What is RDA and why do I need to know?

CLA Montreal Chapter

March 9th, 2011

Chris Oliver

McGill University

chris.oliver@mcgill.ca

RDA = Resource Description and Access

- new metadata standard replaces AACR2

- designed for the web environment

- better metadata to support better resource discovery

Plan

- why do we need a new standard?

- why is RDA important?

- key features of RDA

- when will we start seeing its effect?

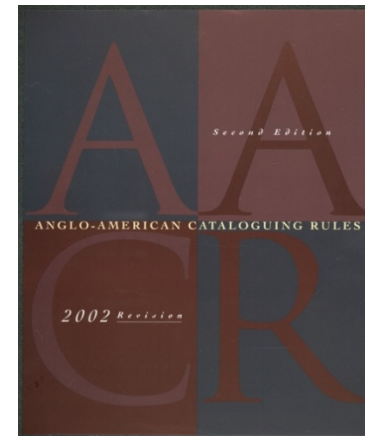 - implementation

  - after implementation

Why a new standard?

It's about time!

1978

Aging AACR2

- AACR2: 1978

- some patching and grafting through revisions

- robust amendment process, but too many underlying problems, many rooted in AACR2's structure

- last amendments: 2005

Problems with AACR2

Some examples of the problems:

- written in the context of card catalogues
  - e.g. abbreviations
  - e.g. limited number of contributors that could be included

- inadequate rules for the description of new types of resources, especially electronic resources

- lack of a theoretical framework to act as a reference point when dealing with new situations
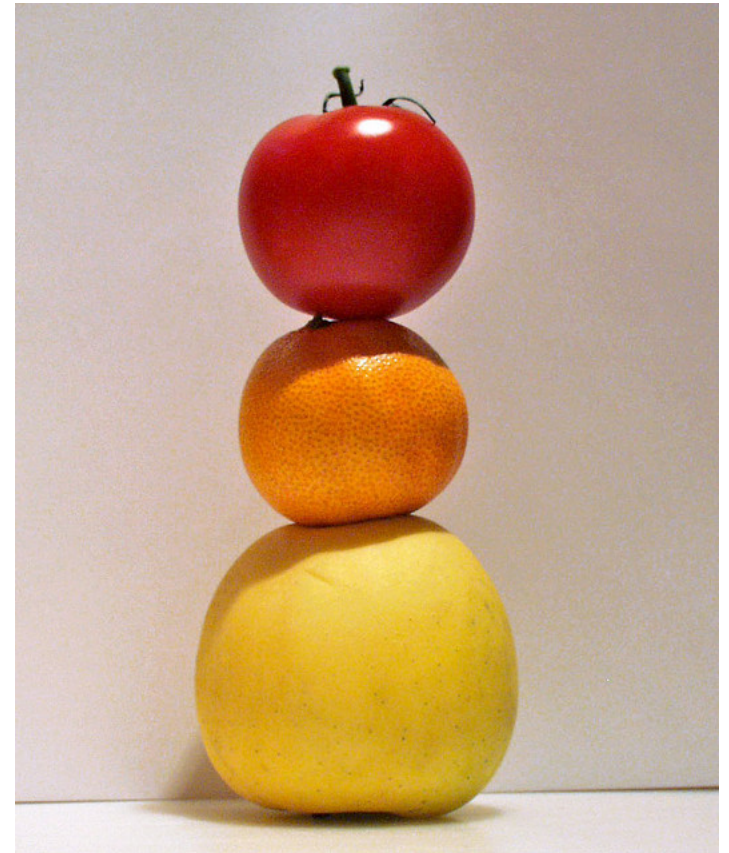
Timeline

1997   problems identified:

International Conference on the Principles & Future Development of AACR, Toronto, Ontario

