

# Putting Library Linked Data into Practice: BIBFRAME and LD4P



Audrey Pearson,  
Yale University Beinecke Rare Book & Manuscript Library

Connecticut Library Association  
2019 Technical Services Section Fall Program

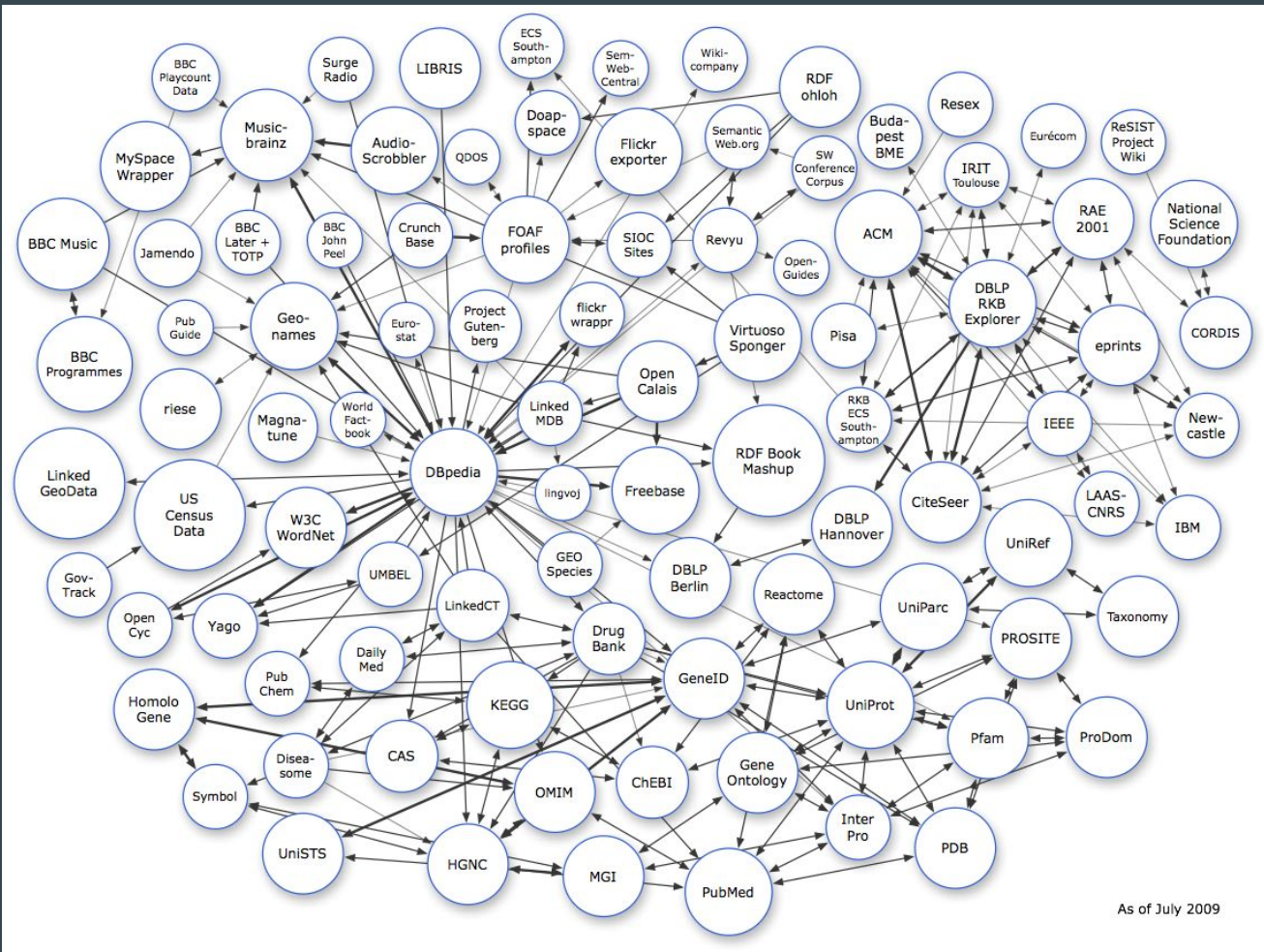
# Outline

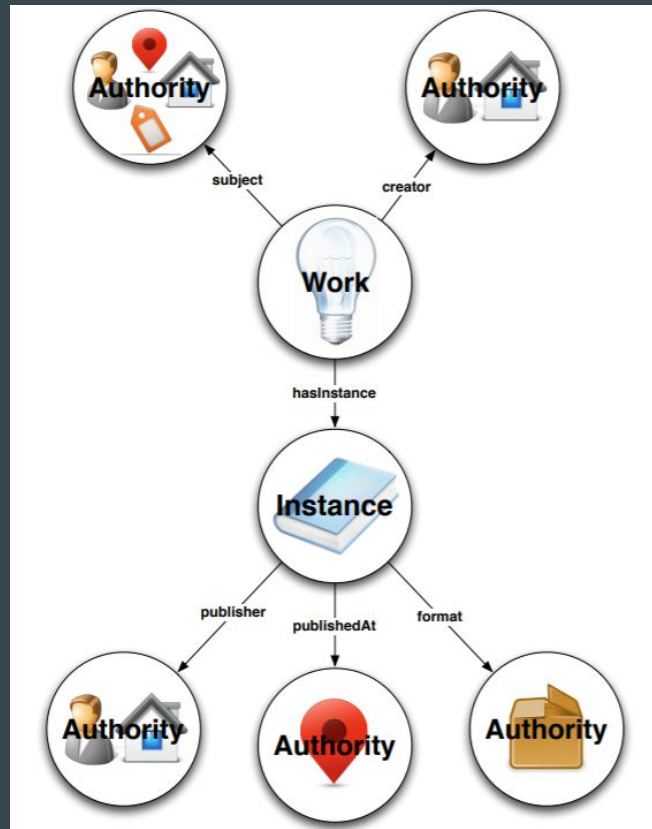
- What is Linked Data?
- BIBFRAME
- LD4P
  - Yale's LD4P Project
- Sinopia Demo
- What's next?

# Linked Data

Tim Berners-Lee's four principles of linked data (2009):

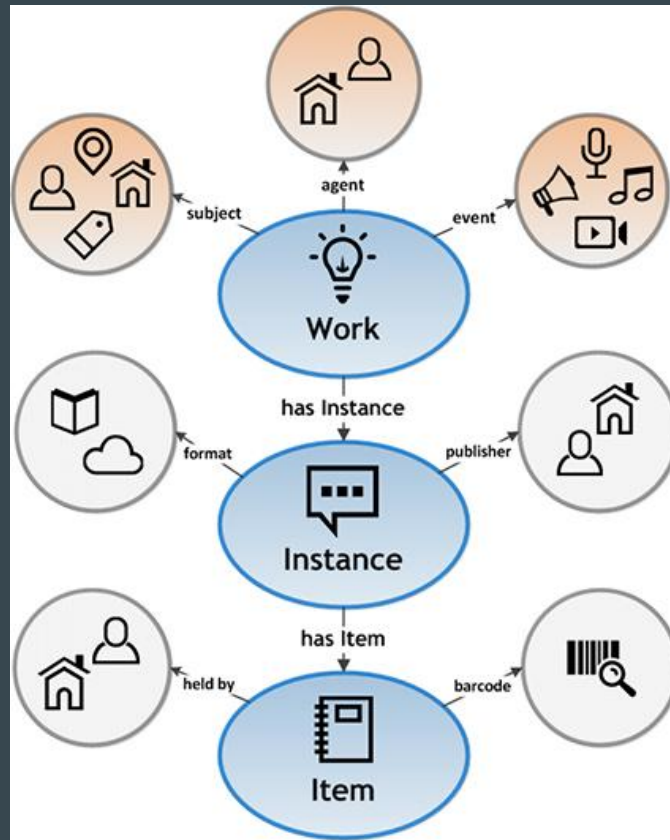
- Use URIs as names for things
- Use HTTP URIs so that people can look up those names
- When someone looks up a URI, provide useful information, using the standards
- Include links to other URIs so that they can discover more things





## BIBFRAME 1.0 Model

(via <https://www.loc.gov/bibframe/pdf/marclid-report-11-21-2012.pdf>)



## BIBFRAME 2.0 Model

(via <https://www.loc.gov/bibframe/docs/bibframe2-model.html>)

# LC MARC to BIBFRAME Resources

MARC to BIBFRAME Conversion Application:

<https://github.com/lcnetdev/marc2bibframe2>

MARC to BIBFRAME Conversion specifications: <http://www.loc.gov/bibframe/mtbf/>

The Library of Congress > BIBFRAME > MARC 21 to BIBFRAME 2.0 Conversion Specifications

**BIBFRAME**

[Print](#) [Subscribe](#) [Share/Save](#) [Give Feedback](#)

## MARC 21 to BIBFRAME 2.0 Conversion Specifications

The following specifications were developed to support a pilot in the use of BIBFRAME 2.0 at the Library of Congress. They specify the conversion of MARC Bibliographic records to BIBFRAME Work, Instance and Item descriptions, and MARC Authority records for titles and name/titles to BIBFRAME Work descriptions.

