Library Data and Its Future on the Web

Sandy Rodriguez
University of Missouri–Kansas City

Felicity Dykas
University of Missouri–Columbia

Mary Aycock
Missouri University of Science & Technology

MOBIUS Annual Conference 2015
Outline

• Linked Data and Linked Open Data
• BIBFRAME
• Linked Data in Libraries

Sir Timothy John "Tim" Berners-Lee, OM, KBE, FRS, FREng, FRSA, DFCB, also known as TimBL, is an English computer scientist, best known as the inventor of the World Wide Web. He was born on June 8, 1955, in London, United Kingdom.

In 1989, Tim Berners-Lee invented the World Wide Web, an Internet-based hypermedia initiative for global informative sharing while at CERN, the European Organization for Nuclear Research in Geneva, Switzerland.

Books: Weaving the Web: The Original Design and Ultimate Destiny of the World Wide Web by its Inventor

Awards: Marconi Prize, MacArthur Fellowship, Charles Stark Draper Prize, Mountbatten Medal, President's Medal

Profiles:
Linked Open Data: What Is It?
Don’t forget the bots
LINKED DATA

On the web, open license
Machine-readable data
Non-proprietary format
RDF standards
Linked RDF

IS YOUR DATA 5★?
= Resource Description Framework

Standard model for publishing and linking data on the web

Designed for:
RDF Triples
Alice is a friend of BOB, who is interested in The Mona Lisa, which was created by Leonardo Da Vinci. BOB was born on 14 July 1990 and the Mona Lisa is about La Joconde à Washington.

http://www.w3.org/TR/2014/NOTE-rdf11-primer-20140225/
Principles of Linked Data

1. Use URIs to name (identify) things.
2. Use HTTP URIs so that these things can be referred to and looked up ("dereferenced") by people and systems.
3. Provide useful information when a URI is dereferenced, using open standards such as RDF.
4. Include links to other related things, using their URIs, when publishing data to the Web.

http://www.w3.org/DesignIssues/LinkedData.html
RDF is a model notation

Common Serializations:
- N-triples
- Terse RDF Triple Language
- RDF/XML
- JSON-LD
- RDFa

N-triples

```xml
```

Turtle

```turtle
@prefix dc: <http://purl.org/dc/terms/> .

```

http://rdf-translator.appspot.com/
Linked Data: Issues

Data quality
Service failures - Caching
Inefficiency of distributed queries
Schema term definitions

Linked Data Tools
<table>
<thead>
<tr>
<th>#</th>
<th>Name</th>
<th>GivenName</th>
<th>Dates</th>
</tr>
</thead>
<tbody>
<tr>
<td>343</td>
<td>Bell, Quentin, 1910-1996</td>
<td>Quentin Claudian Stephen</td>
<td>1910-1996</td>
</tr>
<tr>
<td>349</td>
<td>Bell, William Blair-</td>
<td>William Blair-</td>
<td>1871-1936</td>
</tr>
</tbody>
</table>

For more information, visit [http://blog.archiveshub.ac.uk/2013/08/16/hub-viaf-namematching/](http://blog.archiveshub.ac.uk/2013/08/16/hub-viaf-namematching/)
### Linked Jazz - Developing Tools

<table>
<thead>
<tr>
<th>Name</th>
<th>Birth - Death</th>
</tr>
</thead>
<tbody>
<tr>
<td>Martin &quot;Chink&quot; Abrahams</td>
<td>1886 - 1981</td>
</tr>
<tr>
<td>Nathan Abshire</td>
<td>1913 - 1981</td>
</tr>
<tr>
<td>Buddy Ace</td>
<td>1936 - 1994</td>
</tr>
<tr>
<td>Pepper Adams</td>
<td>1930 - 1986</td>
</tr>
<tr>
<td>Nat Adderley</td>
<td>1931 - 2000</td>
</tr>
<tr>
<td>Joe Albany</td>
<td>1924 - 1988</td>
</tr>
<tr>
<td>Alvin Alcorn</td>
<td>1912 - 2003</td>
</tr>
<tr>
<td>Alger Texas Alexander</td>
<td>1900 - 1954</td>
</tr>
<tr>
<td>Roland Alexander</td>
<td>1935 - 2006</td>
</tr>
<tr>
<td>Rashied Ali</td>
<td>1933 - 2009</td>
</tr>
</tbody>
</table>

**Pepper Adams**

Pepper Frederick "Pepper" Adams III (October 8, 1930 – September 10, 1986) was a jazz baritone saxophonist and composer. He composed 43 pieces, was the leader on eighteen albums spanning 28 years, and participated in 600 sessions as a sideman.