1998-2004   revisions to AACR2

2004   AACR3

2005   new standard: Resource Description and Access

2009   RDA text completed

2010   RDA text + software – standard is a web tool

2011/2012   implementation

Who develops and supports RDA?
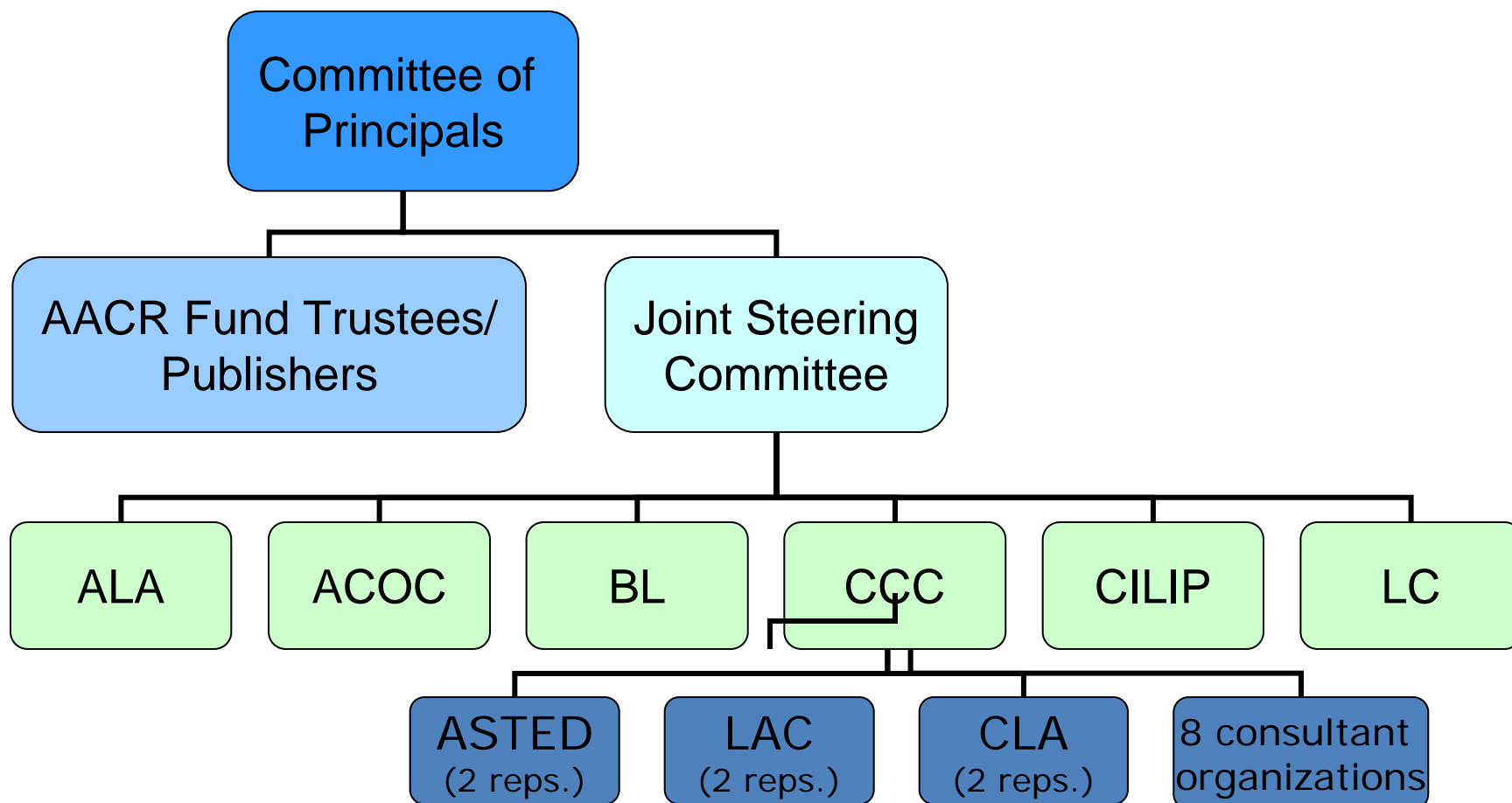
RDA and AACR2

- RDA carries forward some of the strong points of AACR2

- RDA goes beyond AACR2

  - not just for the library

  - designed for current and future environments

  - based on a sound theoretical model

  - flexible and extensible to support the description of all types of resources

  - connecting with other cultural heritage metadata communities

AACR2 ➡ RDA

- continue to record the title

- continue to record the edition

- continue to record the date of publication

But …

- new vocabulary

- new way of thinking about how we do these steps

You will hear …

- entities

- attributes

- relationships

- data elements

- core elements

- FRBR model   *Functional Requirements for Bibliographic Records*

- FRAD model   *Functional Requirements for Authority Data*

New terminology

~~headings~~   access points

~~physical description~~   description of the carrier

~~main entry~~   preferred access point
                                                     representing a work or
                                                     expression

~~see reference~~   variant access point

~~see also reference~~   authorized access point for
                                                     a related entity

Why is RDA important?

- designed to improve resource discovery for the user
  - puts the user at the centre
  - bibliographic data ⟷ how it is used

- designed as a metadata standard
  - data elements
  - instructions on how to record values for those elements
  - controlled vocabularies
  - overall structure governed by a formal model

Why is RDA important?

- get bibliographic data out of the library silo
  - visible on the web
  - interacting with the data of other metadata communities

- RDA is a content standard

  NOT an encoding standard

  RDA = what data should the cataloguer record?

- possible to encode using many encoding systems
  - can be encoded using MARC
  - does **not** have to be encoded using MARC encoding
  - can be used with web friendly XML based encoding schema, such as Dublin Core, MODS

Designed for now and the future

Now

- designed to work in current catalogues

- compatible with AACR2 records

- co-exist with AACR2 records in the same databases

Future

- ready to take advantage of new database structures

- function in the semantic web, linked data environment

- visible in the web alongside other types of metadata

Us and our legacy data

RDA …

- takes us from where we are

- moves us to a new track

- puts us on the right track for the next part of the voyage

Key features of RDA

- set of practical instructions built on the foundation of a theoretical framework/model

- data that responds to user needs

- accurate data

- precise data

- usable data

- visible data

- a standard with an expanded scope

RDA's framework

- explicit conceptual framework

- aligned with the FRBR and FRAD conceptual models

  FRBR   Functional Requirements for Bibliographic Records
  1998

  FRAD   Functional Requirements for Authority Data
  2009

  FRAD is an extension of the FRBR model

- both models developed under the auspices of IFLA

The two models

- widely used data modeling technique:

  entity relationship model
  - entities
  - attributes
  - relationships

- analyze bibliographic and authority data from the point of view of how that data is used

- careful distinction between content and carrier

>>> broad base of international consensus and support

Why are the models important?

broad international support for the explanatory power of the models

common international language and conceptual understanding of the bibliographic universe

as the foundation for a standard:

- easier to apply in international context

- easier for our data to interoperate with other data generated on the basis of a FRBR/FRAD understanding of the bibliographic universe

1st step: how is data used?

