a single-concept fragment -->
    <!-- Scheme: RDA Carrier -->
  - <skos:ConceptScheme rdf:about="http://authoritylists.info/uri/RDACarr">
      <dc:title>RDA Carrier</dc:title>
      <skos:hasTopConcept rdf:resource="http://authoritylists.info/uri/RDACarr/1012"/>
      <skos:hasTopConcept rdf:resource="http://authoritylists.info/uri/RDACarr/1011"/>
      <skos:hasTopConcept rdf:resource="http://authoritylists.info/uri/RDACarr/1023"/>
      <skos:hasTopConcept rdf:resource="http://authoritylists.info/uri/RDACarr/1032"/>
      <skos:hasTopConcept rdf:resource="http://authoritylists.info/uri/RDACarr/1034"/>
      <skos:hasTopConcept rdf:resource="http://authoritylists.info/uri/RDACarr/1044"/>
      <skos:hasTopConcept rdf:resource="http://authoritylists.info/uri/RDACarr/1047"/>
      <skos:hasTopConcept rdf:resource="http://authoritylists.info/uri/RDACarr/1053"/>
    </skos:ConceptScheme>
    <!-- Concept: USB flash drive  -->
  - <skos:Concept rdf:about="http://authoritylists.info/uri/RDACarr/1001">
      <skos:inScheme rdf:resource="http://authoritylists.info/uri/RDACarr"/>
      <skos:prefLabel>USB flash drive</skos:prefLabel>
    - <skos:definition>
        USB flash drives are NAND-type flash memory data storage devices integrated with a USB (universal serial bus) interface.
      </skos:definition>
      <skos:altLabel>Thumb drive</skos:altLabel>
      <skos:altLabel>Flash drive</skos:altLabel>
      <skos:altLabel>Jump drive</skos:altLabel>
      <skos:historyNote>Definition from Wikipedia.</skos:historyNote>
      <skos:broader rdf:resource="http://authoritylists.info/uri/RDACarr/1011"/>
    </skos:Concept>
  </rdf:RDF>
```

When accessed by a program (as opposed to a human being), the registry provides a machine-readable record (in this instance in XML, but other formats could be generated). This means that library systems, and other bibliographic systems, throughout the Web can retrieve this data whenever it is needed. Catalogers and users can be shown definitions or related terms, and program requests can get a response that is directly usable as code.

| | RDA element | Classification | \<indecs\> attribute type | Domain | Value surrogate | Value string | Syntax encoding scheme | Vocabulary encoding scheme |
|---|---|---|---|---|---|---|---|---|
| 1 | | | | | | | | |
| 2 | **IDENTIFYING MANIFESTATIONS AND ITEMS** | | | | | | | |
| 3 | **Title** | element | label | manifestation | - | - | - | - |
| 4 | Title proper | element sub-type | label | manifestation | literal | plain | n/a | n/a |
| 5 | Parallel title | element sub-type | label | manifestation | literal | plain | n/a | n/a |
| 6 | Alternative title | element sub-type | label | manifestation | literal | plain | n/a | n/a |
| 7 | Parallel alternative title | element sub-type | label | manifestation | literal | plain | n/a | n/a |
| 8 | Other title information | element sub-type | label | manifestation | literal | plain | n/a | n/a |
| 9 | Parallel other title information | element sub-type | label | manifestation | literal | plain | n/a | n/a |
| 10 | Variant title | element sub-type | label | manifestation | literal | plain | n/a | n/a |
| 11 | Earlier/later variant title | element sub-type | label | manifestation | literal | plain | n/a | n/a |
| 12 | Key title | element sub-type | label | manifestation | literal | plain | n/a | n/a |
| 13 | Abbreviated title | element sub-type | label | manifestation | literal | plain | n/a | n/a |
| 14 | Devised title | element sub-type | label | manifestation | literal | plain | n/a | n/a |
| 15 | **Statement of responsibility** | element | label | manifestation | - | - | - | - |
| 16 | Statement of responsibility relating to title | element sub-type | label | manifestation | literal | plain | n/a | n/a |
| 17 | Parallel statement of responsibility relating to title | element sub-type | label | manifestation | literal | plain | n/a | n/a |
| 18 | **Edition statement** | element | label | manifestation | - | - | - | - |
| 19 | *Statement designating edition* | sub-element | label | manifestation | literal | plain | n/a | n/a |
| 20 | *Parallel statement designating edition* | sub-element | label | manifestation | literal | plain | n/a | n/a |
| 21 | *Statement of responsibility relating to the edition* | sub-element | label | manifestation | literal | plain | n/a | n/a |
| 22 | *Parallel statement of responsibility relating to the edition* | sub-element | label | manifestation | literal | plain | n/a | n/a |

Work on the RDA vocabularies is based on a JSC document that lists the RDA "elements." Although we have the list of elements there are interesting issues that come up in trying to format these as a formal vocabulary.