[http://id.loc.gov/authorities/names/n79115718](http://id.loc.gov/authorities/names/n79115718)
Be
Bold
Is there a place for libraries in a linked data environment?
Bibliographic Framework Initiative

- Initiated by the Library of Congress
- Development of a bibliographic framework, called BIBFRAME
  - provide a foundation for the future of bibliographic description, both on the web, and in the broader networked world.
  - replacement for MARC
  - serve as a general model for expressing and connecting bibliographic data
- A major focus of the initiative will be to determine a transition path for the MARC 21 formats while preserving a robust data exchange that has supported resource sharing and cataloging cost savings in recent decades.

http://www.loc.gov/bibframe/
BIBFRAME

Bibliographic Framework Initiative

Initiated by the Library of Congress, BIBFRAME provides a foundation for the future of bibliographic description, both on the web, and in the broader networked world. This site presents general information about the project, including presentations, FAQs, and links to working documents. In addition to being a replacement for MARC, BIBFRAME serves as a general model for expressing and connecting bibliographic data. A major focus of the initiative will be to determine a transition path for the MARC 21 formats while preserving a robust data exchange that has supported resource sharing and cataloging cost savings in recent decades.
• Zepheira is a privately held company driven by a team of knowledge management and specific domain area experts ...
• Zepheira’s expert team has led the development and deployment of the Semantic Web ...
• ... has provided much of the vision and technical foundation for the next generation of the Web.
• ... deeply experienced in successfully implementing these technologies ...

https://zepheira.com/about/
Result of Zepheira’s Work

• “Bibliographic Framework as a Web of Data: Linked Data Model and Supporting Services”
  November 21, 2012
• 42 pages
• New, proposed model called BIBFRAME
• 18 pages about the model, including illustrations and examples
• 5 pages about linked data
• 9 pages about related library initiatives
• Conclusion
• Glossary
#1. BIBFRAME is not a ready-to-go, off-the-shelf standard that you can pick up and start using

Library of Congress FAQ
• “When should we move to BIBFRAME?”

• BIBFRAME is far from an environment that you could move to yet. The model and its components are still in discussion and development -- a work in progress. When it is more mature, vendors and suppliers will need time to adjust services to accommodate it. And then we can expect a mixed environment for some time.”

http://www.loc.gov/bibframe/faqs/#q08
“So, Can I Ignore it for now?”
No. Stay alert!

Although the BIBFRAME model is a draft and expected to change, the Library of Congress wants to share it now with the community not only so that it is informed of progress being made but also to engender conversation and constructive feedback. The Library is leading this initiative, but it is important that the library community as a whole work in concert to create an environment for bibliographic description and data exchange that recognizes and leverages the resources and scale of a global network of data.

Bibliographic Framework as a Web of Data, 2012, Introduction
BIBFRAME Includes ...

- BIBFRAME model
- BIBFRAME vocabularies
  - Classes
  - Properties
- BIBFRAME profiles
<table>
<thead>
<tr>
<th>Work (Creative Work)</th>
<th>Instance</th>
<th>Authority</th>
<th>Annotation</th>
</tr>
</thead>
<tbody>
<tr>
<td>Covers FRBR Work and Expression</td>
<td>FRBR Manifestation and Item</td>
<td>A thing or concept associated with a BIBFRAME Work or Instance</td>
<td>Used to give more information about a BIBFRAME Work, Instance, or Authority</td>
</tr>
<tr>
<td>The conceptual essence of a resource</td>
<td>An individual, material embodiment of a Work</td>
<td>Provides access points</td>
<td>Examples: Library holdings, Reviews, Cover Images, Tables of Contents, Descriptions</td>
</tr>
<tr>
<td>Serves as a Web based control point and reference point</td>
<td>Can be physical or digital</td>
<td>Includes: Agents (people, organizations, and jurisdictions), Places, Subjects</td>
<td></td>
</tr>
<tr>
<td>BIBFRAME Properties (Selected)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>--------------------------------</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Work (Creative Work)</strong></td>
<td><strong>Instance</strong></td>
<td><strong>Authority</strong></td>
<td><strong>Annotation</strong></td>
</tr>
<tr>
<td>absorbedBy</td>
<td>accompaniedBy</td>
<td>Authorized AccessPoint</td>
<td>annotates annotation AssertedBy</td>
</tr>
<tr>
<td>classificationLcc</td>
<td>colorContent</td>
<td>identifier</td>
<td></td>
</tr>
<tr>
<td>contributor</td>
<td>contentNotes</td>
<td>label</td>
<td></td>
</tr>
<tr>
<td>creator</td>
<td>duration</td>
<td>relatedTo</td>
<td></td>
</tr>
<tr>
<td>genre</td>
<td>edition</td>
<td></td>
<td></td>
</tr>
<tr>
<td>language</td>
<td>extent</td>
<td></td>
<td></td>
</tr>
<tr>
<td>series</td>
<td>formatOfMusic</td>
<td></td>
<td></td>
</tr>
<tr>
<td>subject</td>
<td>frequency</td>
<td></td>
<td></td>
</tr>
<tr>
<td>title</td>
<td>isbn10</td>
<td></td>
<td></td>
</tr>
<tr>
<td>workTitle</td>
<td>title</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
Instances as defined by this XML serialization is shown in Figure 5.
Focus on Data Elements