Bibliographic data

- find
- identify
- select
- obtain

Authority data

- find
- identify
- clarify
- understand

Bibliographic entities: of interest to user

work

expression

manifestation

item

---

person

family

corporate body

---

concept

object

event

place

FRBR Group 1

products of intellectual or artistic endeavor

FRBR Group 2

responsible for group 1 entities

FRBR Group 3

subjects   (includes group 1 & 2)

Attributes

- characteristics of the entity
  data to be recorded about the entity

examples of attributes:

work:	title, genre
manifestation:	publisher, date of publication, extent
item:	identifier (e.g. barcode), provenance
person:	dates, gender, a title of rank or office
corporate body:	place, dates, address
object:	term

Relationships: links between entities

expression	*translation of*	work
item	*exemplar of*	manifestation

work	*created by*	person
item	*owned by*	family
manifestation	*produced by*	corporate body

work	*based on*	work
manifestation	*electronic reprod.*	manifestation

|		
person	*member of*	family
family	*founded*	corporate body

Group 1 entities

work = a distinct intellectual or artistic creation

expression = intellectual or artistic realization of a work in the form of alpha-numeric, musical or choreographic notation, sound, image, object, movement, etc., or any combination of such forms

manifestation = physical embodiment of an expression of a work

item = single exemplar or instance of a manifestation

4 entities – 1 resource
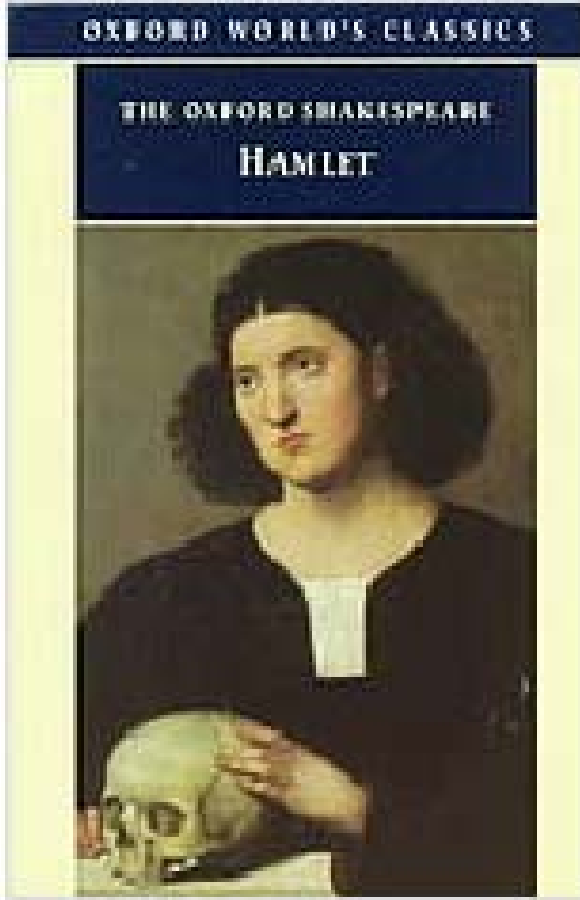
- an item

- an exemplar of the Oxford 1998 manifestation

- an embodiment of the original English expression

- a realization of the work, Hamlet

As a user ...

- I need to know Shakespeare's play Hamlet for an exam
  *I need to find the **work***

- I need to read Hamlet, and my mother tongue is French, so I want a French translation of Hamlet
  *I need to find a particular **expression***

- Our professor put the print 1998 Oxford publication on the reading list
  *I need a particular **manifestation***

- I need to find the copy I borrowed because it is overdue
  *I need a particular **item***

As library staff ...

- we build our collections by selecting **manifestations** that we want to buy
- we select manifestations
  - for the **expression** that the manifestation embodies
  - for a **manifestation** attribute such as the type of carrier (*e.g. print volume or online resource*)
- our bibliographic records are records for **manifestations**
- we buy, receive, label, barcode and loan **items** that are exemplars of the manifestation
- our subject headings are for the **work** or **expression**
- uniform titles bring together **manifestations** of the same **work** or **expression**
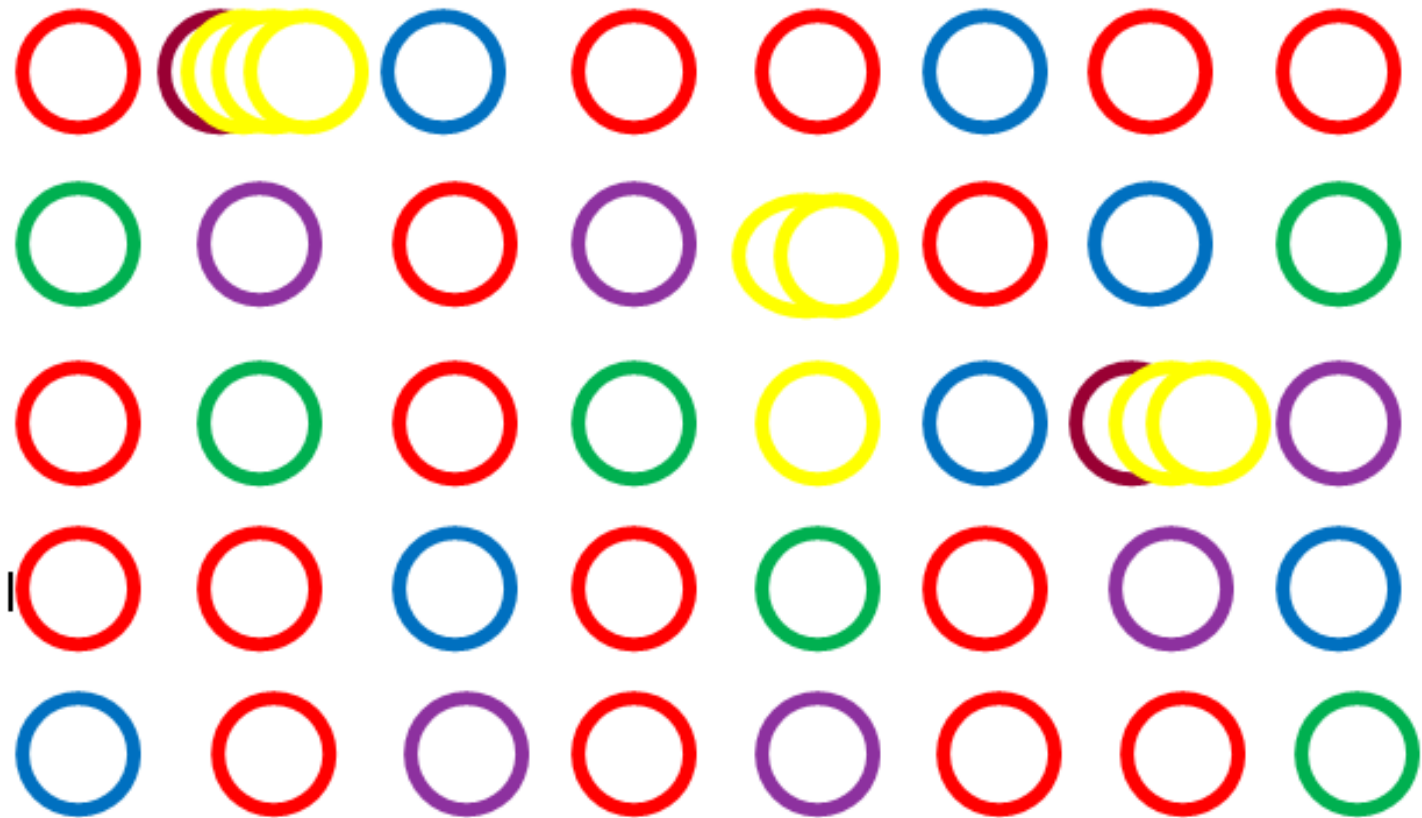
FRBR: leads to better grouping