**bf: BIBFRAME**

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

**Model & Vocabulary**

- BIBFRAME model
- BIBFRAME vocabulary
  - Category view
  - List view
  - RDF View
  - Extension list view

### MARC Bibliographic Conversion Specifications

- Fields 001-005, 007 – Control, physical description (Excel, 38 KB, 06/10/2019)
- Fields 006, 008 - Fixed fields (Excel, 33 KB, 06/10/2019)
- Fields 010-048 – Identifiers, etc. (Excel, 23 KB, 06/10/2019)
- Field 048 - Instrument codes (Excel, 14 KB, 06/10/2019)
- Fields 050-088 - Class/call numbers, etc. (Excel, 15 KB, 06/10/2019)
- Fields 1XX, 6XX, 7XX, 8XX; X00, X10, X11 – Names (Excel, 12 KB, 06/10/2019)
- Fields 200-24X, except 240 – Titles (Excel, 14KB, 06/10/2019)
- Fields 240, X30, etc. – Uniform titles (Excel, 12 KB, 06/10/2019)
- Fields 250-270 – Edition, imprint, etc. (Excel, 14 KB, 06/10/2019)

### Usage Notes

These specifications are written from the perspective of MARC so that each element in MARC would at least be considered, even if not converted. The specifications are presented in MS Excel files accompanied by a few explanatory specifications in MS Word. If there is little or no use of an element in the Library of Congress records, the Specifications usually say "nac" (no attempt to convert) in the conversion column. Others may want to augment the specifications to address "nac" elements. Some elements in a MARC record are not relevant outside the MARC environment so they are simply marked "ignore".

A shorthand is used for the specification of the conversion from MARC to BIBFRAME RDF. The shorthand uses W for

# BIBFRAME Comparison Tool Compare MARCXML to BIBFRAME2

Identifier  Bib ID  LCCN **Serialization**  Turtle  RDF XML

Lookup Name URIs  No  Yes

## MARC

```
01181cam a22002774a 4500
001 14273006
005 20070131200014.0
008 060222s2006 nyua 000 0 eng
010 $a 2006041192
020 $a006089721X (pbk.)
020 $a9780060897215
035 $(OCOLC)ocm64442944
035 $(OCOLC)64442944
040 $aDLC$cDLC$dYDX$dBAKER$dOCO$dNPL$dBTCTA$dYDXCP$dDLC
042 $apcc
050 00 $aPN1992.77.I87$bM46 2006
082 00 $a791.45/72$d222
100 1 $aMendelson, Lee.
245 10 $aIt's the Great Pumpkin, Charlie Brown :$bthe making of a television classic /$cby Lee Mendelson, with reflectio
ns by Bill Melendez.
250 $a1st ed., special 40th anniversary ed.
260 $aNNew York :$bHarper,$cc2006.
300 $aviii, 148 p. :$bill. (chiefly col.) ;$c23 cm.
630 00 $aIt's the Great Pumpkin, Charlie Brown (Television program)
906 $a7$bcbcf$corignew$d1$eocipf$20$gy-gencatlg
925 0 $aacquire$b2 shelf copies$xpolicy default
955 $apc14 2006-02-22 to ASCD$c1g31 2006-02-27$d1g22 2006-02-27$e1g22 2006-02-27 to Dewey$aaa05 2006-02-28$aps13 2007
-01-30 2 copies rec'd., to CIP ver.$fpv21 2007-01-31 CIP ver. two copies to BCCD
```