• Like FRBR and RDA, BIBFRAME focuses on data elements
• Not records
MARC fields with multiple data elements

264 1 $a Chicago : $b ALA Editions, an imprint of the American Library Association, $c 2015.

Date elements

Date of publication = 2015
Publisher = American Library Association
Place of publication = Chicago
http://www.niso.org/news/events/2013/dcmi/bibframework
BIBFRAME Vocabulary

Defines a set of classes and properties

- Class: main categorization
- Properties
  - belong to classes
  - used to describe BIBFRAME resources
Argh! Will I have to learn rdf, triples, and BIBFRAME classes and properties to make this work?

BIBFRAME Profiles
BIBFRAME tools

• [http://bibframe.org/tools/](http://bibframe.org/tools/)
Tools

BIBFRAME Editor (BFE)

For input of data directly into BIBFRAME, try out the demo or download and experiment.

1. BIBFRAME Editor - Demonstration
   This site is for demonstration purposes. It enables recording of a description using the BIBFRAME vocabulary but at this time does not store the input.

2. BIBFRAME Editor - Download
   This link is to the freely downloadable software for the Editor via Github. Please read the documentation for the Editor before attempting to experiment with the software.

MARC to BIBFRAME transformation tools

There are two tools designed to help you evaluate MARC Bibliographic data in the BIFRAME model. The transformation software that powers these tools is also available for download.

1. Comparison service
   Enter the bibliographic identifier (MARC field 001) of a Library of Congress MARC record to view a before and after presentation of a MARC record from the Library of Congress's database as BIBFRAME resources.

2. Transformation service
   Submit your own MARC Bibliographic records (as MARC/XML) and view them as BIBFRAME resources in Exhibit. The resulting data are also available for download.

NOTICE: Most of the class and property names in the data should match the names of classes and properties in the vocabulary. There may be some misalignment because of a short lag between the time the vocabulary is modified and the transformation code updated.
BIBFRAME Editor

Get the code here. In the meantime, select a Profile below to get started.

<table>
<thead>
<tr>
<th>BIBFRAME - Kitchen Sink Profiles</th>
</tr>
</thead>
<tbody>
<tr>
<td>Work</td>
</tr>
<tr>
<td>Authorized access point</td>
</tr>
<tr>
<td>Simple label</td>
</tr>
<tr>
<td>Work title</td>
</tr>
<tr>
<td>Any title</td>
</tr>
<tr>
<td>Title variation</td>
</tr>
<tr>
<td>Associated agent</td>
</tr>
<tr>
<td>Creator role</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>BIBFRAME - Simple Monograph</th>
</tr>
</thead>
<tbody>
<tr>
<td>New Holding</td>
</tr>
<tr>
<td>New Instance</td>
</tr>
<tr>
<td>New Work</td>
</tr>
</tbody>
</table>

New monograph (New Work, Instance, and Holding)
<table>
<thead>
<tr>
<th>Lookup</th>
<th>einstein, albert</th>
</tr>
</thead>
<tbody>
<tr>
<td>LCNAF</td>
<td></td>
</tr>
<tr>
<td>Einstein, H. A. (Hans Albert), 1904-1973</td>
<td></td>
</tr>
<tr>
<td>Einstein, Albert, 1879-1955 (Spirit)</td>
<td></td>
</tr>
<tr>
<td>Einstein, Albert, 1879-1955</td>
<td></td>
</tr>
<tr>
<td>Brooks, Albert, 1947-</td>
<td></td>
</tr>
<tr>
<td>Authorized access point</td>
<td></td>
</tr>
<tr>
<td>Einstein, Arik</td>
<td></td>
</tr>
<tr>
<td>Authority assigner</td>
<td></td>
</tr>
<tr>
<td>Jurisdiction</td>
<td></td>
</tr>
<tr>
<td>Authority source</td>
<td><a href="http://fd.loc.gov">http://fd.loc.gov</a></td>
</tr>
</tbody>
</table>
MARC to BIBFRAME Transformation Service

This service transforms a file of MARCXML records to BIBFRAME representation. Please note: files must be smaller than 2MB.