- texts

- translations of text

- audio-books

- dramatizations

- adaptations as a motion picture

- criticisms

Search the current online catalog:
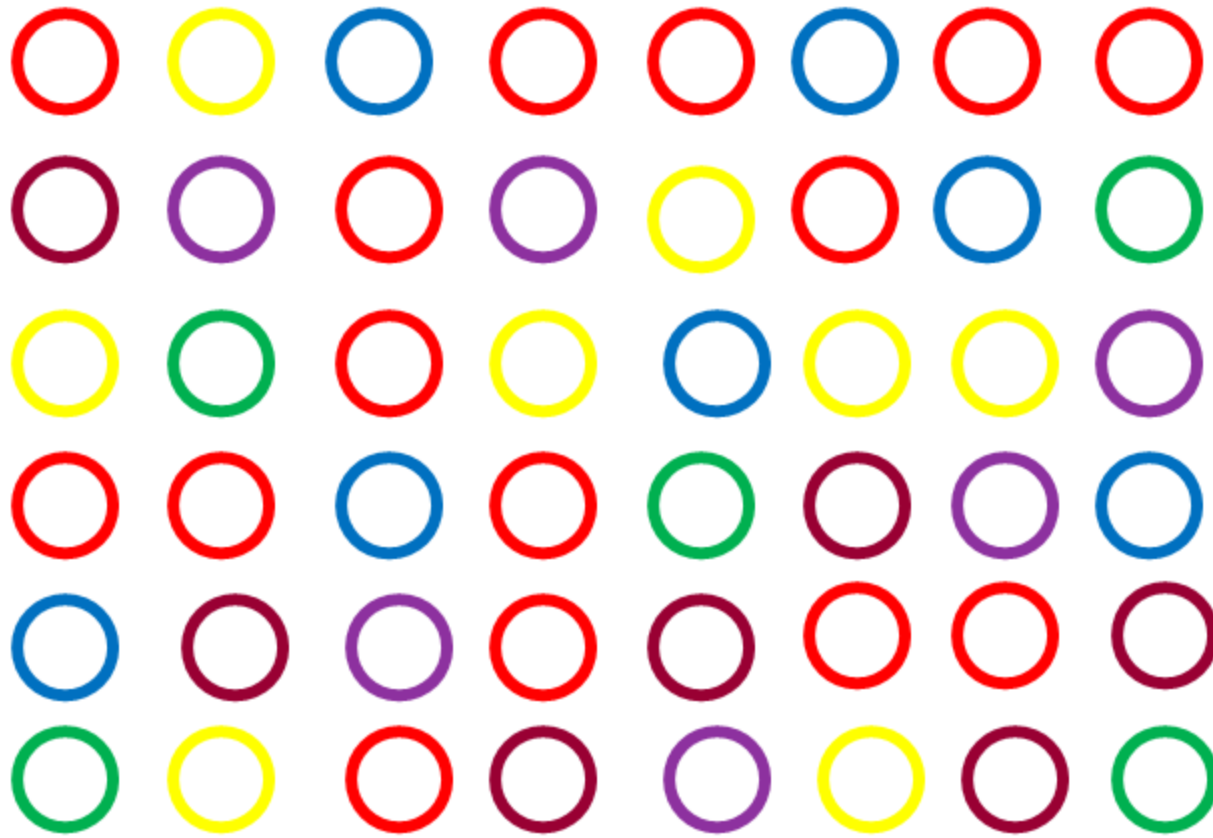
Slightly "FRBR-ized" catalogue

Truly FRBR-ized display of data

- texts

- translations of text

- audio-books

- dramatizations

- adaptations as a motion picture

- criticisms

RDA: responsiveness to user needs

0.4   Objectives and Principles Governing Resource Description and Access

- first objective: responsiveness to user needs

- record data that is important

- it is important because of how it is used

- use is operationalized as "user tasks"

