Control or Chaos:

Embracing Change and Harnessing Innovation in an Ecosystem of Shared Bibliographic Data

Ian Bigelow & Abigail Sparling - University of Alberta Library
Bibliographic Control in the Digital Ecosystem, Firenze, Italy, 8 February 2021
Introduction: What will we cover

In an effort to capture some of the challenges for bibliographic control emerging in the changing landscape for library bibliographic metadata we will focus on several key areas of discussion as they relate to data reuse:

1. RDA and BIBFRAME
2. What surprises do we find in MARC?
3. What should BIBFRAME and BIBFRAME infrastructure look like?
4. Discussion: Harnessing innovation and maintaining control
Context: What lense are we using for this examination

In 2018 strategic planning at UAL resulted in a plan for *Moving Forward with Linked Data at UAL*

“In order to reap the benefits of full participation in the linked open data environment, UAL should continue to take steps towards complete conversion of existing library data to linked open data.”¹

Since 2018 UAL has continued as a member of Share-VDE and was part of the LD4P2 Cohort. We are also a member of the Program for Cooperative Cataloging.

Much of this session is informed by experiences and observations from these projects and our points of reference, particularly as we plan for next steps and strategic priorities for the next 5 years.

1. [Moving Forward with Linked Data at UAL](#)
Context: Bringing it all together

- **Sinopia Connections (3 year goal)**
- **Entity Management in J. Cricket**
- **Sapientia + J Cricket**
- **Bf-bf harmonization**
- **Test LD Discovery; Access to Data**
- **Sirsi**
- **Original and copy cataloguing in Sinopia**
- **Load Bulk MARC files to Sirsi (will be converted to bf)**

The diagram illustrates the workflow from **Sinopia Connections** through **Sapientia + J Cricket**, **Bf-bf harmonization**, **Test LD Discovery; Access to Data**, and finally to **Sirsi**, with **bf2mrc** and **mrc2bf** conversions along the way.
<table>
<thead>
<tr>
<th>Resource Description and Access</th>
<th>BIBFRAME 2010</th>
</tr>
</thead>
<tbody>
<tr>
<td>Chaos and Convergence</td>
<td>BIBFRAME 2011</td>
</tr>
</tbody>
</table>
Approximately a decade of history leading to a few key points:

- From the very outset of RDA in 2010 there was agreement that an alternative to MARC was required to support the extent of RDA.
- BIBFRAME has also had a long development trajectory, with initial work starting in 2011, with the goal of creating a community standard to allow RDA to move beyond MARC.
- In 2010 a test was carried out for the initial version of RDA, but at the time this was focused on application in MARC only, so with the development of BIBFRAME we are really just getting to a point where many components can come together.
- 2017-2020 has seen increased development not just in BIBFRAME, but in the evaluation, testing and analysis of use of RDA in a linked data environment. This acceleration has resulted in beautiful chaos, with further work on data modelling, more maturity in conversion processes, and use case development driving novel extensions and adaptations for the use of RDA in a BIBFRAME environment.
- **Now** is the time to move forward together
Analysing Native BIBFRAME and the use of RDA
<table>
<thead>
<tr>
<th>Resource Template Label</th>
<th>Type</th>
<th>Mandatory</th>
<th>Repeatable</th>
<th>PropertyURI</th>
<th>PropertyLabel</th>
<th>RDA Instruction/Entry Note</th>
</tr>
</thead>
<tbody>
<tr>
<td>UAL Monograph Work (Un-Nested)</td>
<td>resource</td>
<td>false</td>
<td>true</td>
<td><a href="http://id.loc.gov/ontologies/bibframe/expressionOf">http://id.loc.gov/ontologies/bibframe/expressionOf</a></td>
<td>Has Opus</td>
<td></td>
</tr>
<tr>
<td></td>
<td>lookup</td>
<td>false</td>
<td>true</td>
<td><a href="http://id.loc.gov/ontologies/bibframe/hasInstance">http://id.loc.gov/ontologies/bibframe/hasInstance</a></td>
<td>Has Instance</td>
<td>Used with Unspecified</td>
</tr>
<tr>
<td></td>
<td>resource</td>
<td>false</td>
<td>true</td>
<td><a href="http://id.loc.gov/ontologies/bibframe/identifiedBy">http://id.loc.gov/ontologies/bibframe/identifiedBy</a></td>
<td>Work Identifier</td>
<td></td>
</tr>
<tr>
<td></td>
<td>lookup</td>
<td>false</td>
<td>true</td>
<td><a href="http://id.loc.gov/ontologies/bibframe/contribution">http://id.loc.gov/ontologies/bibframe/contribution</a></td>
<td>Contribution</td>
<td></td>
</tr>
<tr>
<td></td>
<td>resource</td>
<td>false</td>
<td>true</td>
<td><a href="http://id.loc.gov/ontologies/bibframe/title">http://id.loc.gov/ontologies/bibframe/title</a></td>
<td>Title Information</td>
<td></td>
</tr>
<tr>
<td></td>
<td>lookup</td>
<td>false</td>
<td>true</td>
<td><a href="http://id.loc.gov/ontologies/bibframe/genreForm">http://id.loc.gov/ontologies/bibframe/genreForm</a></td>
<td>Form of Work</td>
<td></td>
</tr>
<tr>
<td></td>
<td>literal</td>
<td>false</td>
<td>true</td>
<td><a href="http://id.loc.gov/ontologies/bibframe/originDate">http://id.loc.gov/ontologies/bibframe/originDate</a></td>
<td>Date of Work</td>
<td><a href="http://access.rdatoolkit.org/6.4.html">http://access.rdatoolkit.org/6.4.html</a></td>
</tr>
<tr>
<td></td>
<td>lookup</td>
<td>false</td>
<td>true</td>
<td><a href="http://id.loc.gov/ontologies/bibframe/originPlace">http://id.loc.gov/ontologies/bibframe/originPlace</a></td>
<td>Place of Origin of the Work</td>
<td><a href="http://access.rdatoolkit.org/6.5.html">http://access.rdatoolkit.org/6.5.html</a></td>
</tr>
<tr>
<td></td>
<td>literal</td>
<td>false</td>
<td>true</td>
<td><a href="http://id.loc.gov/ontologies/bibframe/geographicCoverage">http://id.loc.gov/ontologies/bibframe/geographicCoverage</a></td>
<td>(Geographic) Coverage of the Content</td>
<td><a href="http://access.rdatoolkit.org/7.3.html">http://access.rdatoolkit.org/7.3.html</a></td>
</tr>
<tr>
<td></td>
<td>lookup</td>
<td>false</td>
<td>true</td>
<td><a href="http://id.loc.gov/ontologies/bibframe/coverage">http://id.loc.gov/ontologies/bibframe/coverage</a></td>
<td>(Time) Coverage of the Content</td>
<td></td>
</tr>
<tr>
<td></td>
<td>resource</td>
<td>false</td>
<td>true</td>
<td><a href="http://id.loc.gov/ontologies/bibframe/intendedAudience">http://id.loc.gov/ontologies/bibframe/intendedAudience</a></td>
<td>Intended Audience</td>
<td>access.rdatoolkit.org/7.7.html</td>
</tr>
<tr>
<td></td>
<td>lookup</td>
<td>false</td>
<td>true</td>
<td><a href="http://id.loc.gov/ontologies/bibframe/intendedAudience">http://id.loc.gov/ontologies/bibframe/intendedAudience</a></td>
<td>Intended Audience</td>
<td></td>
</tr>
<tr>
<td></td>
<td>resource</td>
<td>false</td>
<td>true</td>
<td><a href="http://id.loc.gov/ontologies/bibframe/hasSeries">http://id.loc.gov/ontologies/bibframe/hasSeries</a></td>
<td>In Series</td>
<td><a href="http://access.rdatoolkit.org/7.8.html">http://access.rdatoolkit.org/7.8.html</a></td>
</tr>
<tr>
<td></td>
<td>resource</td>
<td>false</td>
<td>true</td>
<td><a href="http://id.loc.gov/ontologies/bibframe/tableOfContents">http://id.loc.gov/ontologies/bibframe/tableOfContents</a></td>
<td>Contents</td>
<td></td>
</tr>
<tr>
<td></td>
<td>resource</td>
<td>false</td>
<td>true</td>
<td><a href="http://id.loc.gov/ontologies/bibframe/summary">http://id.loc.gov/ontologies/bibframe/summary</a></td>
<td>Summary</td>
<td><a href="http://access.rdatoolkit.org/7.10.html">http://access.rdatoolkit.org/7.10.html</a></td>
</tr>
<tr>
<td></td>
<td>lookup</td>
<td>false</td>
<td>true</td>
<td><a href="http://id.loc.gov/ontologies/bibframe/subject">http://id.loc.gov/ontologies/bibframe/subject</a></td>
<td>Subject of the Work</td>
<td></td>
</tr>
<tr>
<td></td>
<td>lookup</td>
<td>true</td>
<td>true</td>
<td><a href="http://id.loc.gov/ontologies/bibframe/classification">http://id.loc.gov/ontologies/bibframe/classification</a></td>
<td>Classification numbers</td>
<td><a href="http://access.rdatoolkit.org/6.9.html">http://access.rdatoolkit.org/6.9.html</a></td>
</tr>
<tr>
<td></td>
<td>resource</td>
<td>false</td>
<td>true</td>
<td><a href="http://id.loc.gov/ontologies/bibframe/content">http://id.loc.gov/ontologies/bibframe/content</a></td>
<td>Language of Expression</td>
<td><a href="http://access.rdatoolkit.org/6.11.html">http://access.rdatoolkit.org/6.11.html</a></td>
</tr>
<tr>
<td></td>
<td>lookup</td>
<td>false</td>
<td>true</td>
<td><a href="http://id.loc.gov/ontologies/bibframe/illustrativeContent">http://id.loc.gov/ontologies/bibframe/illustrativeContent</a></td>
<td>Illustrative Content</td>
<td><a href="http://access.rdatoolkit.org/7.13.html">http://access.rdatoolkit.org/7.13.html</a></td>
</tr>
<tr>
<td></td>
<td>resource</td>
<td>false</td>
<td>true</td>
<td><a href="http://id.loc.gov/ontologies/bibframe/supplementaryContent">http://id.loc.gov/ontologies/bibframe/supplementaryContent</a></td>
<td>Supplementary Content</td>
<td><a href="http://access.rdatoolkit.org/7.16.html">http://access.rdatoolkit.org/7.16.html</a></td>
</tr>
<tr>
<td></td>
<td>resource</td>
<td>false</td>
<td>true</td>
<td><a href="http://id.loc.gov/ontologies/bibframe/hasExpression">http://id.loc.gov/ontologies/bibframe/hasExpression</a></td>
<td>Related Expressions</td>
<td><a href="http://access.rdatoolkit.org/rda26-55.html">http://access.rdatoolkit.org/rda26-55.html</a></td>
</tr>
</tbody>
</table>
Convergence: The Opus