See also: Comparison Service

External URL File Transformation  Paste-in MARCXML Transformation

URL for MARCXML File

C:\Users\dykasf\Desktop\merlinrecord2015.xml

Submit URL
Examples of Transformations

• 505 0 $a What is RDA? -- RDA and the international context - FRBR and FRAD in RDA -- Continuity with AACR2 -- Where do we see changes? -- Implementing RDA -- Advantages, present and future.

• bf:contentsNote "What is RDA? -- RDA and the international context -- FRBR and FRAD in RDA -- Continuity with AACR2 - Where do we see changes? -- Implementing RDA - Advantages, present and future." ;

• 300 $a vii, 117 p. : $b ill. ; $c 28 cm.

• bf:dimensions "28 cm." ;
Schema.org

- Standard for semantic markup of web data
- Developed by Yahoo!, Google, Bing, and Yandex
- Structures data on the Web

- Schema Bib Extend Community Group
  - https://www.w3.org/community/schemabibex/

- Overlap with and complements BIBFRAME
February 2015 update

• Year of experimentation completed
• Next steps
  • Adjust BIBFRAME vocabulary
  • Continued development of tools
  • Pilots for original cataloging
    • Library of Congress
    • Stanford and five other libraries
  • Further development of MARC to BIBFRAME and BIBFRAME to MARC conversion tools
  • OCLC: more work with OCLC Schema.org and BIBFRAME approaches
  • Zepheira working on tools and working libraries on projects
  • Large scale experimentation

Biographical Sketch: Harry S Truman
www.trumanlibrary.org - Harry S. Truman Presidential Library and Museum
[Harry Truman and family] Harry S. Truman was born in Lamar, Missouri on May 8, 1884, the son of John Anderson Truman and Martha Ellen (Young) Truman.

Harry S. Truman - Wikipedia, the free encyclopedia
en.wikipedia.org/wiki/Harry_S._Truman
Harry S. Truman (May 8, 1884 – December 26, 1972) was the 33rd President of the United States (1945–53). As the final running mate of President Franklin D. Haberdasher - Bess Truman - Margaret Truman - Alben W. Barkley

Harry S. Truman | whitehouse.gov
https://www.whitehouse.gov/1600/presidents/harrystruman - White House
During his few weeks as Vice President, Harry Truman scarcely saw President Franklin Roosevelt, and received no briefing on the development of the atomic ...

Harry Truman - U.S. Presidents - HISTORY.com
www.history.com/topics/us-presidents/harry-truman - History
Find out more about the history of Harry Truman, including videos, interesting articles, pictures, historical features and more. Get all the facts on HISTORY.com.

Harry S. Truman - Biography - U.S. President - Biography.com
www.biography.com/people/harry-s-truman-9511121 - FYI
Sworn in as the 33rd president after Franklin Delano Roosevelt's sudden death, Harry S. Truman presided over the end of WWII and dropped the atomic bomb ...

Facts on President Harry S. Truman | President Truman ... www.trumanlittlewhitehouse.com/... - Harry S. Truman Little White House
Expand your knowledge of presidential history with our facts on President Harry S. Truman.

Harry S. Truman
33rd U.S. President
Harry S. Truman was the 33rd President of the United States. As the final running mate of President Franklin D. Roosevelt in 1944, Truman succeeded to the presidency on April 12, 1945, when Roosevelt died after months of declining health. Wikipedia
Born: May 8, 1884, Lamar, MO
Died: December 26, 1972, Kansas City, MO
Full name: Harry S. Truman
Presidential term: April 12, 1945 – January 20, 1953
Vice president: Alben W. Barkley (1949–1953)
Education: University of Missouri-Kansas City, More

People also search for View 15+ more
Show Me the Linked Data in Libraries!
• A two year IMLS grant project of the University of California Davis Library & Zepheira
• Goal: Switch entire database to graph-based/triple-store
• Underlying model: BIBFRAME
• Software: Kuali Ole (with coding modifications)
• Seeking feedback and comments to make model as universal as possible
  • Use cases
Copy Cataloging for Non-Rare Books Using OCLC Connexion Client

posted on October 25, 2014 by Carl G Stahmer | 2 Comments

The following workflow represents the current work process in place at the UC Davis library when copy cataloging a non-rare book using OCLC connection client. It is one of the work functions that will be converted to a native Linked Data workflow as part of the project.

