

Enhancing Discoverability of Research: The Latest in Metadata and Linked Data Technologies



Gaetan Boisson, independent consultant





www.fromoldbooks.org



FDX

10:32p 22- 48

25 JUN 2013

GOSHEN PUBLIC LIBRARY

08:32am

Dial Pac

Welcome to the Online Public Access Catalog!
Please select one of the options below.

1. TITLE Keywords
2. Exact TITLE
3. AUTHOR Browse
4. SUBJECT Keywords
5. SERIES
6. SUPER Search
7. Newspaper Keyword Search
8. Newspaper Subject Search
9. Best Sellers and Award Books
10. Additional Searches
11. Review Patron Record
12. Logoff

Enter your selection(s) and press <Enter> :

S=Shortcut on, BB=Bulletin Board, ?=Help



New models for medatata



What we know all too well:

MARC21

One bibliographic record, with a number of items and links to authority records



BibLibre

Gaetan Boisson
Enhancing Discoverability of Research:
The Latest in Metadata and Linked Data Technologies

FRBR

(Functional Requirements for
Bibliographic Records)

RFAD

(for Authority Data)

RFSAD

(for Subject Authority Data)



Work

Expression

Manifestation

Item



RDA

(Resource Description and Access)

New cataloguing rules and fields, in
MARC21

336 RDA content types

337 RDA media types

338 RDA carrier types

...



LRM

(Library Reference Model)

Consolidating what was made for FRBR,
FRAD and FRSAD



Table 4.8 Relationships by Entity functioning as the Domain

Rel ID	ID of Domain	Domain	Relationship name	Inverse name	ID of Range	Range
LRM-R13	LRM-E1	Res	has appellation	is appellation of	LRM-E9	Nomen
LRM-R33	LRM-E1	Res	has association with	is associated with	LRM-E10	Place
LRM-R35	LRM-E1	Res	has association with	is associated with	LRM-E11	Time-span
LRM-R1	LRM-E1	Res	is associated with	is associated with	LRM-E1	Res
LRM-R12i	LRM-E1	Res	is subject of	has as subject	LRM-E2	Work
LRM-R20	LRM-E2	Work	accompanies / complements	is accompanied / complemented by	LRM-E2	Work
LRM-R12	LRM-E2	Work	has as subject	is subject of	LRM-E1	Res
LRM-R18	LRM-E2	Work	has part	is part of	LRM-E2	Work
LRM-R22	LRM-E2	Work	is a transformation of	was transformed into	LRM-E2	Work
LRM-R20i	LRM-E2	Work	is accompanied / complemented by	accompanies / complements	LRM-E2	Work
LRM-R21	LRM-E2	Work	is inspiration for	is inspired by	LRM-E2	Work
LRM-R21i	LRM-E2	Work	is inspired by	is inspiration for	LRM-E2	Work
LRM-R18i	LRM-E2	Work	is part of	has part	LRM-E2	Work
LRM-R2	LRM-E2	Work	is realized through	realizes	LRM-E3	Expression
LRM-R19	LRM-E2	Work	precedes	succeeds	LRM-E2	Work
LRM-R19i	LRM-E2	Work	succeeds	precedes	LRM-E2	Work
LRM-R5	LRM-E2	Work	was created by	created	LRM-E6	Agent
LRM-R22i	LRM-E2	Work	was transformed into	is a transformation of	LRM-E2	Work

This is all very new and exciting...



FDX

10:32p 22- 48

25 JUN 2013

GOSHEN PUBLIC LIBRARY

08:32am

Dial Pac

Welcome to the Online Public Access Catalog!
Please select one of the options below.

1. TITLE Keywords
2. Exact TITLE
3. AUTHOR Browse
4. SUBJECT Keywords
5. SERIES
6. SUPER Search
7. Newspaper Keyword Search
8. Newspaper Subject Search
9. Best Sellers and Award Books
10. Additional Searches
11. Review Patron Record
12. Logoff

Enter your selection(s) and press <Enter> :