LRM

- Work
- Expression
- Manifestation
- Item

Is Realized Through

Is Embodied In

Is Exemplified By

Bibframe

- Work
- Has Expression
- Has Instance
- Has Item

Share VDE - Option 4

- svde:Opus
- bf:Work

- svde:Work
- bf:Work

- Instance
- Has Item
- Item

Ford, Kevin. [Share-VDE - Option 4]. Created for the Share-VDE Sapientia Entity Identification Working Group, 2020
Conformance and Questions

1. Based on RDA 3R, what is RDA?
   a. Does it require RDA/RDF and if so is it still a content standard
2. If a community chooses to use unconstrained elements is it conformant with RDA?
3. Importantly, different communities will take different approaches either way:

   “It is important to remember that RDA and RDA/RDF are two different things. RDA instructions will always be more applicable to traditional library resources than to newly emerging material types. We might also consider that given one of our goals for linked data is to communicate and consume data from beyond libraries, our RDF serialization might need to be more approachable than the complexity of RDA/RDF. As such and because we will probably be in a long-term transition away from MARC, PCC will continue to treat RDA as a loose content standard and participate in RDA/RDF and BIBFRAME discussions to assess our ideal linked data output.”

Whither RDA/RDF or BIBFRAME

- While RDA is certainly comparable to BIBFRAME, there are notable differences, for example with some elements having one to many or many to one relationship.
- Nevertheless, as demonstrated by work on Sinopia application profiles, core element sets can clearly be mapped and utilized from one to the other, and this should also support mappings for interchange, or indeed the use of both in a shared data set.
- Minimally if multiple implementation scenarios emerge we need to map out core elements for the purposes of interchange.
- Practically speaking though, according to RDA conformance documentation and pure RDA/RDF, there are few tools for working with RDA in production environments:
  - Does it make sense to duplicate effort for conversion processes, editors, and assessment of resulting data when we can could instead work on this all at once?
  - Recreating similar tools solely for RDA/RDF would take considerable time and likely mean that the new RDA will not be implemented for many years to come.
What surprises do we find in MARC?