Resource discovery = user tasks

Bibliographic data
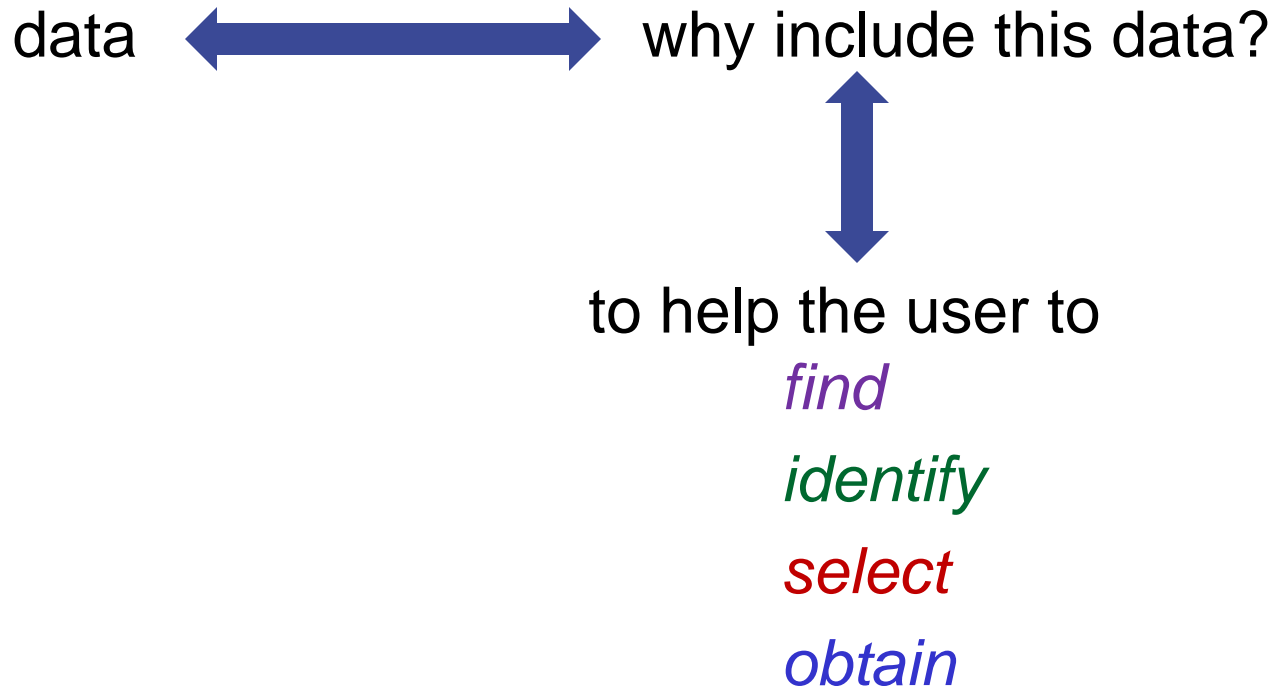
- find
- identify
- select
- obtain

Authority data

- find
- identify
- clarify
- understand

Consistent focus on the user

Throughout RDA:

data ⟷ why include this data?

to help the user to
find

identify

select

obtain

Definition of importance

- record data **if** it is considered important for …

for example, from *3.7 Applied material*

Record the applied material used in the resource if it is considered important for **identification** or **selection** …

Accurate and precise data

- still recording the same kind of information

  title   *edition*   *date of publication*

- record with more accuracy and precision

  - accurate representation

  - relationship designators

  - distinct data elements

  - more controlled vocabulary

Take what you see

- AACR2

Title: The wolrd [sic] of
   television.

 or

Title: The wolrd [i.e. world]
   of television.

Variant title: The world of
   television.

- RDA

Title: The wolrd of
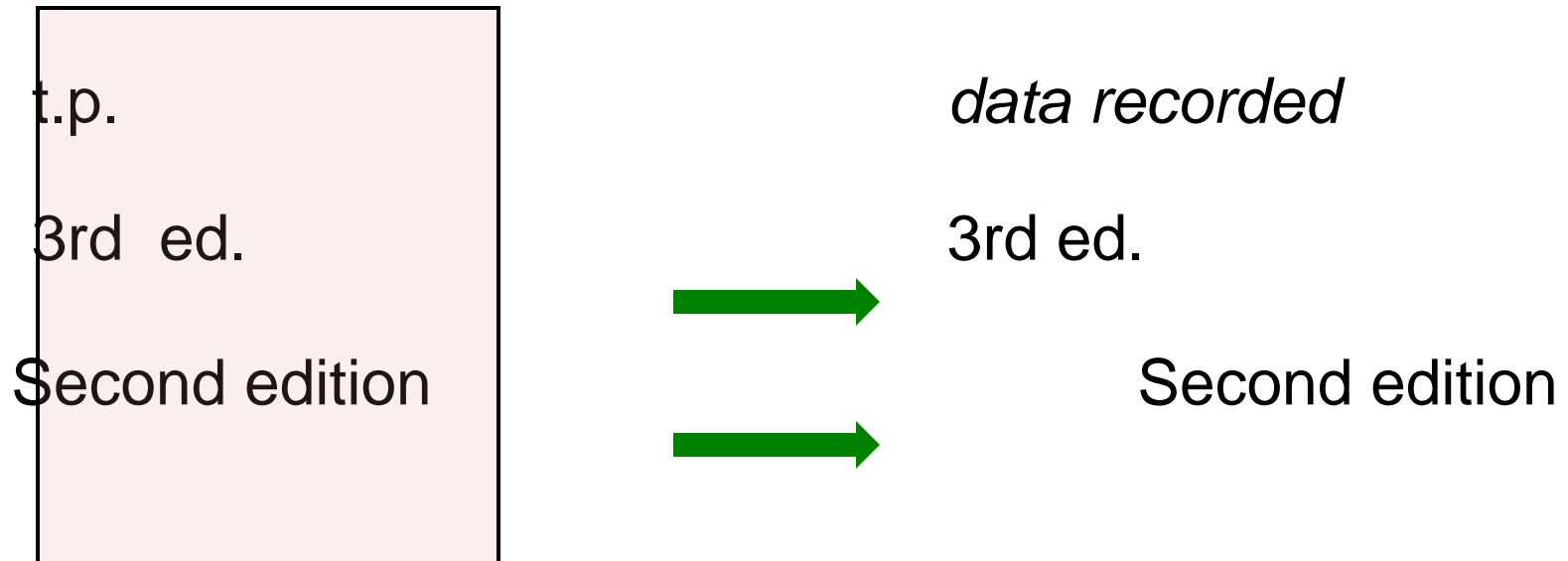   television.

