

# Will you be my bf: forever?



Analysis Techniques for Conversion to BIBFRAME at the University of Alberta

Ian Bigelow, Sharon Farnel and Danoosh Davoodi

# Setting the stage:

Assessing bf: with intent to implement

- How well does bf: transition our data?
- Which flavor of bf: will serve us best?
- How much should be invested in MARC enrichment/development?
- How can we make bf: data discoverable?
- What could workflow look like for bf: pilot/implementation?



# BIBFRAME

## Casalini & @Cult

# BIBFRAME

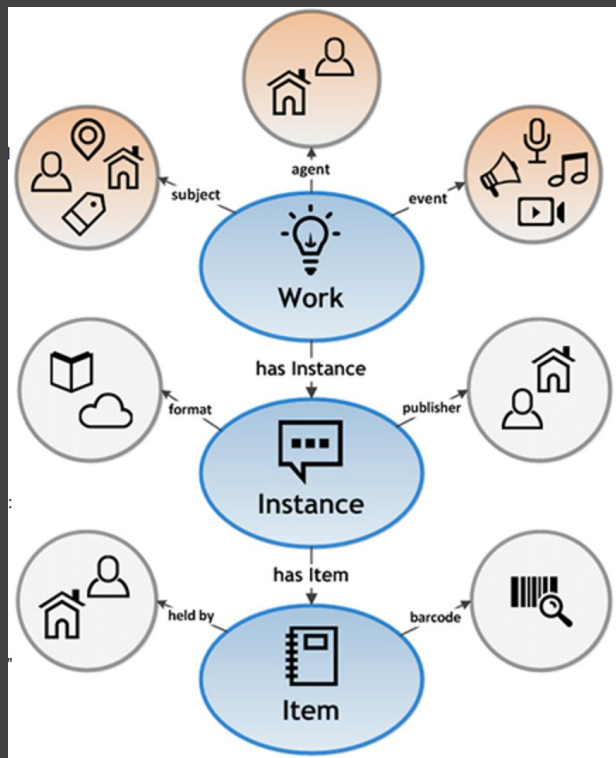
## Library of Congress



UNIVERSITY OF ALBERTA  
LIBRARIES

# Overview of the LC BIBFRAME Converter

- Initiative of LC and the community to provide an option for future bibliographic description on and of the web
- Bibframe.org
- Bf:2.0 XSLT conversion tool released in March 2017
- Available on GitHub:  
<https://github.com/lcnetdev/marc2bibframe2>



```

URTest.xml x
marc:collection
1 <?xml version="1.0" encoding="UTF-8" ?><marc:collection
1 xmlns:marc="http://www.loc.gov/MARC21/slim"
1 xmlns:xsi="http://www.w3.org/2001/XMLSchema-instance"
1 xsi:schemaLocation="http://www.loc.gov/MARC21/slim
1 http://www.loc.gov/standards/marcxml/schema/MARC21slim.x
1 sd">
2 <marc:record><marc:leader>03874nam a2200505 i
2 4500</marc:leader>
3 <marc:controlfield tag="001">6742015</marc:controlfield>
4 <marc:controlfield
4 tag="005">20141120143106.0</marc:controlfield><marc:cont
4 rolfield tag="006">m eo d
4 </marc:controlfield><marc:controlfield tag="007">cr cn
4 |||m|||a</marc:controlfield><marc:controlfield
4 tag="008">141119s2015 caua foab 000 0 eng
4 d</marc:controlfield><marc:datafield tag="020" ind1=" "
4 ind2=" "><marc:subfield
4 code="a">9781627055796</marc:subfield><marc:subfield
4 code="q">ebook</marc:subfield></marc:datafield><marc:dat
4 afield tag="020" ind1=" " ind2=" "><marc:subfield
4 code="z">9781627055789</marc:subfield><marc:subfield
4 code="q">print</marc:subfield></marc:datafield><marc:dat
4 afield tag="024" ind1="7" ind2=" "><marc:subfield
4 code="a">10.2200/800602ED1V01Y201410ICR035</marc:subfiel
4 d><marc:subfield
4 code="2">doi</marc:subfield></marc:datafield><marc:dataf
4 field tag="035" ind1=" " ind2=" "><marc:subfield