## BIBFRAME (Turtle)

```
@prefix bf: <http://id.loc.gov/ontologies/bibframe/> .
@prefix bflc: <http://id.loc.gov/ontologies/bflc/> .
@prefix madsrdf: <http://www.loc.gov/mads/rdf/v1#> .
@prefix rdf: <http://www.w3.org/1999/02/22-rdf-syntax-ns#> .
@prefix rdfs: <http://www.w3.org/2000/01/rdf-schema#> .
@prefix xml: <http://www.w3.org/XML/1998/namespace> .
@prefix xsd: <http://www.w3.org/2001/XMLSchema#> .
@prefix zs: <http://docs.oasis-open.org/ns/search-ws/sruResponse> .

<http://bibframe.example.org/14273006#Item050-15> a bf:Item ;
  bf:itemOf <http://bibframe.example.org/14273006#Instance> ;
  bf:shelfMark [ a bf:ShelfMarkLcc ;
    rdfs:label "PN1992.77.I87 M46 2006" ;
    bf:source <http://id.loc.gov/vocabulary/organizations/dlc> ] .

<http://bibframe.example.org/14273006#Work> a bf:Text,
  bf:Work ;
  rdfs:label "It's the Great Pumpkin, Charlie Brown ." ;
  bf:adminMetadata [ a bf:AdminMetadata ;
    bflc:encodingLevel <http://id.loc.gov/vocabulary/menc1v1/4> ;
    bf:changeDate "2007-01-31T20:00:14"^^xsd:dateTime ;
    bf:creationDate "2006-02-22"^^xsd:date ;
    bf:descriptionAuthentication <http://id.loc.gov/vocabulary/marcauthen/pcc> ;
    bf:descriptionConventions <http://id.loc.gov/vocabulary/descriptionConventions/aacr> ;
    bf:descriptionModifier [ a bf:Agent ;
      rdfs:label "BAVEN" ] .
```

<http://id.loc.gov/tools/bibframe/compare-id/full-ttl?find=14273006>

# BIBFRAME

Pilot Phase 1 (September 2015-March 2016)

Approx. 40 LC Catalogers participated

Pilot Phase 2 (June 2017- )

Approx. 60 LC Catalogers participating at outset

60 new participants joined in June 2019, bringing number of catalogers participating to over 100

LC BIBFRAME catalog data: <http://id.loc.gov/resources/works.html> (Works)

<http://id.loc.gov/resources/instances.html> (Instances)

# LD4P Grants

- Linked Data for Production Phase 1 (LD4P) - 2016-2018
- Linked Data for Production Phase 2 (LD4P2) - 2018-2020



# Ontology Extensions to BIBFRAME

- Performed music: Performed Music Ontology (PMO)
- Cartographic and geospatial: Cartographic Extension
- Moving images: Moving Image Extension
- Art: ArtFrame
- Rare materials: Rare Materials Ontology Extension
  - ArtFrame & Rare Materials Ontology Extension combined: Art and Rare Materials (ARM)



# Linked Data for Production (LD4P)

1st phase (July 2016-June 2018)

- developing standards, guidelines, and infrastructure to communally produce metadata as linked open data,
- developing end-to-end workflows to create linked data in a technical services production environment,
- extending the BIBFRAME ontology to describe library resources in specialized domains and formats, and
- engaging the broader library community to ensure a sustainable and extensible environment.

# Linked Data for Production: Pathway to Implementation

- 2nd phase of Mellon grant-funded project, led by Stanford
- 4 Partner Institutions: Stanford, Harvard, Cornell, Iowa
- 17 cohort academic libraries awarded subgrant
  - University of Alberta
  - Harry Ransom Center, University of Texas at Austin
  - University of Colorado, Boulder
  - University of Chicago
  - Duke University
  - Frick Collection
  - University of Michigan
  - University of Minnesota
  - National Library of Medicine
  - Northwestern University
  - University of Pennsylvania
  - Princeton University
  - Texas A&M University
  - University of California, Davis
  - University of California, San Diego
  - University of Washington
  - Yale University

# Linked Data for Production: Pathway to Implementation

- Work areas:
  - Cloud-based environment for original metadata creation
  - Metadata reuse
  - Linkage to external authorities and web content data
  - Discovery
  - Production workflows for native linked data descriptions
  - Community collaboration

