# Library Data and Its Future on the Web

# Sandy Rodriguez

University of Missouri—Kansas City

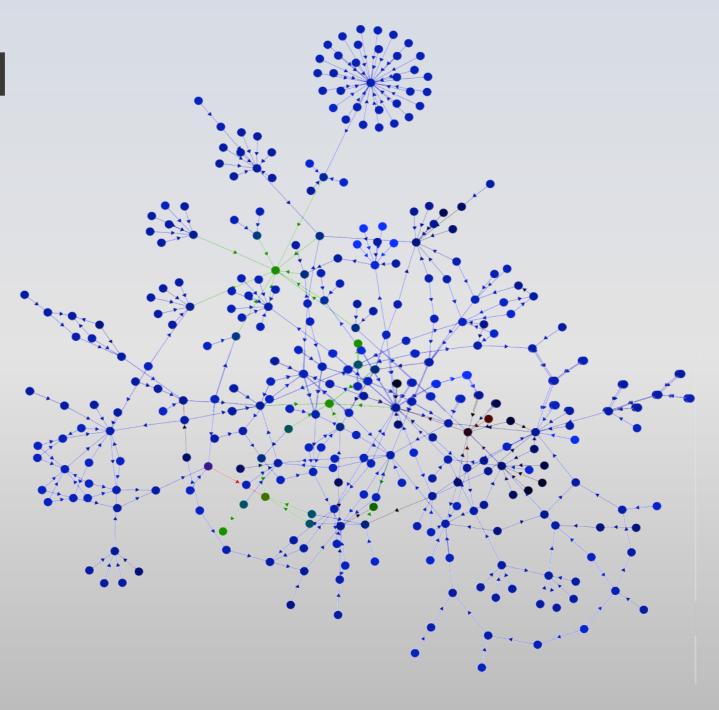
**Felicity Dykas** 

University of Missouri—Columbia

**Mary Aycock** 

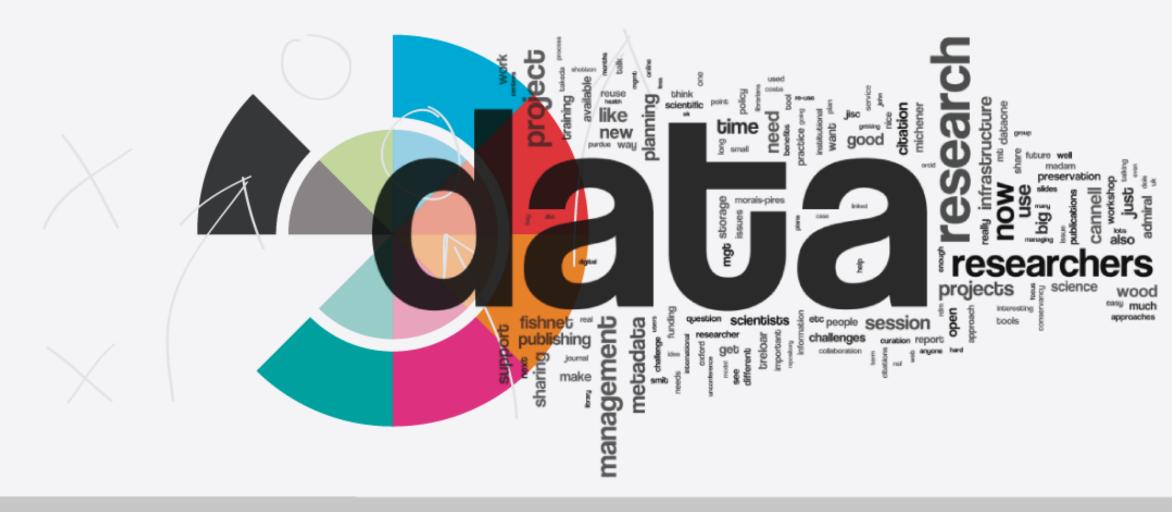
Missouri University of Science & Technology