Variant title: Title should
   read: The world of
   television

or

Title should read: The world
   of television.

Take what you see

- accurate representation
  do not introduce abbreviations

t.p.		*data recorded*
3rd ed.	→	3rd ed.
Second edition	→	Second edition

Accurate representation of content

- describe the manifestation
- identify the work(s) or expression(s)

for example: compilation of two or more works, no collective title

AACR2   • main entry = first work in the compilation

RDA   • do not "name" the resource based on the first work

• treat the works in the compilation in the same way

Accurate data

AACR2:   Author   Polk, Sharon.

            Uniform title   Community band concerts

            Title   Community band concerts / Sharon Polk. Fall harvest festivals / Terri Swanson.

            Contains: Swanson, Terri. Fall harvest festivals

RDA: Title   Community band concerts / Sharon Polk. Fall harvest festivals / Terri Swanson.

            Contains   Polk, Sharon. Community band concerts.

            Contains   Swanson, Terri. Fall harvest festivals

OR   Title   [Two United States Midwest memoirs]

            Contents   Community band concerts / Sharon Polk -- Fall harvest festivals / Terri Swanson.

            Contains   Polk, Sharon. Community band concerts.

            Contains   Swanson, Terri. Fall harvest festivals.

Accurate data: no more "rule of three"

- AACR2

Seeking the sacred / [in conversation with Thomas Moore … [et al.]].

access points for one

- RDA

Seeking the sacred / in conversation with Thomas Moore, Marion Woodman, Roméo D'Allaire, Stephen Lewis, Martin Rutte.

or

Seeking the sacred / in conversation with Thomas Moore … [and four others].

access points for one, some, or all

Relationship designators

- specify roles

  for example   *cartographer*

  performer

  broadcaster

  former owner

- specify the nature of the relationship

  for example   *adaptation of*

  musical setting of

  paraphrased as

Relationship designators

Alice in Wonderland, or, What's a nice kid like you doing in a place like this? / Hanna-Barbera Productions.

Parody of Carroll, Lewis, 1832-1898. Alice's adventures in Wonderland.

Contributor: Harvard University. Botanical Museum, issuing body

Issuing body: Harvard University. Botanical Museum
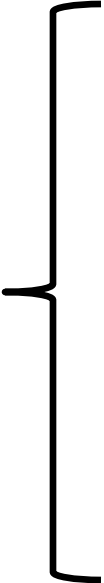
Contributor: Burger, Hillel, photographer

Photographer: Burger, Hillel

Ambiguity in AACR2

AACR2:
- information embedded in "other physical details" (or a non-specific note)
  - displays under the label "***Description***" with data about extent and dimensions
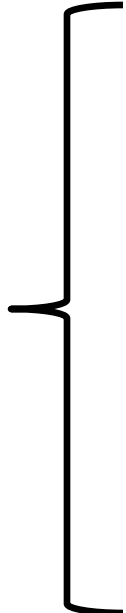
other physical
details

MARC 300 $b

- illustrative content
- encoding format
- production method
- sound content
- applied material
- base material
- reduction ratio

More precision in RDA

defined and distinct elements for each kind of data

different element for each type of data

illustrative content

encoding format

production method

sound content

applied material

base material

reduction ratio

More precision in RDA

controlled vocabulary for many elements

encoding format   DAISY, MP3, Access, XML, JPEG, TIFF, CAD, PDF, Blu-Ray, VCD

production method   blueline, blueprint, engraving, etching, lithograph, photocopy, woodcut

base material   bristol board, canvas, cardboard, ceramic, glass, leather, metal, paper

More precision in RDA

AACR2:   date of publication, distribution, etc.

MARC 21:   260 $c

RDA: 4 distinct elements:

date of production
date of publication
date of distribution
date of copyright

Dates of publication and copyright

- AACR2

Published: New York : Palgrave Macmillan, 2011.

Published: Cambridge, Mass. : Candlewick Press, c2006.

- RDA

Published: New York : Palgrave Macmillan, 2011, ©2011.

Published: Cambridge, Massachusetts : Candlewick Press, [2006], ©2006.