```

```

marc2bibframe2.xsl x
xsl:stylesheet
1 <?xml version='1.0'?>
2 <xsl:stylesheet version="1.0"
3 xmlns:rdf="http://www.w3.org/1999/02/22-rc
4 xmlns:rdfs="http://www.w3.org/2000/01/rdf-
5 xmlns:marc="http://www.loc.gov/MARC21/slim
6 xmlns:bf="http://id.loc.gov/ontologies/bif
7 xmlns:bfloc="http://id.loc.gov/ontologies/b
8 xmlns:madsrdf="http://www.loc.gov/mads/rdf
9 xmlns:xsl="http://www.w3.org/1999/XSLTran
10 exclude-result-prefixes="xsl marc">
11
12 <xsl:output encoding="UTF-8" method="xml" indent="yes"/>
13 <xsl:strip-space elements="*" />
14
15 <xsl:param name="baseuri" select="'http://example.org/'" />
16 <xsl:param name="idfield" select="'001'" />
17 <xsl:param name="idsource" />
18 <!--
19 to run test of idsource, comment out previous line,
20 line and uncomment the test in test/ConvSpec-001-00
21 -->
22 <!-- <xsl:param name="idsource" select="'http://id.loc.g
23 <xsl:param name="serialization" select="'rdfxml'" />
24
25 <xsl:include href="utils.xsl" />
26 <xsl:include href="ConvSpec-ControlSubfields.xsl" />

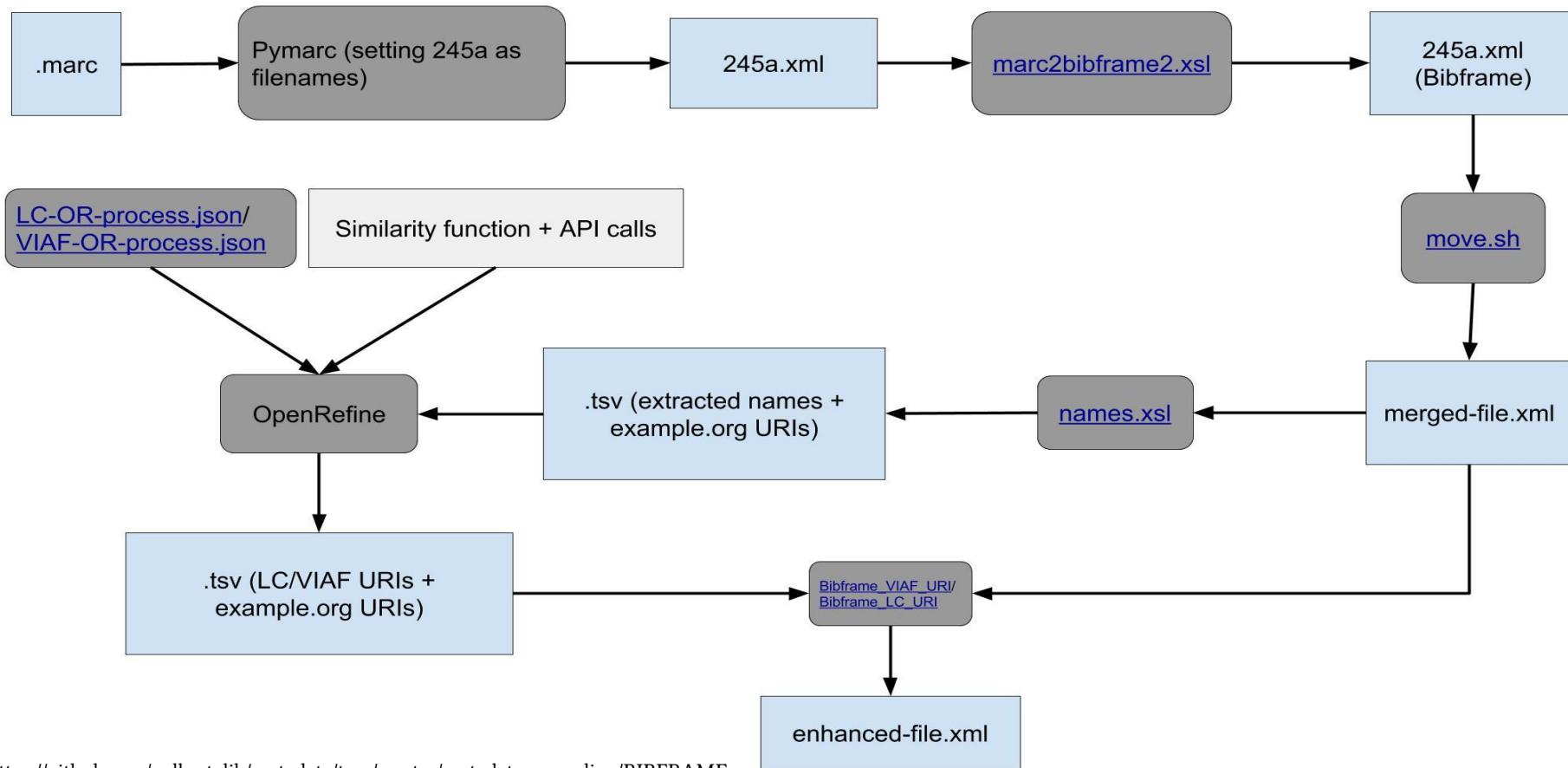
```

```

64664 </bf:shelfMark>
64663 <bf:itemOf rdf:resource="http://example.org/7797
64664 </bf:Item>
64665 </bf:hasItem>
64666 <bf:hasItem>
64667 <bf:Item rdf:about="http://example.org/7797870#Item
64668 <bf:electronicLocator>
64669 <rdfs:Resource>
64670 <bfloc:locator rdf:resource="http://ebookce
64671 <bf:note>
64672 <bf:Note>
64673 <rdfs:label>University of Alberta Ac
64674 </bf:Note>
64675 </bf:note>
64676 <bf:note>
64677 <bf:Note>
64678 <rdfs:label>from Ebook Central</rdfs
64679 </bf:Note>
64680 </bf:note>
64681 </rdfs:Resource>
64682 </bf:electronicLocator>
64683 <bf:itemOf rdf:resource="http://example.org/7797
64684 </bf:Item>
64685 </bf:hasItem>
64686 </bf:Instance>
64687 </rdf:RDF>
64688