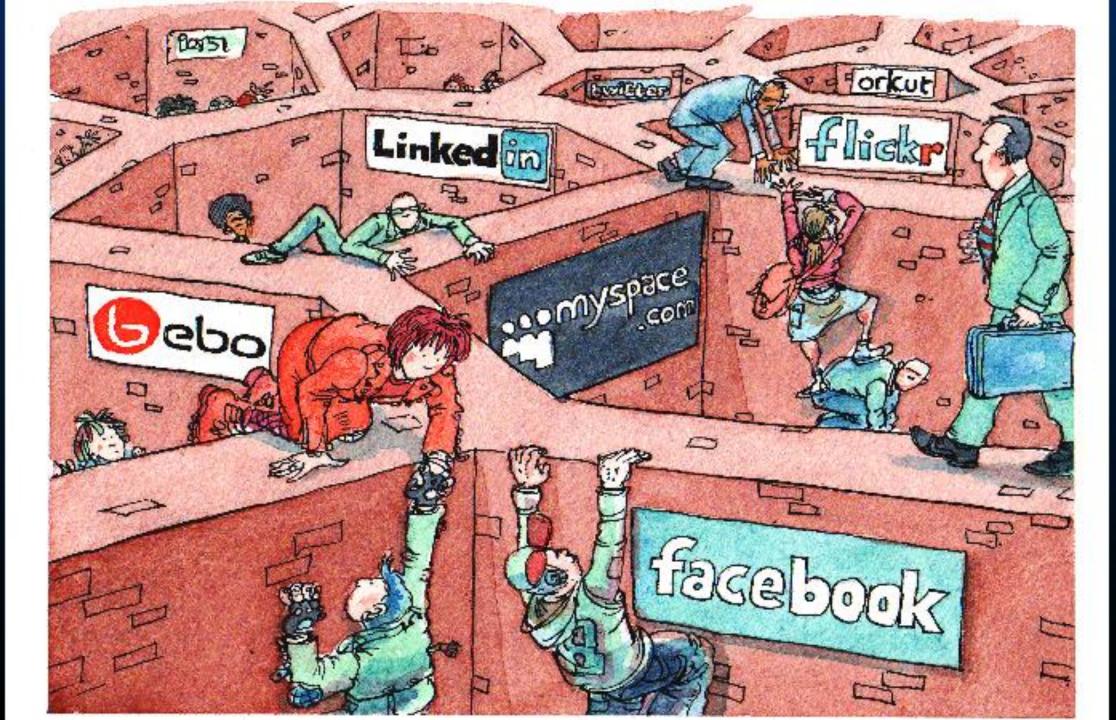
**MOBIUS Annual Conference 2015** 

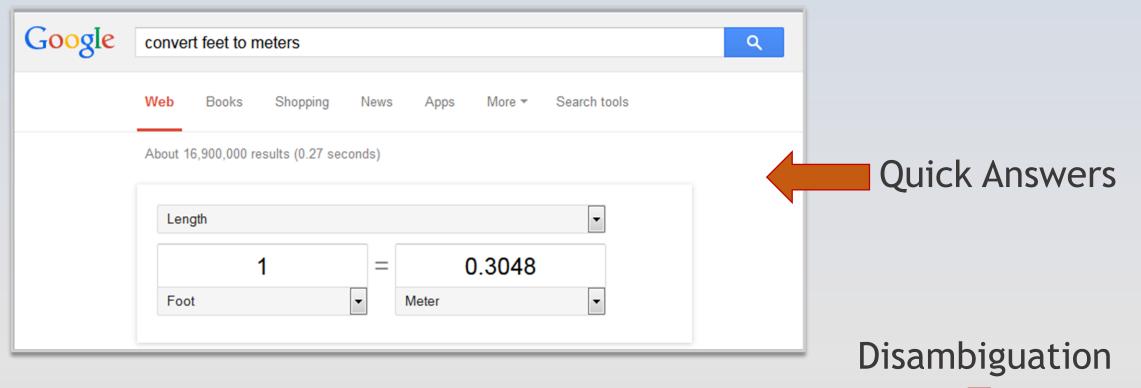


# Outline

- Linked Data and Linked Open Data
- BIBFRAME
- Linked Data in Libraries







# **Enriched Data**



#### See results about

#### Bonobo (Musician)

Born: March 30, 1976 (age 39), Brighton, United Kingdom

Record labels: Ninja Tune, Tru Thoughts

## ghton, United Kingdom ughts

#### Bonobo (Primate)

Higher classification: Chimpanzee Scientific name: Pan paniscus



Feedback

#### Tim Berners-Lee - World Wide Web Consortium

www.w3.org/People/Berners-Lee/ ▼ World Wide Web Consortium ▼ Tim Berners-Lee invented the World Wide Web in 1989. He founded and Directs the World Wide Consortium (W3C) the forum for technical development of the ... Longer Biography - Answers for young people - Frequently Asked Questions - Card

#### Tim Berners-Lee - Wikipedia, the free encyclopedia

en.wikipedia.org/wiki/**Tim\_Berners-Lee** ▼ Wikipedia ▼ Sir Timothy John "Tim" Berners-Lee, OM, KBE, FRS, FREng, FRSA, DFBCS (born 8 June 1955), also known as TimBL, is an English computer scientist, best ... Awards and honours ... - DFBCS - Conway Berners-Lee - Mary Lee Woods

#### In the news



# Tech Time Warp of the Week: A 1930s Futurist Envisioned the Web With a 'Televised Book'

Wired - 2 days ago

Tim Berners Lee created the world wide web in 1989. But the idea of hypertext stretches ...

More news for tim berners lee

#### Tim Berners-Lee | Internet Hall of Fame

internethalloffame.org > Inductees ▼ Internet Hall of Fame ▼ In 1989, Tim Berners-Lee invented the World Wide Web, an Internet-based hypermedia initiative for global information sharing while at CERN, the European ...

#### Tim Berners-Lee | Speaker | TED.com

https://www.ted.com/speakers/tim\_herners\_lee v\_TED\_v



## Tim Berners-Lee

Computer Scientist

Sir Timothy John "Tim" Berners-Lee, OM, KBE, FRS, FREng, FRSA, DFBCS, also known as TimBL, is an English computer scientist, best known as the inventor of the World Wide Web. Wikipedia

Born: June 8, 1955 (age 59), London, United Kingdom

Spouse: Rosemary Leith (m. 2014)

Organizations founded: World Wide Web Consortium, World Wide

Web Foundation

Books: Weaving the Web: The Original Design and Ultimate Destiny of

the World Wide Web by its Inventor

Parents: Mary Lee Woods, Conway Berners-Lee

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 ↑ ?



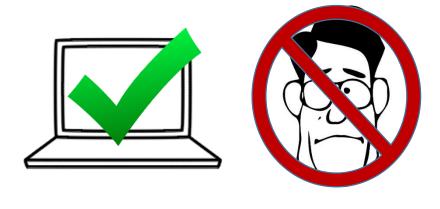


# = Resource Description Framework

Standard model for publishing and linking data on the web

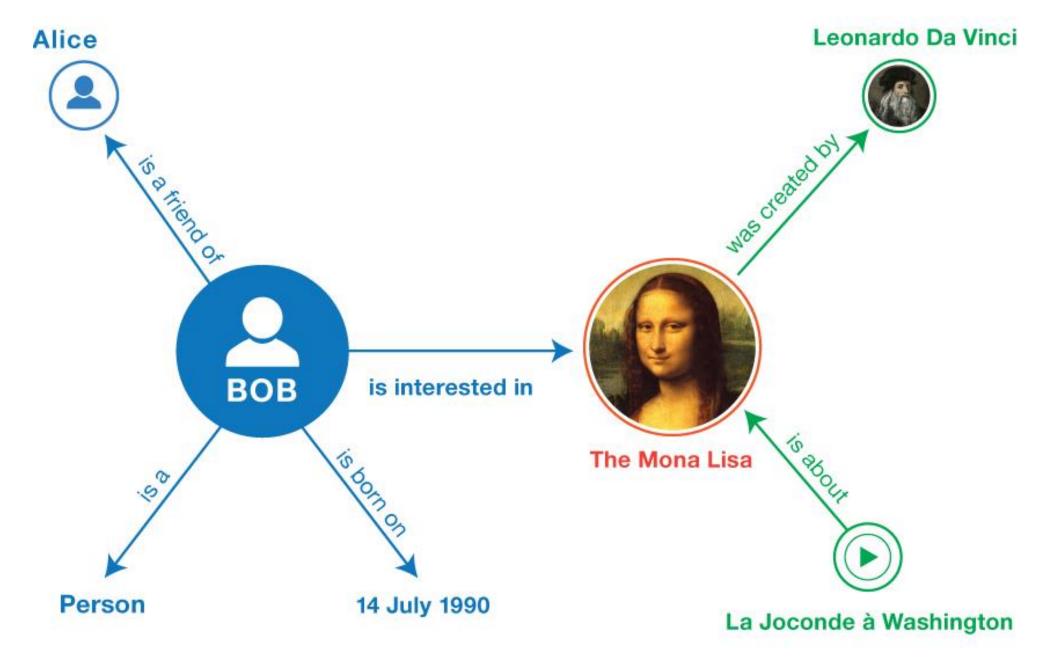


Designed for:



# RDF Triples



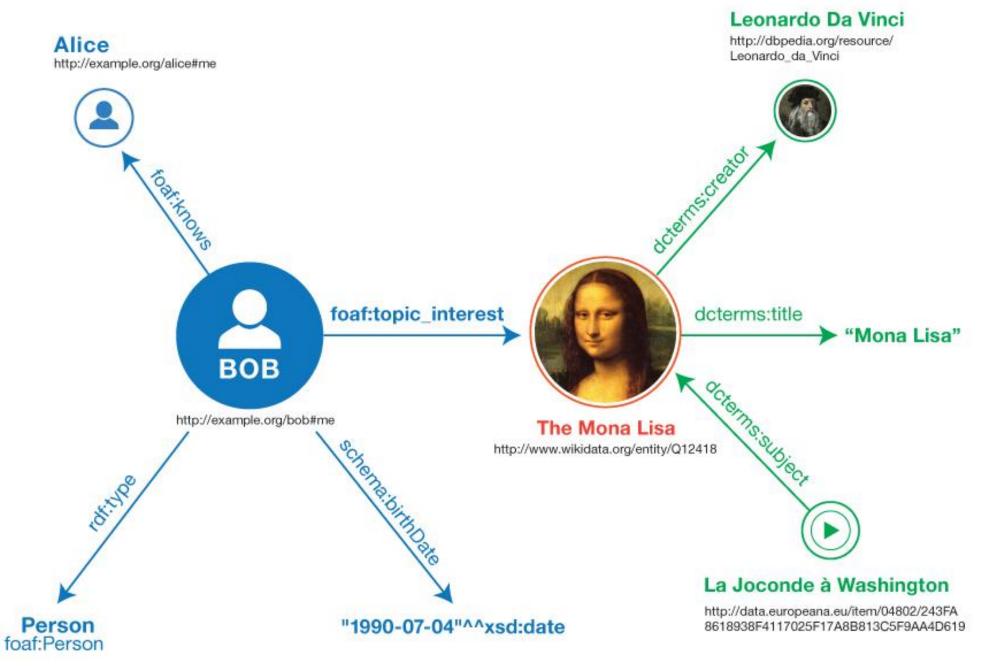


http://www.w3.org/TR/2014/NOTE-rdf11-primer-20140225/

# 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**

<a href="http://www.wikidata.org/entity/Q12418">http://www.wikidata.org/entity/Q12418</a>>

<a href="http://purl.org/dc/terms/creator">http://purl.org/dc/terms/creator</a>

<a href="http://dbpedia.org/resource/Leonardo\_da\_Vinci">http://dbpedia.org/resource/Leonardo\_da\_Vinci</a>.

## **Turtle**

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

<a href="http://www.wikidata.org/entity/Q12418">http://www.wikidata.org/entity/Q12418</a> dc:creator <a href="http://dbpedia.org/resource/Leonardo\_da\_Vinci">http://dbpedia.org/resource/Leonardo\_da\_Vinci</a> .

http://rdf-translator.appspot.com/

# Copyrighted Material Structured data on the Web David Wood Marsha Zaidman Luke Ruth WITH Michael Hausenblas FOREWORD BY Tim Berners-Lee

# Linked Data: Issues

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

Wood, D., Zaidman, M., Ruth, L., & Hausenblas, M. (2014). Linked data: Structured data on the Web.