RDA Element Analysis:
http://www.collectionscanada.gc.ca/jsc/docs/5rda-elementanalysisrev.pdf

# Issues/Problems/Puzzles

**"Title Proper"**   = surrogate for title page
                     = access point
                     = display element
                     = sort order

To begin with, many elements serve more than one function in the bibliographic description, and most of these functions are implicit, not explicit. This has always been the case with library data, and it is definitely the case with data that we have coded in MARC format.

For the creation of the vocabularies we have to ask: how many of these functions need their own element?

# Issues/Problems/Puzzles

**New York, *Random House*, 2007**

**"Publisher name"**

It's natural for those of us in the data processing world to look at the publication statement with place, publisher, and date and visualize "publisher" as an entity in itself, perhaps in a separate record that contains the publisher's address and links to all of the books that it has produced. But in RDA, this isn't the "publisher" it's the string representing the publisher's name that appears on the title page. If the publisher listed there is wrong or fictitous, that's what goes into the description. Elsewhere, a note might say that the real publisher is "X."

We have to look very carefully at how the data elements are defined in RDA.

# Issues/Problems/Puzzles

**Tolkien, J. R. R. (John Ronald Reuel),  1892-1973**

**"Preferred form of name"**

Author names are almost the opposite of the publisher. The publisher name is part of the "surrogate" for the title page. The author name – a horridly ugly thing that we impose on our users – is constructed by catalogers. The interesting question here is: what does this represent? Although the FRBR entity is "person," the entry in the description isn't for a person but for a personal *name*. A real world person could be represented by more than one name (eg. both Mark Twain and Samuel Clemens are each "preferred names" in their own right); a personal name could represent two or more real world persons who write together under a single pseudonym. Yet the entry has birth and death dates that are information about a real world person.

This one is a real puzzle.

# Issues/Problems/Puzzles

- "Identification of the resource based on: Pt. 2, published 1998"

  – RDA includes meta-metadata; that is, data about the description and the cataloging decisions. This is important data, but it doesn't belong in the vocabulary that will define the resource.

- "Publisher statement" = place+publisher+date

  – Some elements are simply combinations of other elements: the publisher statement exists only as a combination of place, publisher, and date. Should the statement itself be in the vocabulary, or just the individual elements?

- "Edition statement" = "5th edition, revised"

  – Some elements are just text strings, even though they may contain more than one data value (the "6 maps on 1 sheet" is another example of this). Do these need to be divided into separate elements?

# Issues/Problems/Puzzles

**The two towers : being the second part of The lord of the rings / by J.R.R. Tolkien ; with a new foreward by the author.**

I have a particularly hard time with the element called "statement of responsibility." This is traditionally a key element in library cataloging, but I can't figure out what to do with it in a vocabulary declaration. It doesn't stand alone: "by J.R.R..." is intended as a continuation of the title. In this case, it looks like the title is doing at least double duty, which may be the problem. Taken together, the title + statement of responsibility are a surrogate for the title page. But the title here is also the title entry that will be used, without the statement of responsibility, for retrieval and sorting.