```

# Lots of data in BIBFRAME 2



```

<xsl:stylesheet xmlns:xsl="http://www.w3.org/1999/XSL/Transform"
  xmlns:xs="http://www.w3.org/2001/XMLSchema"
  xmlns:rdf="http://www.w3.org/1999/02/22-rdf-syntax-ns#"
  xmlns:rdfs="http://www.w3.org/2000/01/rdf-schema#"
  xmlns:bf="http://id.loc.gov/ontologies/bibframe/"
  xmlns:bflc="http://id.loc.gov/ontologies/bflc/" xmlns:mdsrdf="http://www.loc.gov/mads/rdf/v1#"
  exclude-result-prefixes="xs xsl rdf rdfs bf bflc madsrdf" version="2.0">

  <xsl:output indent="yes" media-type="xml" omit-xml-declaration="yes"/>
  <xsl:strip-space elements="*/>

  <xsl:variable name="doc" select="collection(' ../metadata-wrangling/BIBFRAME/CompData/?select=2015*.xml;recurse=yes')"/>

  <xsl:template match="/">
    <xsl:for-each select="$doc/bf:Agent[@rdf:about]">
      <xsl:variable name="c"></xsl:variable>
      <xsl:variable name="name" select="replace(normalize-space(rdfs:label), $c, ' ')">
      <xsl:value-of select="$name"/>
      <xsl:variable name="name-part">
        <xsl:value-of select="normalize-space(replace(rdfs:label, ' ', '%20'))"/>
      </xsl:variable>
      <xsl:text>#09</xsl:text>
      <xsl:for-each select="rdf:type[starts-with(@rdf:resource, 'http://id.loc.gov/ontologies/bibframe/')]">
        <xsl:value-of select="normalize-space(replace(@rdf:resource, 'http://id.loc.gov/ontologies/bibframe/', ''))"/>
        <xsl:variable name="p">
          <xsl:value-of select="normalize-space(replace(@rdf:resource, 'http://id.loc.gov/ontologies/bibframe/', ''))"/>
        </xsl:variable>
        <xsl:text>#09</xsl:text>
        <xsl:choose>
          <xsl:when test="$p = 'Person'">
            <xsl:text>http://www.viaf.org/viaf/search?query=local.personalNames%20all%20%22</xsl:text><xsl:value-of sel
          </xsl:when>
          <xsl:otherwise>
            <xsl:text>http://www.viaf.org/viaf/search?query=local.corporateNames%20all%20%22</xsl:text><xsl:value-of se
          </xsl:otherwise>
        </xsl:choose>
      </xsl:for-each>
      <xsl:text>#09</xsl:text>
      <xsl:value-of select="@rdf:about"/>
      <xsl:text>#xa</xsl:text>
    </xsl:for-each>
  </xsl:template>
</xsl:stylesheet>