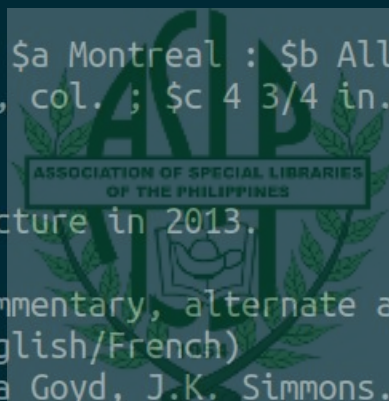
S=Shortcut on, BB=Bulletin Board, ?=Help



```

02505ngm a2200601 a 4500
001 1536863
007 vd cvaizq
008 130522s2013 onc097 e vleng d
020 $a (v.r.)
024 1 $a 065935593638
028 42 $a 593638 $b Entertainment One
035 $a 2477057
040 $a CaOWTO $c CaOWTO
041 0 $a eng $a fre $j eng $j fre $j spa
082 04 $a FIC
090 $a FIC Dark
245 00 $a Dark skies $h [DVD] = $b Ciel obscur / $c a Blumhouse and Robotproof production ; produced by Jason
Stewart. --
246 31 $a Ciel obscur $h [DVD]
260 $a [Toronto] : $b Entertainment One ; $a Montreal : $b Alliance Vivafilm, $c c2013.
300 $a 1 videodisc (ca. 97 min.) : $b sd., col. ; $c 4 3/4 in.
500 $a DVD.
500 $a Disc label title.
500 $a Originally released as a motion picture in 2013.
500 $a Widescreen version.
500 $a Special features include: audio commentary, alternate and deleted scenes with optional commentary.
500 $a Bilingual packaging (reversible English/French)
511 1 $a Keri Russell, Josh Hamilton, Dakota Goyd, J.K. Simmons.
520 $a "As the Barrett family's peaceful suburban life is rocked by an escalating series of disturbing even
ing and deadly force is after them"--www.imdb.com.
521 8 $a CHV rating: 14A.
521 8 $a Classement CV: 14A.
538 $a DVD, Dolby 5.1 surround.
540 $a For home use only.
546 $a English and French language tracks, with English, French and Spanish subtitles.
546 $a SDH (Subtitled for the deaf and hard-of-hearing)
650 0 $a Extraterrestrial beings $v Fiction. $9 48862
655 7 $a Horror films. $2 lcgft. $9 190275

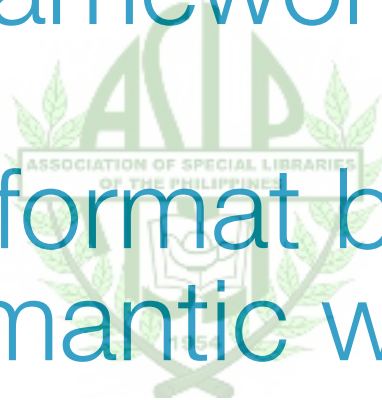
```



BIBFRAME

(Bibliographic Framework initiative)

New cataloging format built for linked data and the semantic web



BIBFRAME

 Print
  Subscribe
  Share/Save
  Give Feedback

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 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.

MARC Bibliographic Conversion Specifications

- > [Fields 001-005, 007 – Control, physical description](#) (Excel, 37 KB, 02/08/2019)
- > [Fields 006, 008 - Fixed fields](#) (Excel, 32 KB, 02/08/2019)
- > [Fields 010-048 – Identifiers, etc.](#) (Excel, 22 KB, 02/08/2019)
- > [Field 048 - Instrument codes](#) (Excel, 14 KB, 02/08/2019)
- > [Fields 050-088 - Class/call numbers, etc.](#) (Excel, 15 KB, 02/08/2019)
- > [Fields 1XX, 6XX, 7XX, 8XX; X00, X10, X11 – Names](#) (Excel, 13 KB, 02/08/2019)
- > [Fields 200-24X, except 240 – Titles](#) (Excel, 14KB, 02/08/2019)
- > [Fields 240, X30, etc. – Uniform titles](#) (Excel, 12 KB, 02/08/2019)
- > [Fields 250-270 – Edition, imprint, etc.](#) (Excel, 15 KB, 02/08/2019)
- > [Fields 3XX – Physical description, etc.](#) (Excel, 24 KB, 02/08/2019)
- > [Fields 490, 510, 530-535 – Links](#) (Excel, 14 KB, 02/08/2019)
- > [Fields 5XX – Notes](#) (Excel, 65 KB, 02/08/2019)
- > [Fields 647-662 – Subjects](#) (Excel, 15 KB, 02/01/2019)
- > [Fields 720, 740-758 – Added entries](#) (Excel, 13 KB, 02/08/2019)
- > [Fields 760-788 - Linking entries](#) (Excel, 14 KB, 02/08/2019)
- > [Fields 841-887 – Holdings, location, alternate graphics, etc.](#)

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 Work, I for Instance, but Item for Item. The Specifications use the vocabularies that are in <http://id.loc.gov> in a number of cases, especially for names, subjects, genre, and 007 and 008 data.