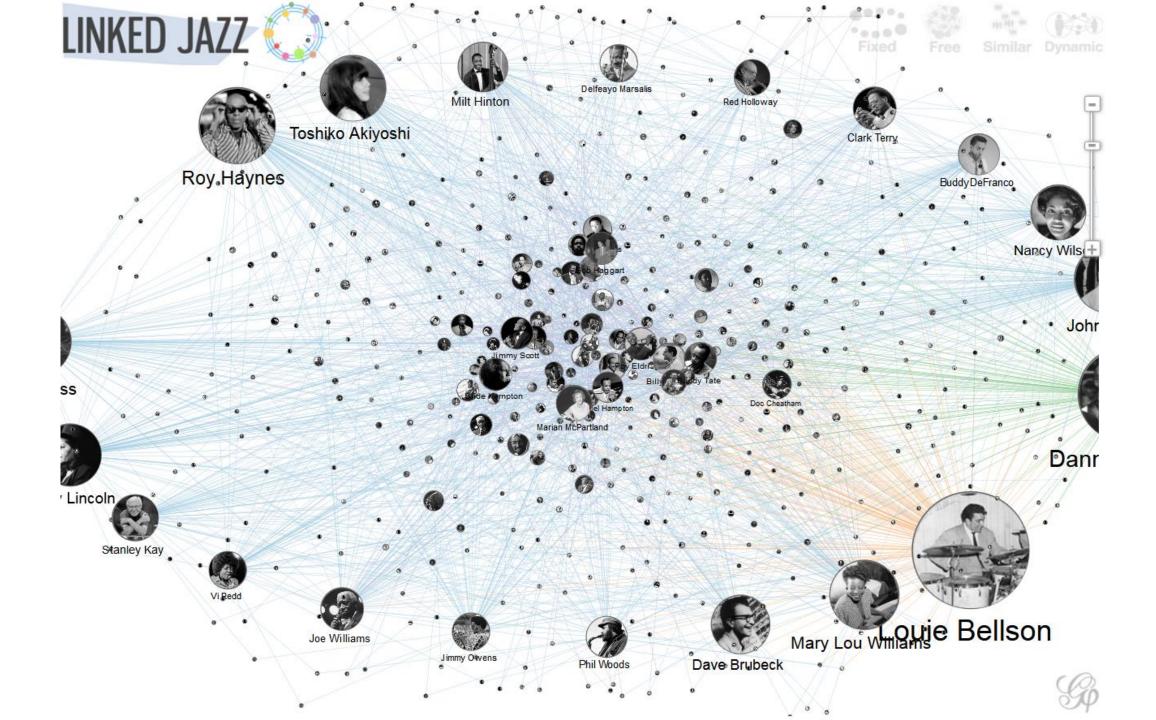
# Linked Data Tools



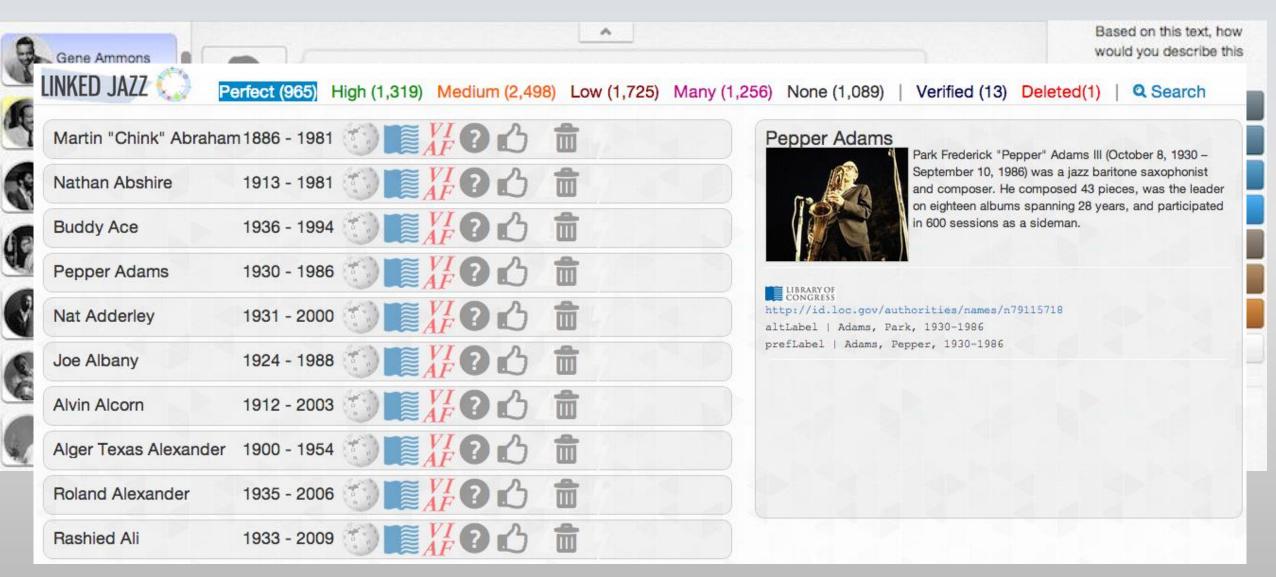
Refresh

new 3

none 3416



# Linked Jazz - Developing Tools



# Be Bold



# Is there a place for libraries in a linked data environment?

# bf: BIBFRAME

# 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/





ASK A LIBRARIAN

DIGITAL COLLECTIONS

LIBRARY CATALOGS

Search

Search Locigov

- Print Subscribe Ø Share/Save ♥ Give Feedback

GO

#### The Library of Congress > BIBFRAME

#### BIBFRAME



- Home
- Frequently asked questions
- Webcasts & presentations
- Contact us

#### Model & Vocabulary

- BIBFRAME model
- > BIBFRAME vocabulary rank
  - ➤ Model view
  - ➤ Catedory view
  - ➤ List view range
- BIBFRAME specifications

### **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.



#### **General Information**

- > Frequently asked questions
- > Webcasts and presentations
- Contact us
- > Document archive

#### Implementation & Testing

> BIBFRAME testbed

#### Model & Vocabulary

- > BIBFRAME model
- BIBFRAME vocabulary라
  - Vocabulary description
  - Terminology & conventions
- > BIBFRAME draft specifications
  - BIBFRAME Profiles
  - BIBFRAME Authorities
  - BIBFRAME Relationships
- Analyses

#### What's New

#### **BIBFRAME Update Forum at ALA**

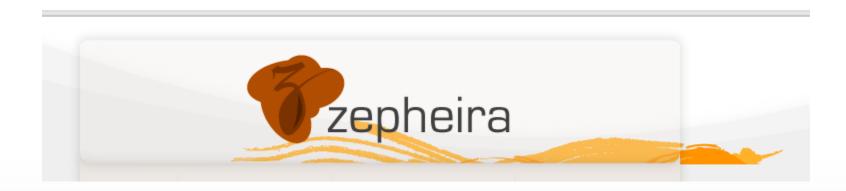
February 1, 10:30-11:30am McCormick Place West, W196a

BIBFRAME at ALA »

#### New White Paper Available

Common Ground: Exploring
Compatibilities Between the Linked

Data Madala of the Library of Congress



- 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

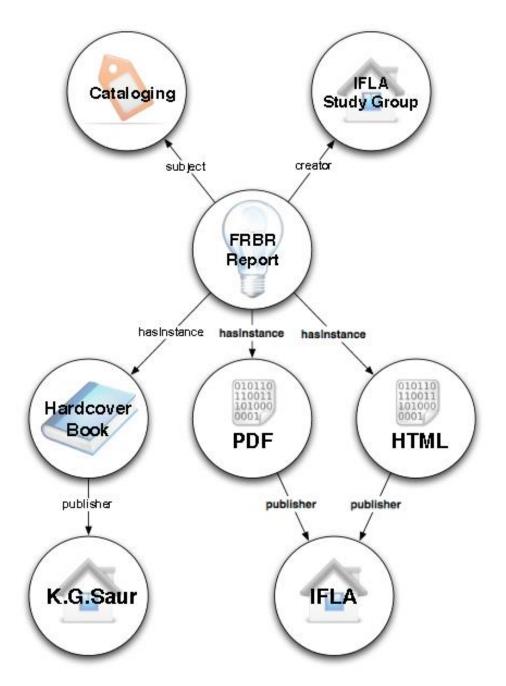
# BIBFRAME Includes ...

- BIBFRAME model
- BIBFRAME vocabularies
  - Classes
  - Properties
- BIBFRAME profiles

Work (Creative Work)	Instance	Authority	Annotation
Covers FRBR Work and Expression	FRBR Manifestation and Item	A thing or concept associated with a BIBFRAME Work or Instance	Used to give more information about a BIBFRAME Work, Instance, or
The conceptual essence of a resource	An individual, material embodiment of a Work	Provides access points	Authority  Examples: Library holdings, Reviews,
Serves as a Web based control point and reference point	Can be physical or digital	Includes: Agents (people, organizations, and jurisdictions), Places, Subjects	Cover Images, Tables of Contents, Descriptions

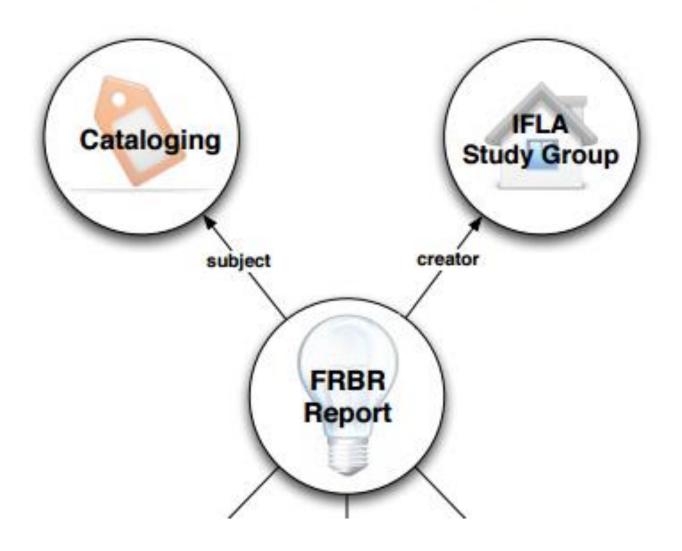
# BIBFRAME Properties (Selected)