```

	A	B	C
1	Maseda Daigaku. Engeki Hakubutsukan.	Organization	http://example.org/1007604#Agent710-21
2	Japan Cultural Institute in Rome.	Organization	http://example.org/1007604#Agent710-22
3	Biennale di Venezia (36th : 1972)	Meeting	http://example.org/1007604#Agent711-23
4	Metropolitan Museum of Art (New York, N.Y.)	Organization	http://example.org/1084147#Agent110-10
5	Bean, Jacob.	Person	http://example.org/1084147#Agent700-19
6	National Gallery of Art (U.S.)	Organization	http://example.org/1078393#Agent110-11
7	Block, Leigh B., 1905-1987	Person	http://example.org/1078393#Agent600-20
8	Block, Leigh B., Mrs., 1904-	Person	http://example.org/1078393#Agent600-21
9	Los Angeles County Museum of Art.	Organization	http://example.org/1078393#Agent710-23
10	National Science Film Library (Canada)	Organization	http://example.org/1019628#Agent110-10
11	Davis, Judith A.,	Person	http://example.org/1019628#Agent700-21
12	Canadian Film Institute.	Organization	http://example.org/1019628#Agent710-22
13	Brunschwig, Chantal, 1943-	Person	http://example.org/1074866#Agent100-11
14	Calvet, Louis-Jean, 1942-	Person	http://example.org/1074866#Agent700-20
15	Klein, Jean-Claude, 1943-	Person	http://example.org/1074866#Agent700-21
16	Jacob, André,	Person	http://example.org/1094234#Agent100-10
17	Germany (West). Auswärtiges Amt.	Jurisdiction	http://example.org/1048801#Agent110-11
18	Germany. Auswärtiges Amt	Jurisdiction	http://example.org/1048801#Agent610-19
19	Germany (West). Auswärtiges Amt	Jurisdiction	http://example.org/1048801#Agent610-20
20	Sasse, Heinz Günther, 1906-	Person	http://example.org/1048801#Agent700-22
21	Eickhoff, Ekkehard.	Person	http://example.org/1048801#Agent700-23
22	Fischer, Irving, 1867-	Person	http://example.org/1018226#Agent100-11
23	Comstock, Helen.	Person	http://example.org/1088457#Agent100-11
24	Decker, Isabelle M.	Person	http://example.org/1067852#Agent100-12
25	Sack, Allan.	Person	http://example.org/1036859#Agent100-9
26	Youman, Jack.	Person	http://example.org/1036859#Agent700-21
27	Lochhead, Douglas,	Person	http://example.org/1064932#Agent100-22
28	Souster, Raymond, 1921-	Person	http://example.org/1064932#Agent700-36
29	Wheatley, John H.	Person	http://example.org/1061087#Agent100-9
30	Coon, Herbert L.	Person	http://example.org/1061087#Agent700-22
31	Landmark 4-H Home Ec. Club.	Organization	http://example.org/1024886#Agent710-22
32	Tyson, Geoffrey, 1898-1971.	Person	http://example.org/1017248#Agent100-10
33	National and Grindlays Bank Limited.	Organization	http://example.org/1017248#Agent610-18
34	Hillier, Bevis, 1940-	Person	http://example.org/1085410#Agent100-13
35	Mattleman, Marciene Schreiber.	Person	http://example.org/1060609#Agent100-9
36	Trese, Leo J. (Leo John), 1902-1970.	Person	http://example.org/1036570#Agent100-10
37	Stern, Renée Bernd, 1875-1940.	Person	http://example.org/1031905#Agent100-11
38	Newton, David E.	Person	http://example.org/1008728#Agent100-9