The Specifications also use properties with strings for data that will be used when trying to match and merge data to change the data to conform to the BIBFRAME Work and Instance models and to identify labels and RWOs in the MADS name and subject files. These local elements from

BIBFRAME 2.0 Monograph Work ⓘ

Creator of Work

Primary Contribution

Austen, Jane, 1775-1817, author. ✎ 🗑

Title Information

Work Title

Work Title Variation

Transliterated Title

Pride and prejudice ✎ 🗑

Pride and prejudice ✎ 🗑

Pride and prejudice ✎ 🗑

Form of Work

Form/Genre

Fiction. ✎ 🗑

Novels. ✎ 🗑

Love stories. ✎ 🗑

Domestic fiction. ✎ 🗑

bibliographies ✎ 🗑

Date of Work

+

Place of Origin of the Work

Place Associated with a Work

(Geographic) Coverage of the Content

Geographic coverage

(Time) Coverage of the Content

+

Intended Audience

Intended Audience

Contribution

Contribution

Rogers, Pat, 1938-, editor. ✎ 🗑

Austen, Jane., Contributor ✎ 🗑

Austen, Jane, 1775-1817, Contributor ✎ 🗑

Subject of the Work

Subject components

Young women. ✎ 🗑

English. ✎ 🗑

Sisters. ✎ 🗑

Roman. ✎ 🗑

Social classes. ✎ 🗑

Courtship. ✎ 🗑

Young women--Fiction. ✎ 🗑

Courtship--Fiction. ✎ 🗑

Social classes--Fiction. ✎ 🗑

Sisters--Fiction. ✎ 🗑



What is linked data?



You need to understand linked data and RDF, because it is the basis for BIBFRAME.



Linked data

A W3C initiative.

The web we know is made of linked documents.

Linked data defines standards to create a web of interlinked data : the semantic web.



RDF

(Resource Description Framework)

A W3C standard.

Defines how we will build linked data,
through the use of RDF triples.



RDF triples

Subject Predicate Object .

José Rizal was born in Calamba .



RDF triples

Subject Predicate Object .

José Rizal was born in Calamba .



A group of RDF triples creates a graph

José Rizal was born in Calamba .

Calamba was founded on 28 August 1742 .



This is getting interesting

José Rizal was born in Calamba .

Calamba was founded on 28 August 1742 .

José Rizal was born on 19 June 1861 .



José Rizal

was born in

Calamba

was born on

was founded on

"19 June 1861"

"28 August 1742"



This all seems very flexible... How can computers possibly be good at this?

Limited and precisely defined vocabularies

Unique identifiers for all the entities used



Unique Resource Identifiers (URI) can be used in our triples

wd:Q1500 wdt:P19 wd:Q75978 .

In Wikidata:

Q1500 is José Rizal

Q75978 is Calamba

P19 is the « born in » property





SPARQL is the language used to query linked data

```
SELECT ?subject ?predicate ?object  
WHERE {  
  ?subject ?predicate ?object .  
}
```




We make the query by using URIs in the WHERE clause

```
SELECT ?subject ?predicate ?object
WHERE {
  ?subject was born in Calamba .
}
```



We make the query by using URIs in the WHERE clause

```
SELECT ?subject
WHERE {
  ?subject wdt:P19 wd:Q75978 .
}
```



In MARC records too, we use unique identifiers for authorities!

But we don't have a standard way of communicating and reusing them.

And we duplicate data in a lot of places in our records.



RDF is not a library standard but the standard for linked data

Libraries can use this standard along with a defined vocabulary designed for them, this is what BIBFRAME aims to be.



|| BIBFRAME 2 List View || [BIBFRAME 2 Category View](#) || [BIBFRAME LC Extension List View](#) ||

This single-page view presents the entire vocabulary on one page with an ordered list of classes and properties, followed by more detailed information panels further down the document. Links are internal to the page, and the browser back button can reverse navigation on the page.

Subclasses and subproperties of classes and properties are not included in the Class List and the Property List, but are listed at the panel for the class or property to which they relate. A class or property with subclasses is marked in the list with a "+" before the class or property name (and a separate list of these is also made in a note below each list).

Changes to a vocabulary element (name, expected value, or definition) are noted in the panel for the element. Classes and properties that have been deprecated from the vocabulary are listed after the panels [here](#), and may be reviewed in the Vocabulary archive at <https://github.com/lcnetdev/bibframe-ontology>.

To view the vocabulary in rdf, use [bibframe.rdf](#).