Actors

1. Cataloging Specialist – Medium level of experience.
2. Copy Cataloger – Lowest level of experience.
3. Supporting systems and tools (e.g. harvesting daemon).

Description

User Story: “As a library cataloger I need to find and claim a record in OCLC for a published book to include in my integrated library system with additional metadata for its location in library collection.”

Success Metric(s)

1. The new set of bibliographic metadata elements will be present and discoverable by library users of the OCLC Melvyl System.
"I believe everyone benefits from the visibility of libraries and their content on the Web."

- The Libhub Initiative Pledge

- One of founding sponsors: Zepheira
- Other vendor sponsors: Atlas, Innovative, SirsiDynix
- Goal: “publish BIBFRAME resources to the Web” and enhance discoverability
Linked Data for Libraries (LD4L)

• Collaboration of Cornell, Harvard and Stanford
• Funded by nearly $1 million two-year grant (Andrew W. Mellon Foundation) in February 2014
• Goal of project: Create a triple store called the Scholarly Resource Semantic Information Store (SRSIS) model
  • Intent to use existing ontologies
  • Also to use open source software
  • One deliverable: to build a Blacklight interface to facilitate searching of triples
LODLAM
Linked Open Data in Libraries Archives and Museums

Latest News
Below, you'll find all of the latest news and goings on in the LODLAM community, including upcoming events, published reports, etc. You can also keep an eye on #LODLAM on the Twitter stream below and to the right.

LODLAM Resources
You can find resources about LODLAM in the full resources page. Here you can find archived talks, presentations, and articles to get up to speed on Linked Open Data with a particular emphasis on libraries, archives, and museums.

SUMMIT 2015
Save the date for the 3rd International LODLAM Summit, taking place in Sydney, June 29-30, 2015.
More details available in the coming weeks.
• Inspiring because presenters have emphasized that their “Project led, implemented & managed by 2 busy faculty librarians”

• Library first formed a study group on linked data

• Then after a year, worked on prototype
  • Enabled them to verify the ability to generate LOD without investment in new systems
  • Instrumental in acquiring administrative support for project
Design Phase

- Selecting Technologies
- Data model
- Mapping
- Determining rules
  to create URI’s
Implementation

Initial Data prep in ContentDM repository
• Further Metadata Work
• Reconciled Data
• Generated triples/RDF schema & exported theses
Import linked data into triple stores

• Runs parallel to their ContentDM repository

• Finished result: Video of content in Pivot Viewer:
  • https://youtu.be/-83FTKEkYZ0
How Feasible is it to Implement Linked Data?

• Independent project requires
  • Technical knowledge or willingness to learn (RDF, linked data principles)
  • Group project
  • Administrative support/buy-in
  • IT/Technical Support (triple store)
When Can We Expect to See Linked Data in Our Libraries?

• Pioneers/early adopters are paving the way
• Not financially feasible to leap until details have been worked out
• Momentum may propel linked data era forwards
References

Will be posted on the MOBIUS Conference Website
Questions

• Humorous Disclaimers:
  • We’re 3 busy faculty librarians
  • We’re librarians, not “linked data specialists,” Jim
  • Our crystal balls work imperfectly
Zepheira Training

Linked Data and BIBFRAME Practical Practitioner Training

Curriculum

- Suggested Schedule
  - 5-6 Calendar Weeks
- 5 modules (Self-Paced On-demand Videos)
  1. History of the Web and Linked Data
  2. BIBFRAME Foundations
  3. BIBFRAME - Data Modeling Essentials
  4. Interacting with BIBFRAME Materials and Profiles via Zepheira's Training Tool Suite
  5. Readiness Assessment, Change Management and Governance
- Exercises, Discussions
- Regular Office Hours with Colleagues and Instructors

https://zepheira.com/solutions/library/
Contact Information

MOBIUS 2015

Sandy Rodriguez
employedBy UMKC
hasEmail Address RodriguezSan @umkc.edu

Felicity Dykas
employedBy MU
hasEmail Address DykasF@missouri.edu

Mary Aycock
employedBy MS&T
hasEmail Address aycockM@mst.edu