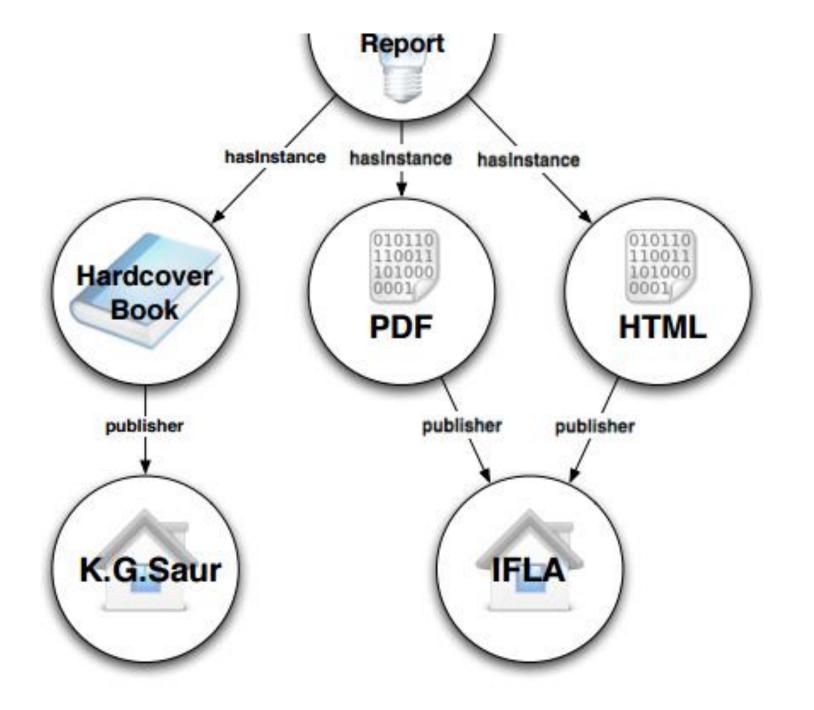
Work (Creative Work)	Instance	Authority	Annotation
absorbedBy classificationLcc contributor creator genre language series subject title workTitle	accompaniedBy colorContent contentNotes duration edition extent formatOfMusic frequency isbn10 title	Authorized    AccessPoint identifier label relatedTo	annotates annotation AssertedBy

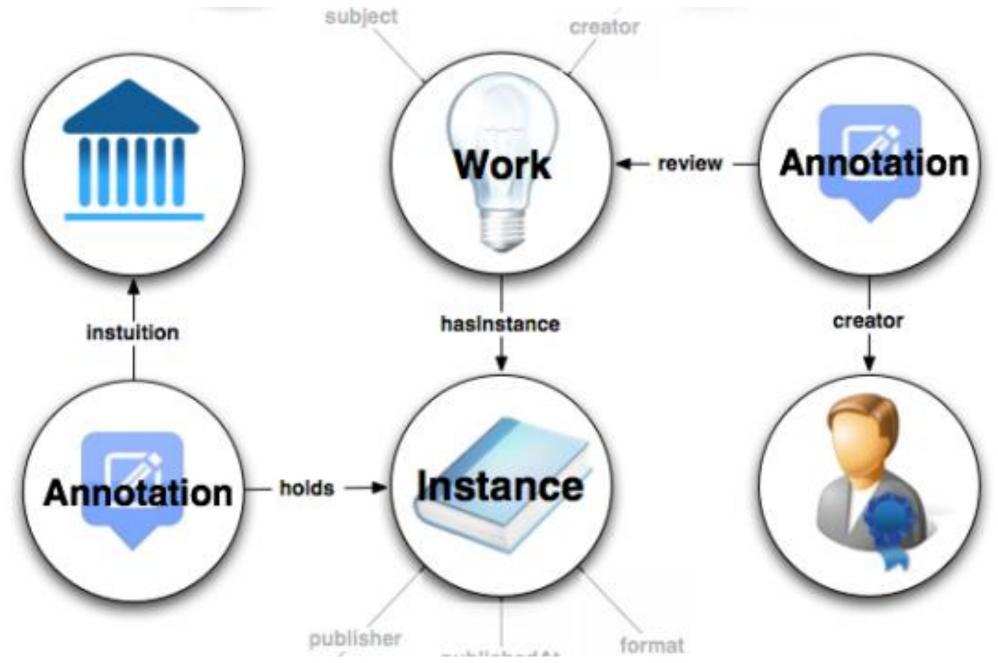


http://www.loc.gov/bibframe/pdf/marcld-report-11-21-2012.pdf

Instances as defined by this XML serialization is shown in Figure 5.







http://www.loc.gov/bibframe/pdf/marcld-report-11-21-2012.pdf

### Focus on Data Elements

- Like FRBR and RDA, BIBFRAME focuses on data elements
- Not records

```
LEADER 00000cam a2200457 i 4500
001 885148293
003
      OCOLC
005 20150406155922.0
008 14072152015 iluab b 001 0 eng
010 2014017912
016 7 016057710 2Uk
019 756280902 a757462586 a803890454 a891691753 a893977550
      9780838911310 galk, paper
       (OCOLC)885148293 z(OCOLC)756280902 z(OCOLC)757462586
       z(OCOLC)803890454 z(OCOLC)891691753 z(OCOLC)893977550
       z(OCoLC)894201357
      DLC|beng|erda|cDLC|dOCLCO|dTXA|dYDXCP|dFDA|dERASA|dTOH
       |dwau|dBTCTA|dBDX|dUKMGB|dCDX|dOGK|dutorBLW
049 UMSL
050 00 Z695.6|b.A555 2015
082 00 025.3/46 223
100 1 Andrew, Paige G., eauthor.
245 10 RDA, Resource description & access and cartographic
       resources / cPaige G. Andrew, Susan M. Moore, Mary Lynette
       Larsgaard.
246 3 RDA, Resource description and access and cartographic
       resources.
246 10 Resource description & access and cartographic resources.
246 18 RDA and cartographic resources
264 1 Chicago : | bALA Editions, an imprint of the American
      Library Association, c2015.
300 viii, 144 pages :|billustrations ;|c26 cm
336 text|btxt|2rdacontent.
337 unmediated bn 2rdamedia.
338 volume|bnc|2rdacarrier.
504 Includes bibliographical references and index.
505 0 The past is prologue -- RDA and FRBR entities as applied
       to cartographic resources : an overview -- Comparing
       standards : continuing, different and added practices --
       Navigating RDA to catalog cartographic resource elements -
       - Cartographic resources cataloging : moving forward.
520 8 As the cataloging universe moves into the era of PDA:
```