1. Changes across time:
   a. The evolution of standards and practices (AACR2-RDA-RDA3R for instance)
      i. Ex. 245$h[GMD] vs MARC 336/7/8

2. Changes across space
   a. Communities of practice and approach to use of elements of description
      i. Ex. 700$a$t vs usage in 76X-78X

Work in the Share-VDE community with a wide range of institutions in North America and Europe has highlighted the need to account for such differences for conversion specifications. A prime example included testing of 1985 and 2015 data separately through conversion to account for changing standards.

To handle such differences requires decisions, specific solutions, and sometimes compromises. If each community makes such decisions separately, what will it mean for bibliographic control?
What Should BIBFRAME and BIBFRAME Infrastructure Look Like?

To support the move from MARC to BIBFRAME we need:

- BIBFRAME data output from each MARC to BIBFRAME conversion process to be interoperable with BIBFRAME data created natively in RDF
- BIBFRAME data created natively in RDF by any one community to be interoperable with other BIBFRAME stores
- BIBFRAME data in various flavors to be converted to MARC with similar consistency to meet MARC use case requirements
- New tools and processes need to support various serializations of BIBFRAME (RDF XML, n-triples, n-quads, turtle, JSON-LD), or a community decision on which to use for development
<table>
<thead>
<tr>
<th>Label</th>
<th>ID</th>
<th>Class</th>
<th>Context</th>
</tr>
</thead>
</table>
Unable to load the entire resource. The unused triples are:

Format: N-Triples

<http://share-vde.org/sharevde/rdf8libframe/Agent/270150> <http://id.loc.gov/ontologies/bfic/name0MatcherKey> "Lucius Annaeus Seneca." .
<http://share-vde.org/sharevde/rdf8libframe/Agent/270150> <http://id.loc.gov/ontologies/bfic/name0MatcherKey> "Aurelius, Marcus." .
<http://share-vde.org/sharevde/rdf8libframe/Agent/270150> <http://id.loc.gov/ontologies/bfic/name0MatcherKey> "Aurelius, Marcus." .
<http://share-vde.org/sharevde/rdf8libframe/Title_506/6dfcfe3d-b5ad-d4bf-b1aa-e4aa825b7f90> <http://www.w3.org/2000/01/rdf-schema#label> "Meditations" .
BIBFRAME “Shape” and Serializations

- Differences in BIBFRAME modeling across projects and communities results in data and tooling incompatibilities
  - Ex. LC’s BIBFRAME to MARC converter XSLT released in May 2020 non workable for converting BIBFRAME data created using Sinopia application profiles
  - Ex. Differences between Sinopia and LC modeling or Sinopia and Share-VDE as highlighted in the Task Group on PCC Sinopia Application Profiles Final Report

- BIBFRAME serialization choices impact tool development and usability
  - Ex. LC’s BIBFRAME to MARC converter XSLT only works with BIBFRAME data serialized in RDF/XML
Harnessing Innovation

“[The BIBFRAME] model, like MARC, must be able to accommodate any number of content models and specific implementations, but still enable data exchange between libraries. It needs to support new metadata rules and content standards that emerge, including the newest library content standard - RDA (Resource Description & Access). The BIBFRAME model must therefore both broaden and narrow the format universe for exchange of bibliographic data.”

Harnessing Innovation

To date, experimentation and innovation by libraries and library partners around BIBFRAME has resulted in:

- BIBFRAME ontology updates
- Unique community extensions
- Format specific application profiles
- Mappings between emergent and project-specific library linked data models
- Systems to support the conversion, creation, management, and discovery of BIBFRAME data
Maintaining Control

- Define core BIBFRAME elements necessary for resource description
- Define standard BIBFRAME model and “shape” to support conversion and data re-use
- Define MARC use cases in a BIBFRAME environment
- Define implementation scenarios for the use of RDA 3R in BIBFRAME
Maintaining Control

- Develop a community of practice and ensure open feedback loops between LC, large scale projects and BIBFRAME implementers, and the wider linked data community
- Prioritize transparency around ongoing and future developments to the BIBFRAME ontology and technical infrastructure (along with supporting analyses and user testing data)
- Develop and coordinate implementation timelines for RDA and BIBFRAME
  - This does not need to reflect when everyone moves to linked data, but when both standards are supported for application and exchange.
Thank You!

Questions?

Ian Bigelow - bigelow@ualberta.ca
Abigail Sparling - ajsparli@ualberta.ca