# Issues/Problems/Puzzles
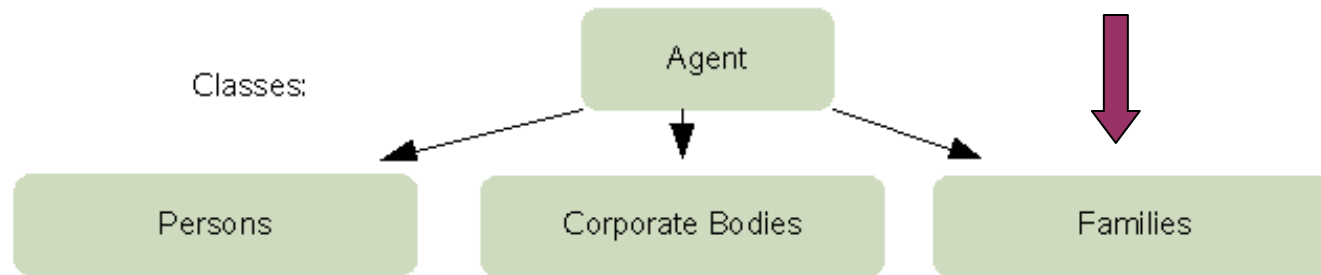
## Bartholomew world travel series, v. 5

Relationships, like that of a book to a series, are not expressed relationally in RDA but are text strings within a description of a resource, like this series statement. Text strings are "links" only in that humans can read them and search for the related bibliographic item. They aren't viable links for machine processing of data.

## Original title: L'éducation sentimentale

Even the key relationships like Expression to Work aren't always made clear. While some catalogs may have a "Work title" record in their authority file, it is also acceptable to indicate these relationships with a note. This isn't a question of right or wrong, but it tells me that we can't expect to impose a strict entity/relationship model on library bibliographic data.

# Issues/Problems/Puzzles

Top Level View – RDA Agent Domain

Classes:

Agent

Persons        Corporate Bodies        Families

Name(s)         Name(s)                Name(s)
Date(s)         Date                   Date
Details         Type                   Type
  Country       Place                  Details
  Language      Details                  Members
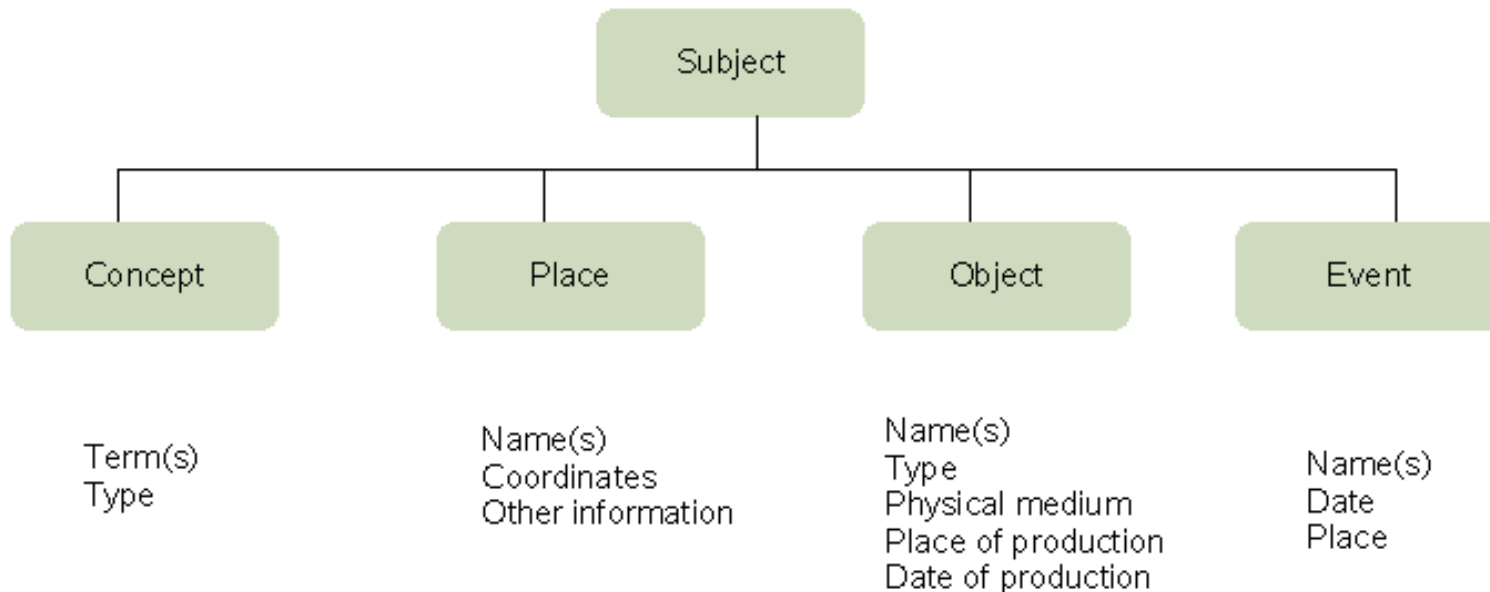  Place of birth    Address                Family history
  etc.            Language                 etc.
                  etc.