#### 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**

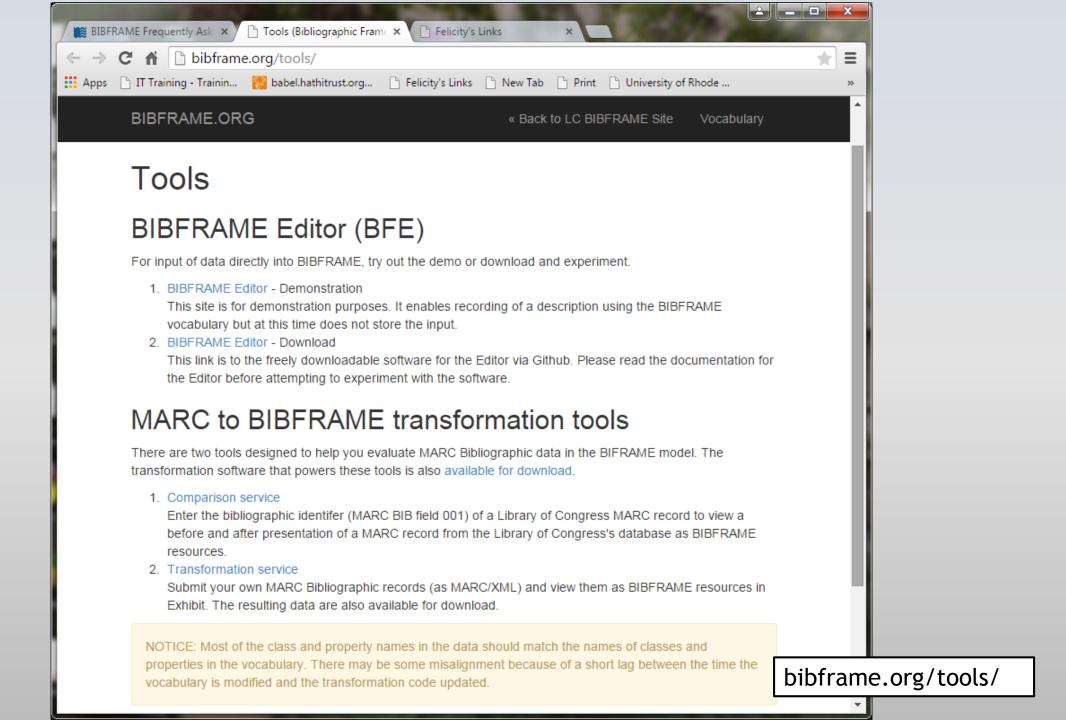
Defines a set of classes and properties

- 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?

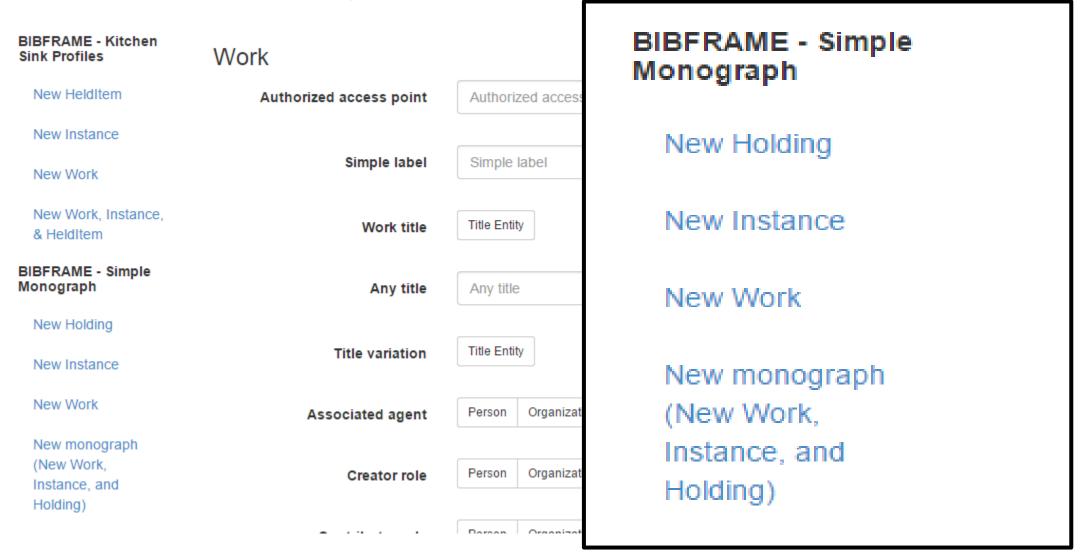
BIBFRAME Profiles
BIBFRAME tools

http://bibframe.org/tools/



#### BIBFRAME Editor

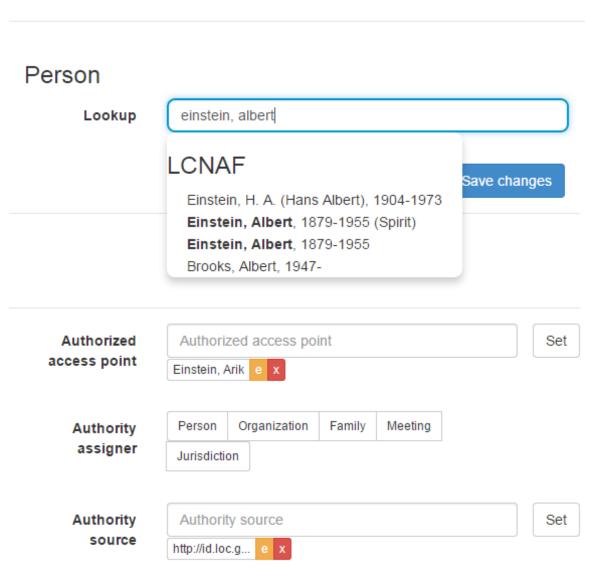
Get the code here. In the meantime, select a Profile below to get started.