More precision in RDA

- AACR2

Matheus, Michael

Brunhoff, Jean de, 1899-1937. Babar en famille. English

Beard, Henry

- RDA

Matheus, Michael, 1953-

Brunhoff, Jean de, 1899-1937. Babar en famille. English. Spoken word

Miss Piggy, author

Beard, Henry, author

More precision in RDA

- precise access points
- access points for an automated environment

AACR2	Aesop's fables. Polyglot.
RDA	Aesop's fables. Greek
	Aesop's fables. Latin
	Aesop's fables. English
	Aesop's fables. German

AACR2	Aesop's fables. English & German
RDA	Aesop's fables. English
	Aesop's fables. German

AACR2

- data embedded in long character strings
- data in ambiguous elements

- assumption is that a human will read and interpret information in record

  ➢ cannot be used reliably to search or to limit a search

  ➢ cannot be used for automated processing

  ➢ cannot be used to generate a meaningful display
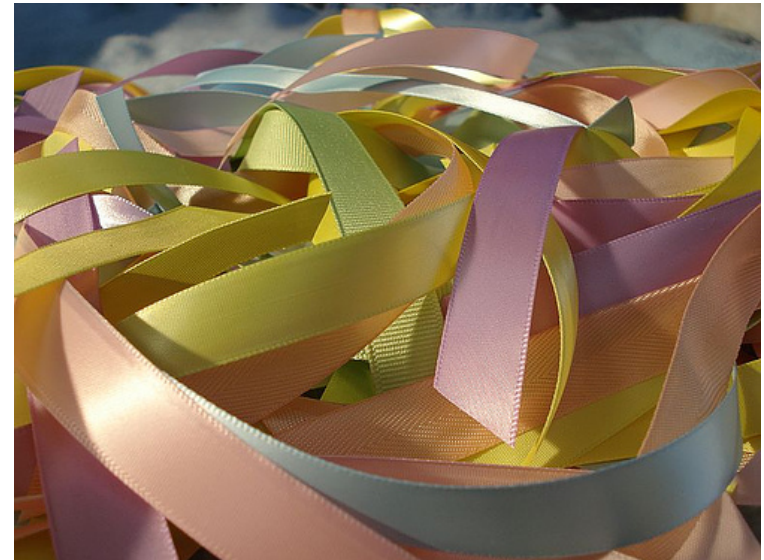
Precise data = Usable data

RDA

- each element is distinct and precisely defined
- each element contains only one kind of data
- controlled vocabulary in many elements

➢ each element has the potential to be usable:

　　to index
　　to search
　　to build meaningful displays of data

➢ data in any element can be used:　　by humans

　　　　　　　　　　　　　　　　　by computers

Usable data

identify an entity:

either

       eye-readable data   name

                                                date   …

       Shields, Carol, 1953-2003

and/or

       machine actionable data   identifier

           0101A6635

           http://viaf.org/viaf/4944537/#Shields,_Carol

URI = basic building block for linked data in the semantic web

Usable data

- RDA is a content standard

  not an encoding standard

  not a display standard

- take the data and create useful displays

RDA says: record person's date of death   *2003*

Display?   *died 2003*

               d. 2003

               - 2003

               date of death: 2003

User friendly: less abbreviations

- AACR2

Description: [37] p. : col. ill. ; 28 cm.

Description: 86, [21] p. : ill., 1 folded map : 24 cm.

- RDA

Description: 37 unnumbered pages : colour illustrations ; 28 cm

Description: 86 pages, 21 unnumbered pages : illustrations, 1 folded map ; 24 cm

User friendly: no Latin abbreviations

- AACR2

Title: … / by Susan Brown … [et al.].

Published: [S.l. : s.n.], 1852.

- RDA

Title: … / by Susan Brown … [and four others]

Published: [Place of publication not identified] : [publisher not identified], 1852

Day 1 of implementation

- day 1 of implementation will be in the current environment: MARC 21

  at least 12 different RDA data elements map to 300 $b

  at least 4 different types of dates map to 260 $c

- day 1 of implementation, our data will still be muddy:

  ambiguous

  imprecise

  less than desired for large scale machine processing

Phase 1: "shoe-horning" into MARC

Author	▶ Stanfield, James Ronald, 1945- , author.
Title	John Kenneth Galbraith / by James Ronald Stanfield and Jacqueline Bloom Stanfield.
Published	Houndmills, Basingstoke, Hampshire : Palgrave Macmillan, 2011, ©2011.
Description	xi, 251 pages ; 23 cm.
Series	▶ Great thinkers in economics series.
Location Info	HB119 G33 S733 2011 [Regular Loan] Humanities and Social Sciences - McLennan Bldg ⓘ
LC Subject	▶ Galbraith, John Kenneth, 1908-2006.
LC Subject	▶ Economists -- United States -- Biography.
	▶ Economics -- United States -- History.
Browse the Shelf	▶ HB119 G33 S733 2011
Bibliography Note	Includes bibliographical references and index.
Contributor	▶ Stanfield, Jacqueline Bloom, 1947- , author.
ISBN	▶ 9780230242685 (hardback)
	▶ 0230242685 (hardback)

Title	El escritor y el intelectual entre dos mundos : lugares y figuras del desplazamiento / Cécile Chantraine-Braillan, Norah Giraldi Dei Cas, Fatiha Idmhand (eds.).
Published	Madrid : Iberoamericana ; Frankfurt am Main : Vervuert, 2010, ©2010.
Description	916 pages ; 23 cm
Location Info	PQ8520.1 I5 Z65 2010 [Regular Loan] Humanities and Social Sciences - McLennan Bldg ⓘ
LC Subject	► Ainsa, Fernando, 1937- -- Criticism and interpretation.
LC Subject	► Spanish literature -- History and criticism.
	► Latin American literature -- History and criticism.
Browse the Shelf	► PQ8520.1 I5 Z65 2010
Note	Festschrift in honor of Fernando Ainsa.
Bibliography Note	Includes bibliographical references.
Contributor	► Ainsa, Fernando, 1937- honouree.
	► Chantraine-Braillan, Cécile, editor.
	► Giraldi dei Cas, Norah, 1947- editor.
	► Idmhand, Fatiha, editor.
ISBN	► 9788484895206
	► 8484895203