There is the question of what it means to say that RDA follows FRBR. In general, the FRBR "attributes" are not reflected in RDA. It also isn't clear how the lessons learned in creating RDA (probably the first real "test" of FRBR) will be reflected back in the FRBR model.

As an example, RDA includes "Families" as agents. This comes from FRAD, but not from FRBR. Is this significant? Is it important to keep these models synchronized? If not, what is their purpose as models?
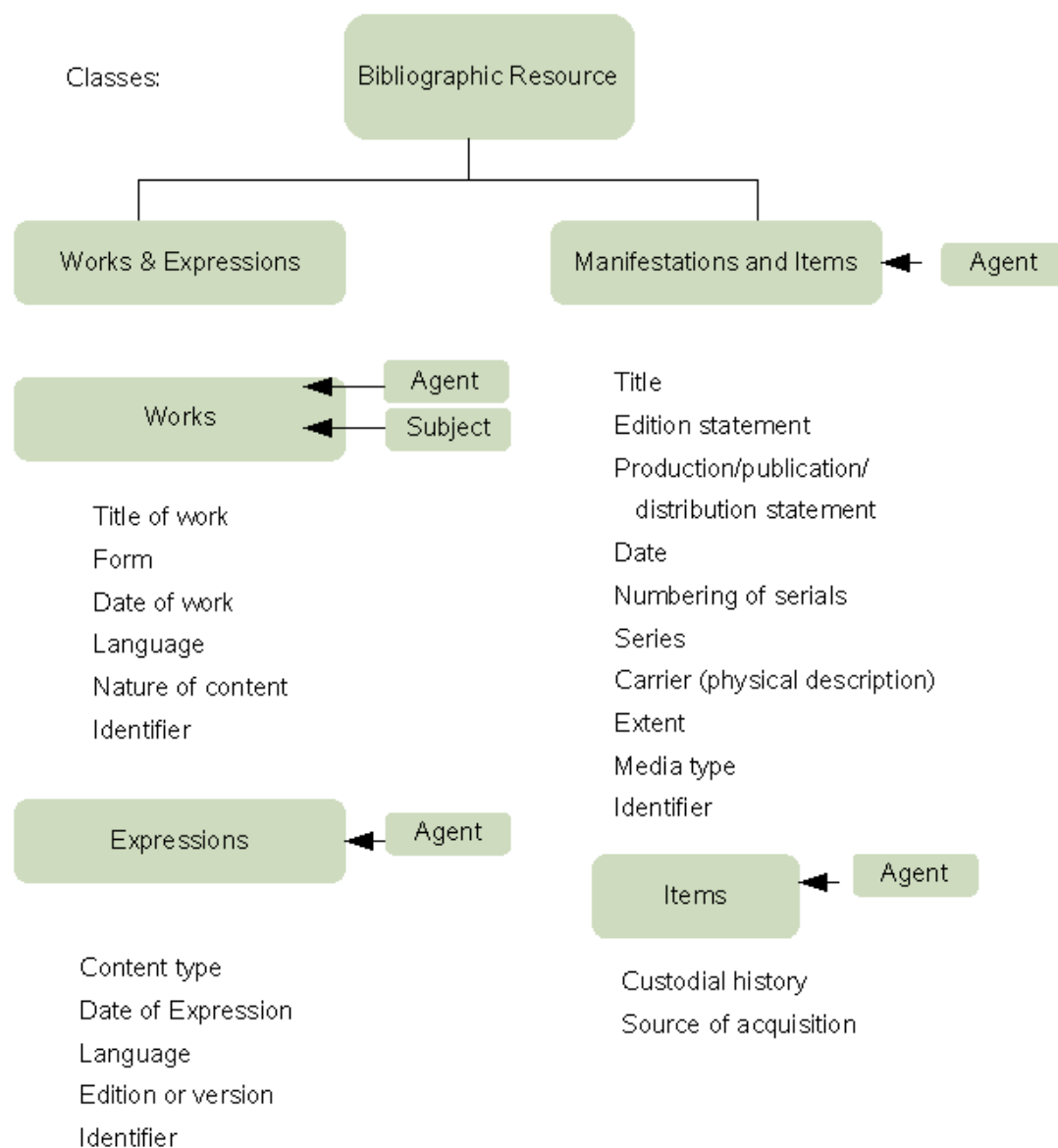
# Issues/Problems/Puzzles

Top Level View – RDA Subject Domain

```
                          ┌──────────┐
                          │ Subject  │
                          └────┬─────┘
          ┌──────────────┬─────┴──────┬──────────────┐
     ┌────┴────┐    ┌─────┴───┐   ┌────┴────┐    ┌────┴────┐
     │ Concept │    │  Place  │   │ Object  │    │  Event  │
     └─────────┘    └─────────┘   └─────────┘    └─────────┘
```

Concept
Term(s)
Type

Place
Name(s)
Coordinates
Other information

Object
Name(s)
Type
Physical medium
Place of production
Date of production

Event
Name(s)
Date
Place

RDA doesn't include subject analysis, although it has place-holders in the element table for the FRBR Group 3 entities. This is a very important aspect of library catalog records; how will this gap be filled in, and by whom?

## Top Level View – RDA Bibliographic Resource Domain

Classes:

**Bibliographic Resource**

- **Works & Expressions**
- **Manifestations and Items** ← Agent

**Works** ← Agent
← Subject

Title of work
Form
Date of work
Language
Nature of content
Identifier

**Expressions** ← Agent

Content type
Date of Expression
Language
Edition or version
Identifier

Title
Edition statement
Production/publication/
    distribution statement
Date
Numbering of serials
Series
Carrier (physical description)
Extent
Media type
Identifier

**Items** ← Agent

Custodial history
Source of acquisition

There's always a lot of discussion about the FRBR Group 1 entities, in particular about the boundaries between them. RDA has recently changed how it presents these. Group 1 has been divided into Content (Work and Expression) and Carrier (Manifestation and Item). There is no bright line between the two entities in each group. This isn't necessarily a problem, but it does bring up the question of whether the RDA elements can be definitively assigned to one Group I entity. If an element can be used either for a Work or an Expression, is that one vocabulary element or two different ones?

# FUQ*

- ## Is FRBR the right model?

  - I don't know, but RDA could be the real test of FRBR concepts.

- Is RDF the right format?

- Does this replace MARC21?

- Who will be in charge? How will things be decided?

- WHAT WERE YOU THINKING?!