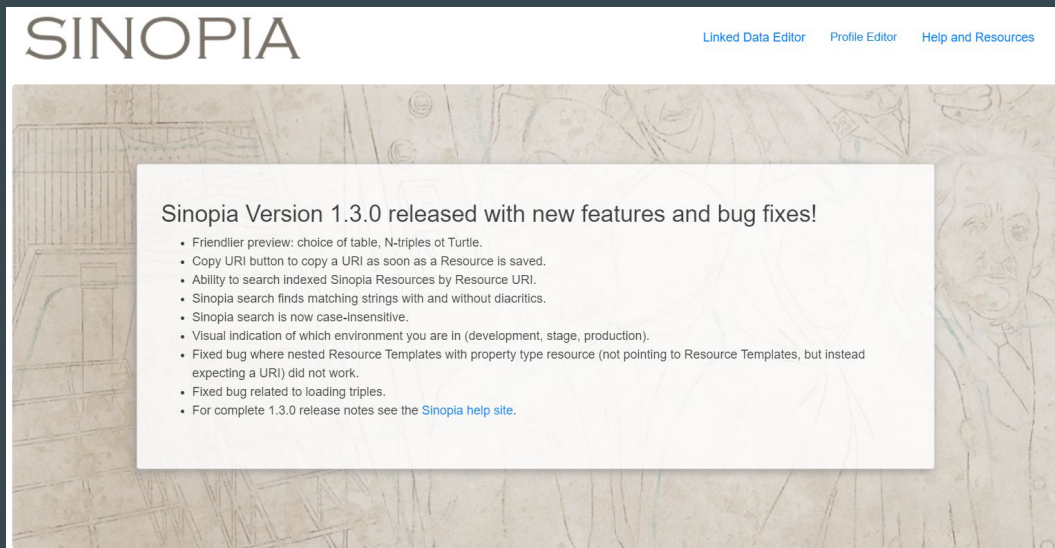
# Sinopia

- Based on Library of Congress's BIBFRAME Editor and BIBFRAME Profile Editor
- Run on Amazon Web Services by Stanford with Cornell's Questioning Authority search server
- Minimal viable product was released August 2019

SINOPIA

# Sinopia

- Development: <https://development.sinopia.io/>
- Stage: <https://stage.sinopia.io/>
- Production: <https://sinopia.io/>



The screenshot shows the Sinopia website header with the title "SINOPIA" and navigation links for "Linked Data Editor", "Profile Editor", and "Help and Resources". The main content area features a background image of a parchment document with a portrait of a man on the right. A white text box in the center contains the following text:

**Sinopia Version 1.3.0 released with new features and bug fixes!**

- Friendlier preview: choice of table, N-triples or Turtle.
- Copy URI button to copy a URI as soon as a Resource is saved.
- Ability to search indexed Sinopia Resources by Resource URI.
- Sinopia search finds matching strings with and without diacritics.
- Sinopia search is now case-insensitive.
- Visual indication of which environment you are in (development, stage, production).
- Fixed bug where nested Resource Templates with property type resource (not pointing to Resource Templates, but instead expecting a URI) did not work.
- Fixed bug related to loading triples.
- For complete 1.3.0 release notes see the [Sinopia help site](#).

# Yale's LD4P Project

## Goals:

1. Lay the foundation for local cataloging workflows using linked data standards
2. Create and enhance descriptive metadata using BIBFRAME and the Art and Rare Materials ontology
3. Explore the search, retrieval, and data visualization implications of linked data

# Project Scope

- Metadata production
  - Create original and complex copy cataloging using BIBFRAME and Sinopia for “medium rare” materials (20% cataloging time)
  - Collaborate with Black Bibliography Project on ontology development and metadata enhancement of existing JWJ catalog records (<https://blackbibliog.org/>)
- Install metaphactory software to explore data search and retrieval
- Participate in next phase of SHARE Virtual Discovery Environment, developed by Casalini Libri and @Cult
- Build on relationships in rare books community through collaboration and reporting on project outcomes

# Original Timeline

- October 2018-April 2019: Planning
- April 2019-May 2019: Training
- April/May 2019: Data entry begins
- April 2020: Data entry wraps up
- April 2020-June 2020: Project evaluated and outcomes reported
- June 30, 2020: Project end date

# Revised Timeline

- October 2018-August 2019: Planning
- August 2019-September 2019: Profile creation
- October 2019-November 2019: Training in Sinopia
- November 2019: Data entry begins
- May 2020: Data entry wraps up
- May 2020-June 2020: Project evaluated and outcomes reported
- June 30, 2020: Project end date

# Project Success Criteria

- Document workflows to use as future templates
- Create substantial metadata using BIBFRAME and Art and Rare Materials (ARM) ontology extension, and contribute it to LD4P cloud environment
- Catalog a representative subset of “medium rare” materials from collections identified as project priority in YUL and Beinecke Library backlog collections
- Install metaphactory and successfully develop a view for SPARQL-based querying
- Convert Yale catalog data as part of SHARE-VDE project
- Contribute to development of a rare materials BIBFRAME/ARM profile for use in Sinopia
- Measure staff self-assessed knowledge and comfort with linked data before, during, and after project

# Sinopa Demo

# LD4 Conference

2020 LD4 Conference: Texas A&M University, May 13-14, 2020; <https://bit.ly/32iYEFK>

2019 LD4 Conference: <https://bit.ly/2J6z2pm>

Questions?

[audrey.pearson@yale.edu](mailto:audrey.pearson@yale.edu)