RDA: preparing for tomorrow

Visible data

- users expect that   all metadata is on the web

    library data needs to be visible on the web

- online catalog = closed database
    - invisible to web search engines "dark data"

- MARC 21 = library specific record format
    - web cannot access and use MARC data
    - not used in other cultural heritage communities

Visible data: out of the library silo

Making library data visible

- release library data from MARC 21 record structure

- make library data available on the web

  - use XML schema

  - support use in semantic web with declared name spaces and registered vocabularies

- library data that can be discovered by web search engines

Expanded scope

not just for libraries

- connecting with other cultural heritage communities

  e.g. additions and changes for archives

- instructions designed to describe a wide variety of resources

- possibility for other communities to adopt/adapt

  - release from MARC 21 record format also helps build connections to other communities

  - awareness of practices and standards in other metadata encoding communities

Expanded scope

beyond "Anglo-American"

- options for use of
  - other languages
  - other scripts
  - other calendars
  - other numeric systems

- interest of all countries using AACR2
- interest of countries that never used AACR

Where is the future?

- data that enables machine-actionable processing of data

- data that enables resource discovery on the web

- data that can be stored and used in new database structures

- data that can connect us with other communities

- data that responds to user needs

Poised to jump …
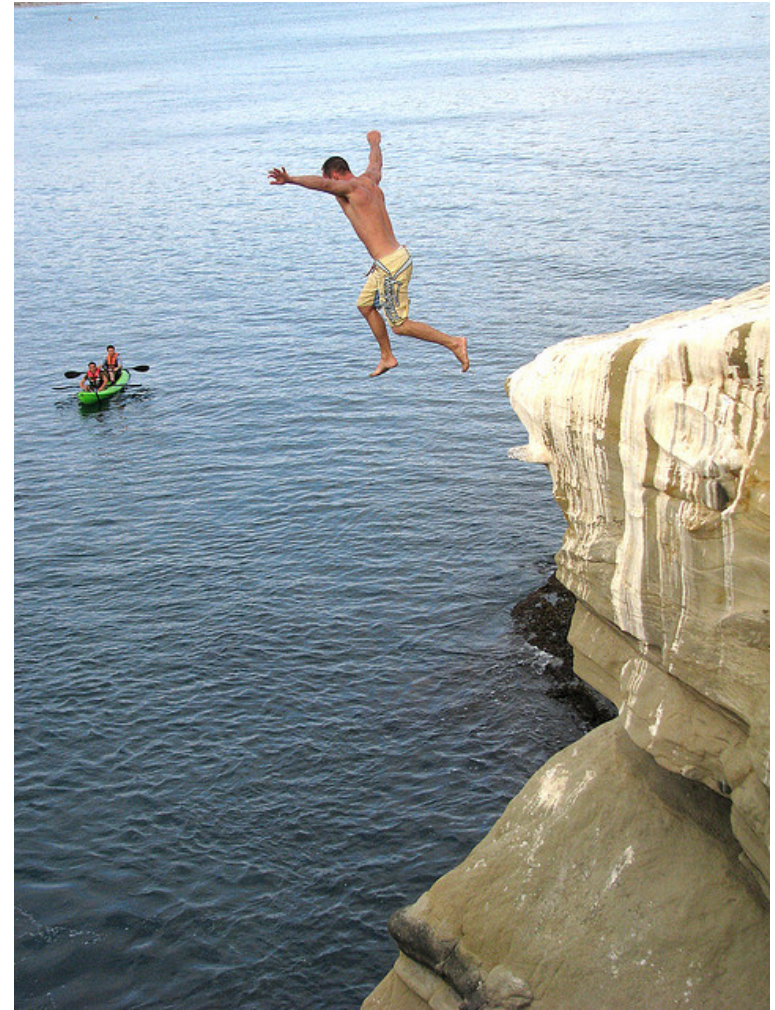
Day 1: small splash

Next step: the big wave …

Flickr credits: creative commons attribution

Indexing by kyz

http://www.flickr.com/photos/kyz/3962573662/

1978 Chevrolet El Camino by Collector Car Ads

http://www.flickr.com/photos/48718277@N06/4624577836/

Apple III introduction by Ballistik Coffee Boy

http://www.flickr.com/photos/spaceageboy/3132377142/

Fruit pyramid by Mario's Planet

http://www.flickr.com/photos/33511186@N00/58642708/

Fresh veggies by Brendahallowes

http://www.flickr.com/photos/11115086@N00/4293059669/

Flickr credits: creative commons attribution

Conversation by "clairity"

http://www.flickr.com/photos/clairity/154640125/

Dinner train by Slideshow Bruce

http://www.flickr.com/photos/springfieldhomer/95614622/

Cross track – iPhone wall paper by CJ Schmit

http://www.flickr.com/photos/cjschmit/4623783487/

Pastel ribbons by Sakurako Kitsa

http://www.flickr.com/photos/kitsa_sakurako/414007954/

Complete family bike by Sumanth Garakarajula

http://www.flickr.com/photos/photocracy/2861747022/

Flickr credits: creative commons attribution

Shard frame by Matt from London

http://www.flickr.com/photos/londonmatt/4221934301

Oregon silo by TooFarNorth

http://www.flickr.com/photos/toofarnorth/4597980984/

Cliff jumping/Diving prohibited by Better than Bacon

http://www.flickr.com/photos/slurm/1118367520/

Splash by Elsie esq.

http://www.flickr.com/photos/elsie/5025266033/

Splash! by Dplanet

http://www.flickr.com/photos/dplanet/141378340/

Questions …