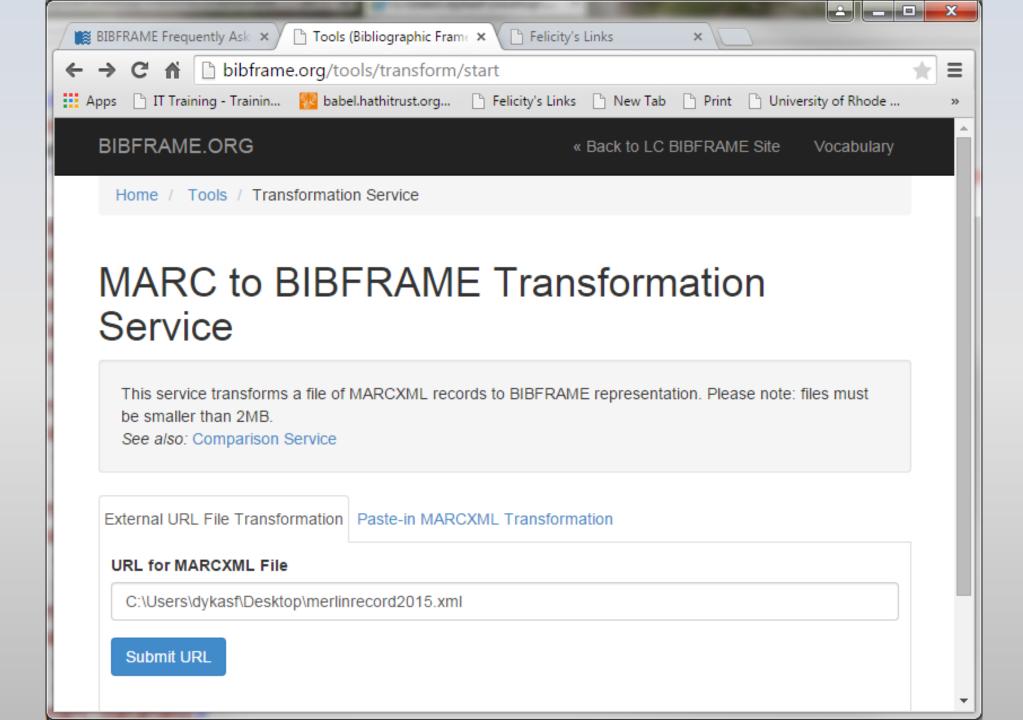


#### Work - Monograph

Authorized access	Authorized access point						Set
point							
Title	Title						Set
Title variation	Title Entit	ty					
Author	Person	Organization	Family	Meeting	Jurisdiction		
			I			1	
Editor	Person	Organization	Family	Meeting	Jurisdiction		
Translator	Person	Organization	Family	Meeting	Jurisdiction		
Date of work	Date of work						Set
	Diago						
Place of creation	Place						

Person





# **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.";

```
DI.CLASSITICACIONECC \NCCP.//IU.IUC.guv/auchulites/Classiticaciun/2024.ij.N4
bf:contributor [ a bf:Organization ;
        bf:authorizedAccessPoint "American Library Association.";
        bf:hasAuthority [ a madsrdf:Authority ;
                 madsrdf:authoritativeLabel "American Library Association." ] ;
        bf:label "American Library Association." ];
bf:creator [ a bf:Person ;
        bf:authorizedAccessPoint "Oliver, Chris, 1951-";
        bf:hasAuthority [ a madsrdf:Authority ;
                 madsrdf:authoritativeLabel "Oliver, Chris, 1951-" ];
        bf:label "Oliver, Chris, 1951-" ];
bf:derivedFrom <http://id.loc.gov//resources/bibs/16266874.marcxml.xml> ;
bf:hasAnnotation [ a bf:Annotation ;
        bf:annotates <http://id.loc.gov//resources/bibs/16266874> ;
        bf:changeDate "2011-02-02T15:44";
        bf:derivedFrom <a href="http://id.loc.gov//resources/bibs/16266874.marcxml.xm">http://id.loc.gov//resources/bibs/16266874.marcxml.xm</a>
```

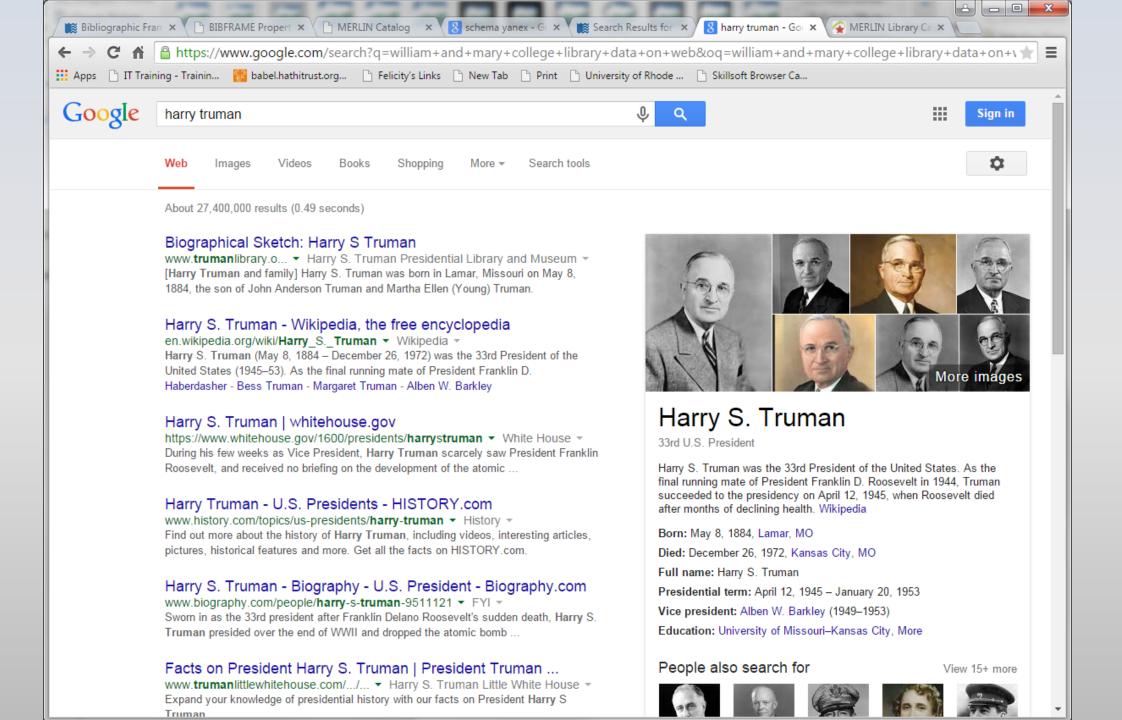
# Schema.org

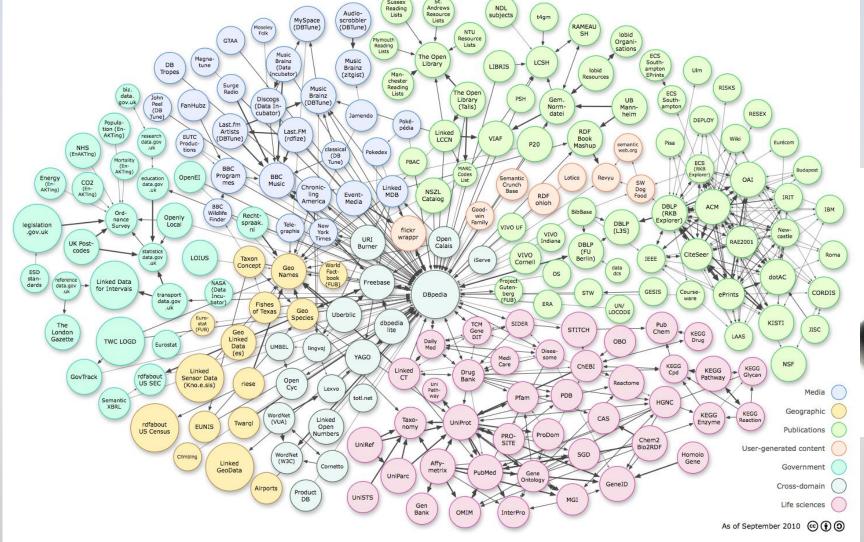
- Standard for semantic markup of web data
- Developed by Yahoo!, Google, Bing, and Yandex
- Structures data on the Web