- How can I help?

*FREQUENTLY UNANSWERED QUESTIONS

# FUQ*

- Is FRBR the right model?

- **Is RDF the right format?**

  – We're working with RDF because 1) it expresses entities and relationships 2) it is being actively worked on by W3C groups so there is support for it 3) people on the project are familiar with RDF. Our data will not be limited to RDF expression. Got a favorite format? Come chat with us.

- Does this replace MARC21?

- Who will be in charge? How will things be decided?

- WHAT WERE YOU THINKING?!

- How can I help?

*Frequently Unanswered Questions

# FUQ*

- Is FRBR the right model?

- Is RDF the right format?

- **Does this replace MARC21?**

  – No. What we're working on is not a record format. However, the vocabularies should make it possible to develop a new library data carrier, or multiple carriers, fairly easily.

- Who will be in charge? How will things be decided?

- WHAT WERE YOU THINKING?!

- How can I help?

*Frequently Unanswered Questions

# FUQ*

- Is FRBR the right model?

- Is RDF the right format?

- Does this replace MARC21?

- Who will be in charge? How will things be decided?

  - This is particularly difficult. We submitted a grant proposal to NSF to develop a registry maintenance methodology, including creating the community structures that would be needed. Their reply was: this will change everything in the library world, but it's not technically interesting. (Most technology is less interesting than its social implications.)

- WHAT WERE YOU THINKING?!

- How can I help?

*FREQUENTLY UNANSWERED QUESTIONS

# FUQ*

- Is FRBR the right model?

- Is RDF the right format?

- Does this replace MARC21?

- Who will be in charge? How will things be decided?

- **WHAT WERE YOU THINKING?!**

  - I often wonder about that myself. This is a huge task, yet one that some of us think is extremely important. And that leads us to the next question:

- How can I help?

*Frequently Unanswered Questions

# http://dublincore.org/dcmirdataskgroup

## Dublin Core Metadata Initiative®
*Making it easier to find information.*

> FrontPage

## DCMI/RDA Task Group Wiki

This Task Group is for collaborative work to enable broader use of the Resource Description and Access (RDA), building on agreements made at a meeting held at the British Library April 30/May 1, 2007. Participants in the meeting came from DCMI and other Semantic Web groups, and the RDA development effort. The Task Group is led by Diane Hillmann of Cornell University and Gordon Dunsire of Strathclyde University.

The planning for the London meeting was originally located on the DC-Libraries Wiki Some information gathered for the meeting, including documents, agendas, a meeting packet, and other information of historical interest remains at that location.

### Wiki Pages

- Use cases Use Cases
- Analysis task 1 analysisTask1
- Analysis task 2
  - List of in-line vocabularies from RDA RDAVocab
  - List of in-line vocabularies from MARC (for information) MARCVocab

### Charter and Work Plan

Follow the work on the wiki …

# http://jiscmail.ac.uk/archives/dc-rda.html



**List Archives**

Subscriber's Corner
Server Archives
List Archives

List Management
List Moderation
Server Management

Help
Log off

Archive Search

CONTACTS:

List Owner(s)

JISCmail Helpline

**Archives of DC-RDA@JISCMAIL.AC.UK**

*List for discussion on Resource Description and Access (RDA)*

- Search the archives
- Post to the list
- Join or leave the list (or change settings)
- Manage the list (list owners only)

- February 2008
- January 2008
- December 2007
- August 2007
- July 2007
- June 2007
- May 2007
- April 2006
- February 2006
- January 2006
- December 2005

JISCmail Tools

📁 Files Area

Surveys

Discussion Room (Help)

List Information

Subscribers:

Last Post:

Files:

Active Surveys:

Discussions:

And join the discussion group to participate in the project.

# Thank You

kcoyle@kcoyle.net