Scrutinizing BIBFRAME from the perspective of RDA Metadata Schema and mapping to MARC 21 Format for Bibliographic Data

Tholakele Xulu
txulu@hsrc.ac.za

FOR IGBIS FORUM 2019 – LIASA ENTITY
Leriba Hotel, Centurion, Pretoria – South Africa
IN THIS SESSION

• BACKGROUND
  • MARC – History & Challenges
  • BIBFRAME - Development

• MODELS
  • BIBFRAME

• MAPPING MODELS
  • RDA
  • MARC
  • BIBFRAME

• MAPPING ENTIES - RELATIONSHIPS
  • RDA
  • BIBFRAME
GOALS

• Provide context for the genesis of a MARC replacement
• Provide an introduction to BIBFRAME
• Provide LC BIBFRAME updates
• Provide pointers for you to continue to follow this dynamic work
• How to get involved
<table>
<thead>
<tr>
<th>YEAR</th>
<th>DESCRIPTION</th>
</tr>
</thead>
<tbody>
<tr>
<td>1960</td>
<td>MARC – Library of Congress</td>
</tr>
<tr>
<td>1999 – Date</td>
<td>MARC 21 Formats – Made up of #5 Standards</td>
</tr>
</tbody>
</table>
MARC FORMATS

- Originally used as way for LC to disseminate and print cards more easily and quickly;
- MARC records largely look like electronic cards and function that way as well (more difficult for computer systems to process strings of data);
- Cataloguing community is closely tied to MARC tags and MARC was always mixed up with cataloguing descriptive rules,
MARC FORMATS, *Cont...*

- MARC is used exclusively by libraries which keep the library community isolated from the other communities.
- Moreover, every statement in MARC depends mainly on the entire record for context and meaning as none of the statement of MARC record including the fields and subfields can stand alone.
- Therefore it ended up as a carrier and descriptive schema.
MARC record models

- **Bibliographic record**
  - Describes all facets of the resource
  - Stand alone, can represent an item
  - Distinguished from all types of MARC record by specific codes in Leader

- **Holdings record**
  - Detail on physical location, holdings, and condition
  - Can be partially imbedded

- **Authority record**
  - Assist in maintaining consistency across bibliographic records
  - Provides reference structure that helps catalogue users
  - But not essential for exchange of bibliographic since bibliographic records can stand alone
MARC Successes

• Can carry data formulated by different cataloguing rules and conventions:-
  • Multiple descriptive rules, different principles and models
  • Different subject thesauri
  • Multiple languages and scripts

• Cooperation in record exchange has resulted in widespread use and costs savings

• Richness of MARC records supports multifaceted retrieval:-
  • Coded data
  • Parsed data - syntax analysis, or syntactic analysis - a process of analyzing a string of symbols, either in natural language, computer languages or data structures, conforming to the rules of a formal grammar
Why Change from MARC?

- Serves as a syntax defined by an international standards for communications with #2 expressions:
  - Classic MARC (MARC 2709)
  - MARCXML
- A data element set defined by content designation and semantics
- Many data elements are defined by external content rules; a common misperception is that it is tied to AACR2
- MARC does not specify internal storage and institutions do not store MARC 21
Problems with MARC

• MARC 2709 syntax problems:-
  • Limitation of available fields, subfields, indicator values, etc. – This prevents complete cross-walking
  • Redundant data (fixed vs variable fields)
  • Longevity of format complicates reuse of data tags; redundancies have built up overtime
  • Ability to link relationships is somewhat limited
  • Lack of explicit hierarchical levels
  • Failure to handle and link name headings that have multi-forms and multi-scripts
  • Many changes were considered to accommodate RDA
  • Lacks essential checks and balances to assure that appropriate granularity is achieved when coding a record.
Why Change a Standard? - *Changes in description requires*

- Incorporate linked data support
- Use of URIs where possible
- New library cataloguing rules
- Emphasis on relationships
- More media and electronic resources

Transcription pros and cons - *One of the serious problems inherited in MARC cataloging is the handling and linking of name headings.*

- that have multi-forms and multi-scripts.
- Non-traditional material
- E-Resource access management
- E-Resource object management
LINKING DATA ERA

It never ends - from Cataloguers to CATALINKS.
DEEMED TO BE A REPLACEMENT - BIBFRAME

• 2002, October – MARC must die (Roy Tennant).
• 2006, TBL’s - (Tim Berners-lee) Linked Data: Design issues.
• 2008, January – On the Record – (LC report) – suggested that MARC was no longer “fit for purpose”.
• 2011, June – Reports and recommendations of the U.S. RDA test Coordinating Committee. – Reported that, “most felt any benefits of RDA would be largely unrealized in a MARC environment”.
• 2011, October – A Bibliographic Framework for the Digital Age (LC Report). LC concludes from the two reports above to commit to a “new bibliographic framework”.
• 2012, May – LC begins work with Zepheira on model.
• 2014, Stabilization of model and experimentation. Still developed and argued over – reviewed time-to-time.
BIBFRAME - *is* ...

- A data model for resource description
- Built upon Linked data models interpreted through Resources Description Framework (RDF) modelling practice
- Not based on RDA or FRBR, but mostly compatible with them
- Still under development, but getting closer
- One hailstone in a metadata maelstrom
- Its current draft specification is version 2.0.
BIBFRAME – is intended to . . .