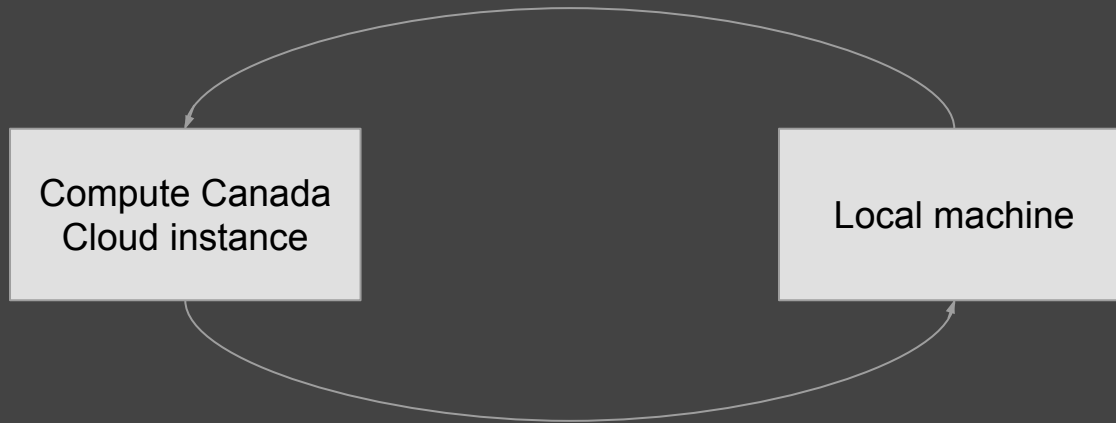
All	Names	match-ID-F	score-lc	Names1	LCP	match-ID	score	Cor	LCC	score-c	Corp/personal	Column 3	Ingest Key
11.	Conservatorio Superior de Música <input checked="" type="checkbox"/> Conservatorio Superior de Música de Sevilla (84) <input checked="" type="checkbox"/> Real Conservatorio Superior de Música de Madrid (79) <input checked="" type="checkbox"/> Conservatorio Superior de Música Santa Cruz de Tenerife (72) <input checked="" type="checkbox"/> Create new topic	n84186607	84	Conservatorio Superior de Música <input checked="" type="checkbox"/> Carmona, José Carlos, 1963- (0.188) <input checked="" type="checkbox"/> Pons, Jordi, 1960- (0.219) <input checked="" type="checkbox"/> Vidal Corella, Vicente (0.125) <input checked="" type="checkbox"/> Create new topic				Conservatorio Superior de Música <input checked="" type="checkbox"/> Conservatoire national supérieur de musique et de danse de Paris (0.406) <input checked="" type="checkbox"/> Real Escuela Superior de Arte Dramático (Spain) (0.313) <input checked="" type="checkbox"/> Real Conservatorio Superior de Música de Madrid (0.681) <input checked="" type="checkbox"/> Create new topic			Organization		http://example.org/1018
12.	Puiggrós, Rodolfo, 1906-1980 Choose new match	n88232712	98	Puiggrós, Rodolfo, 1906-1980 <input checked="" type="checkbox"/> Frank, André Gunder, 1929-2005 (0.167) <input checked="" type="checkbox"/> Puiggrós, Rodolfo, 1906-1980 (1) <input checked="" type="checkbox"/> Create new topic	n88232712	n88232712	1	Puiggrós, Rodolfo, 1906-1980 <input checked="" type="checkbox"/> Create new topic			Person		http://example.org/10411
13.	Silva Neto, Serafim da, 1917-1960 Choose new match	n80152925	100	Silva Neto, Serafim da, 1917-1960 <input checked="" type="checkbox"/> Silva Neto, Serafim * da (1917-1960) (0.941) <input checked="" type="checkbox"/> Create new topic	n80152925	n80152925	0.9411764705882353	Silva Neto, Serafim da, 1917-1960 <input checked="" type="checkbox"/> Create new topic			Person		http://example.org/10010
14.	Font y Sagué, Norbert, 1873-1910 <input checked="" type="checkbox"/> Create new topic			Font y Sagué, Norbert, 1873-1910 <input checked="" type="checkbox"/> Create new topic				Font y Sagué, Norbert, 1873-1910 <input checked="" type="checkbox"/> Create new topic			Person		http://example.org/10010