- 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







# Show Me the Linked Data in Libraries!

# BIBFLOW AN IMLS PROJECT OF THE UC DAVIS LIBRARY & ZEPHEIRA HOME ABOUT RESEARCH QUESTIONS FINDINGS REPORTS TEAM CONTACT WORKFLOW ANALYSIS

- 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

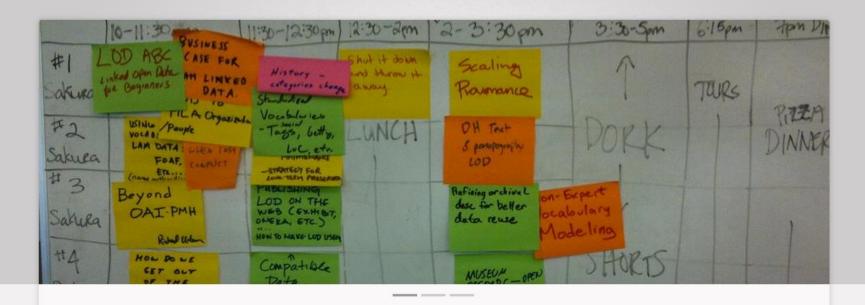
# Linked Data for Libraries (LD4L)

- 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

About Resources RSS Community

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.

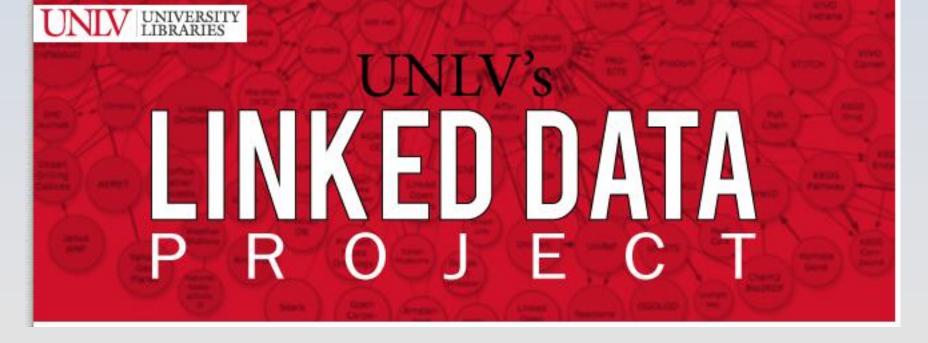


#### **LODLAM Resources**

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



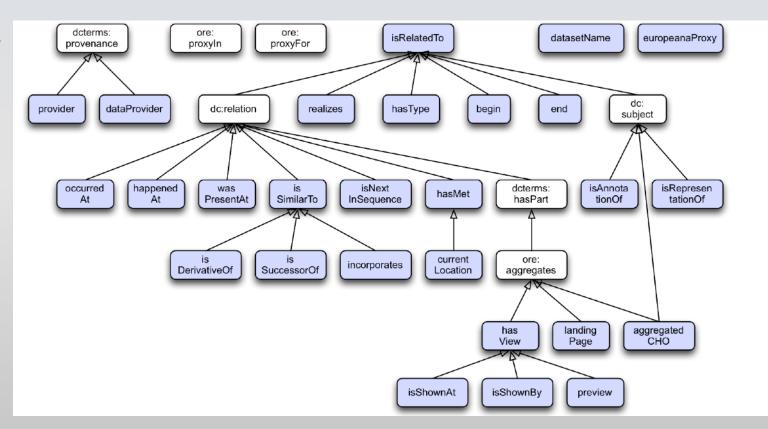
Save the date for the 3rd International LODLAM Summit, taking place in Sydney, June 29-30, 2015. More details available in the coming weeks.



- 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





Costume design drawing,
yellow calypso costume, circa
 1945-1955



2. Costume design drawing, samba costume pencil sketch, circa 1945-55



■ Grid | I List | III Tile

3. Costume design drawing, samba costume sketch, circa 1945-55



4. Costume design drawing, pink American Indian headdress, circa 1945-55



5. Costume design drawing, pink American Indian headdress and shield, circa 1945-55



6. Costume design drawing, template of nude female dance poses, circa 1945-55



 Costume design drawings, various cultures, circa 1945-55



8. Costume design drawings, music, food, and drink, circa 1945-55

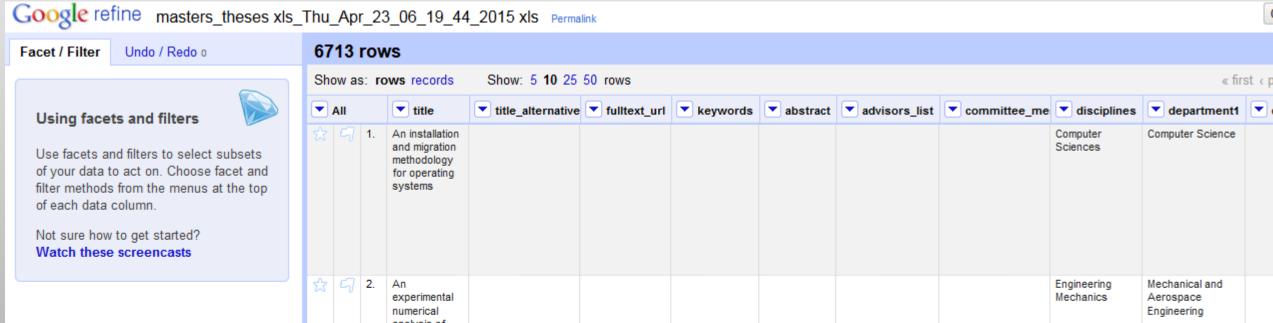


 Set design drawing, eight productions for the Jean
 Devlyn Dancers, circa 1945-55



A free, open source, powerful tool for working with messy data

- 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

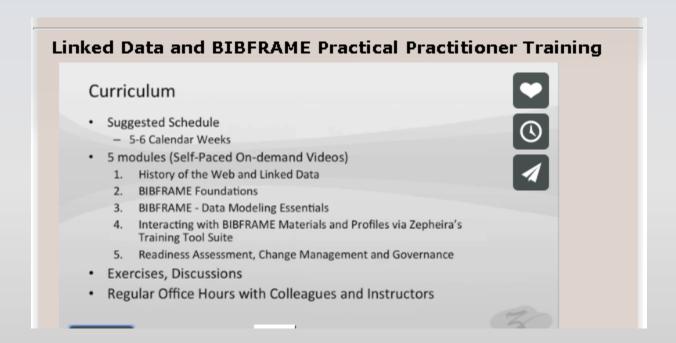
Will be posted on the MOBIUS Conference Website

# Questions

- Humorous Disclaimers:
  - We're 3 busy faculty librarians
  - We're librarians, not "linked data specialists," Jim
  - Our crystal balls work imperfectly



# Zepheira Training



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