• Replace MARC
• Break apart the bibliographic record into atomistic metadata statements
• Play well with other metadata standards
• Make library resources more easily discoverable on the open web
BIBFRAME - Goals

• Supply search engines with description in a form they can exploit
• Use/exploit linking
  • traditional = textual, identifiers
  • semantic technology = URIs (URI – Uniform Resource Identifier)
• MARC transition
  • enable reuse of data from MARC
  • provision of transformations to new models
BIBFRAME GOALS – Cont ...

- Extensibility to new and broader content
- Accommodate needs for different types of libraries
- New views of different types of metadata
descriptive, authority, holdings
coded data, classification data, subject data
- preservation, rights, technical, archival
- Reconsideration of the data related activities exchange, internal storage, input interfaces and technique
Bibliographic Framework Transition Initiative

- Rethinking bibliographic control because of technological and environmental changes
- Content and packaging of RDA suggest that a different carrier is needed to fully exploit it
- Reevaluate use of scarce resources and provide efficiencies in creating and sharing bibliographic metadata
- Analyze present and future bibliographic control environments
- Identify components of the bibliographic framework to support users
- Plan for an evolution to a future framework.
BIBFRAME - Core levels

BIBFRAME Model consists of three core levels:-

- **Work.** The highest level of abstraction, a Work, in the BIBFRAME context, reflects the conceptual essence of the cataloged resource: authors, languages, and what it is about (subjects).

- **Instance.** A Work may have one or more individual, material embodiments, for example, a particular published form. These are Instances of the Work. An Instance reflects information such as its publisher, place and date of publication, and format.

- **Item.** An item is an actual copy (physical or electronic) of an Instance. It reflects information such as its location (physical or virtual), shelf mark, and barcode.
BIBFRAME HIGH LEVEL MODEL - Relationship to RDA and Mapping to MARC

- bf:Authority – Subjects (topics, etc.), Places, Person/Corporate Names, etc. (creators, contributors, editors, publishers, etc.)

- bf:Item - Summary
- bf:Item - Cover art
- bf:Item – Holding Information

Data in MARC fields
- RDA Work
- RDA Express -ion
- RDA Manifest -ation
- RDA Item
RESULTS – in short...

- All MARC Title and Name/title authority records converted to bf:Work entities
- RDA/frbr Work to RDA/frbr Expression relationships expressed in bf:Works
Additional Key Concepts

- **Agents**: Agents are people, organizations, jurisdictions, etc., associated with a Work or Instance through roles such as author, editor, artist, photographer, composer, illustrator, etc.

- **Subjects**: A Work might be “about” one or more concepts. Such a concept is said to be a “subject” of the Work. Concepts that may be subjects include topics, places, temporal expressions, events, works, instances, items, agents, etc.

- **Events**: Occurrences, the recording of which may be the content of a Work.

*These KEY concepts have relationships to the core classes*
IN ESSENCE

BIBFRAME vocabulary consists of RDF classes and properties.

• Classes include the three core classes listed above as well as various additional classes, many of which are subclasses of the core classes.

• Properties describe characteristics of the resource being described as well as relationships among resources.
  • For example: one Work might be a “translation of” another Work; an Instance may be an “instance of” a particular BIBFRAME Work.

• Other properties describe attributes of Works and Instances.
  • For example: the BIBFRAME property “subject” expresses an important attribute of a Work (what the Work is about), and the property “extent” (e.g. number) expresses an attribute of an Instance.
MANIFESTATION ACCORDING TO RDA TERMINOLOGY

Linkable entities

- E.g. author, composer
- E.g. festival, concert

Work
(Has instance)

- E.g. work place, time-span

Instance*
(Has item)

- E.g. volume, videodisc

Item

- E.g. library

Carrier type

Publisher

E.g. person, corporate body

Event

Held by

Local identifier

E.g. barcode

* Manifestation, according to RDA terminology
COMPATABILITY WITH VARIOUS MODELS - Modelling activity – BIBFRAME

- Functional Requirements for Bibliographic Records (FRBR)
- `<indecs>` metadata framework (INDECs)
- Online Audiovisual Catalogers model (OLAC)
- CIDOC Conceptual Reference Model (CIDOC – CRM)
- Describing Archives Content Standard (DACS)
- Resource Description Framework (RDF)
Next Steps for US?

**ILS Issues – backend and public displays:**

- Introductory sessions – SA context
- Realization of the 4th Industrial Revolution in our day-to-day operations
- Develop training materials depending on trusted sources
- Learn all about the new models then put it to the test – *Do not wait! – for knowledge base and comparison purposes*
- Within your institution or across a cooperative, Cooperate!
- Decide upon your library status: Are you ready to adopt *(or not adopt!)* But if you are using OCLC WorldShare you are bound to move with change!!!
Conclusion

• More information required, through series of presentation
• Much efforts will go into the new initiative
• There are widely differing views
• More questions than answers
  • How much of MARC will be retained?
  • Will the new format look like MODS, a derivative, or will it be completely new?
  • How will supporting data be accommodated?
  • How will systems change?
  • How long will it take?
Resources


• Linked data principles and resources - https://www.slideshare.net/vdeboer/linked-data-principles-and-examples


• 2015/2016 - https://www.loc.gov/bibframe/docs/index.html

• 2016 - https://www.loc.gov/bibframe/docs/bibframe2-model.html - BIBFRAME 2.0 Release

Thank you

- Follow the progress: [www.loc.gov/marc/transition](http://www.loc.gov/marc/transition)
- Join the discussion: [bibframe@listserv.loc.gov](mailto:bibframe@listserv.loc.gov)