Namespace

<http://id.loc.gov/ontologies/bibframe/>

Version

2.0.1

Version IRI

<http://id.loc.gov/ontologies/bibframe-2-0-1/>

Prior Version IRI

<http://id.loc.gov/ontologies/bibframe-2-0-0/>

Class List

AbbreviatedTitle	Generation	PlaybackCharacteristic
AccessPolicy	GenerationProcess	PlayingSpeed
AcquisitionSource	GenreForm	Polarity
+ AdminMetadata	GeographicCoverage	PostalRegistration
+ Agent	GrooveCharacteristic	PresentationFormat
Ansi	Gtin14Number	Print
AppliedMaterial	Hdl	Production
Archival	+ Identifier	ProductionMethod



Well when are we starting to use BIBFRAME?

Who knows...

But we can already leverage linked data!



Well when are we starting to use BIBFRAME?

Who knows...

But we can already leverage linked data!



Application examples





- Main page
- Community portal
- Project chat
- Create a new Item
- Create a new Lexeme
- Recent changes
- Random item
- Query Service
- Nearby
- Help
- Donate
- Tools
- What links here
- Related changes
- Special pages
- Permanent link
- Page information
- Concept URI
- Cite this page

Item **Discussion**

Read **View history**



Wiki Loves Earth photo contest: Upload photos of natural heritage sites in the Philippines to help Wikipedia and win fantastic prizes!

José Rizal (Q1500)

Filipino writer, ophthalmologist, polyglot and nationalist

edit

Jose Rizal | José Protasio Rizal Mercado y Alonso Realonda | Jose Protasio Rizal Mercado y Alonso Realonda

▼ In more languages [Configure](#)

Language	Label	Description	Also known as
English	José Rizal	Filipino writer, ophthalmologist, polyglot and nationalist	Jose Rizal José Protasio Rizal Mercado ... Jose Protasio Rizal Mercado ...
French	José Rizal	poète, romancier, et artiste philippin	José Protacio Rizal Mercado ...
Tagalog	José Rizal	Pilipinong manunulat, doktor at dalubwika na itinuturing pambansang bayani ng Pilipinas	Pepe Dimasalang Laong Laan
Spanish	José Rizal	escritor, médico y héroe nacional filipino	Rizal Rizal, José Jose Rizal

[All entered languages](#)

Statements

instance of



human



edit

► 3 references



+ add statement

Identifiers

SELIBR ID	<div><div>📏</div><div>197695</div><div>▼ 0 references</div></div> <div><div>✎ edit</div><div>+ add reference</div><div>+ add value</div></div>
NLA (Australia) ID	<div><div>📏</div><div>35454499</div><div>▼ 0 references</div></div> <div><div>✎ edit</div><div>+ add reference</div><div>+ add value</div></div>
VIAF ID	<div><div>📏</div><div>41845763</div><div>▶ 2 references</div></div> <div><div>✎ edit</div><div>+ add value</div></div>
SUDOC authorities ID	<div><div>📏</div><div>027102572</div><div>▼ 0 references</div></div> <div><div>✎ edit</div><div>+ add reference</div><div>+ add value</div></div>
Library of Congress authority ID	<div><div>📏</div><div>n80051794</div><div>▶ 2 references</div></div> <div><div>✎ edit</div></div>





Douglas Adams

Adaptateur, Scénariste, Auteur du texte, Auteur adapté

♥ Vie

1946-2010

📖 Biographie

"Douglas Noel Adams (11 March 1952 – 11 May 2001) was an English author, scriptwriter, essayist, humorist, satirist and dramatist. Adams was author of The Hitchhiker's Guide to the Galaxy, which originated in 1978 as a BBC radio comedy before developing into a 'trilogy' of five books that sold more than 15 million copies in his lifetime and generated a television series, several stage plays, comics, a computer game, and in 2005 a feature film. Adams's contribution to UK radio is commemorated in The Radio Academy's Hall of Fame.[1] Adams also wrote Dirk Gently's Holistic Detective Agency (1987) and The Long Dark Tea-Time of the Soul (1988), and co-wrote The Meaning of Liff (1983), The Deeper Meaning of Liff (1990), Last Chance to See (1990), and three stories for the television series Doctor Who; he also served as script editor for the show's seventeenth season in 1979. A posthumous collection of his works, including an unfinished novel, was published as The Salmon of Doubt in 2002. Adams was an advocate for environmentalism and conservation, a lover of fast cars,[2] technological innovation and the Apple Macintosh, and a radical atheist.

📌 Thèmes

Science-fiction

Comédie

Satire

👤 A été inspiré par

👤 A travaillé avec

[Jean Bonnefoy](#), [Jean Rosenthal](#), [Tony Akins](#)

📁 Ressources web

[BnF](#)

[Wikidata](#)

📍 Lieux associés

Gaetan Boisson

Enhancing Discoverability of Research:
The Latest in Metadata and Linked Data Technologies



ByWater
SOLUTIONS

BibLibre