15.	Eyzaguirre, Jaime Choose new match	n80090270	100	Eyzaguirre, Jaime (0.944) <input checked="" type="checkbox"/> Barros van Buren, Mario (0.261) <input checked="" type="checkbox"/> Eyzaguirre, Jaime, 1935- (0.708) <input checked="" type="checkbox"/> Create new topic	n80090270	n80090270	0.9444444444444444	Eyzaguirre, Jaime <input checked="" type="checkbox"/> Universidad Católica de Chile. Oficina Coordinadora de Investigaciones "Profesor Jaime Eyzaguirre" (0.111) <input checked="" type="checkbox"/> Create new topic			Person		http://example.org/10012
16.	Valdeavellano, Luis G. de (Luis García de), 1904- Choose new match	n79081511	99	Valdeavellano, Luis G. de (Luis García de), 1904- Choose new match	n79081511	n79081511		Valdeavellano, Luis G. de (Luis García de), 1904- <input checked="" type="checkbox"/> Create new topic			Person		http://example.org/10011
17.	Tudebodus, Petrus <input checked="" type="checkbox"/> Tudebodus, Petrus, active 1096-1100 (65) <input checked="" type="checkbox"/> Tudebodus, Petrus, active 1096-1100 (65) <input checked="" type="checkbox"/> Create new topic			Tudebodus, Petrus <input checked="" type="checkbox"/> Tudebodus, Petrus, active 1096-1100 (0.486) <input checked="" type="checkbox"/> Create new topic				Tudebodus, Petrus <input checked="" type="checkbox"/> Create new topic			Person		http://example.org/10012
18.	Palavicini, Félix Fulgencio, 1881- Choose new match	n82032394	98	Palavicini, Félix Fulgencio, 1881- Choose new match	n82032394	n82032394	1	Palavicini, Félix Fulgencio, 1881- <input checked="" type="checkbox"/> Create new topic			Person		http://example.org/10011
19.	Mexico, Congreso Constituyente (1916-1917) Choose new match	n83120024	100	Mexico, Congreso Constituyente (1916-1917) <input checked="" type="checkbox"/> Create new topic		n83120024		Mexico, Congreso Constituyente (1916-1917) Choose new match	n83120024	1	Jurisdiction		http://example.org/10020
20.	Mexico Choose new match	n81013960	100	Mexico <input checked="" type="checkbox"/> Eisenstein, Sergei, 1898-1948 (0.067) <input checked="" type="checkbox"/> Cortés, Hernán, 1485-1547, (0.038) <input checked="" type="checkbox"/> Maximilian, Emperor of Mexico, 1832-1867 (0.15) <input checked="" type="checkbox"/> Create new topic				Mexico <input checked="" type="checkbox"/> Universidad nacional autónoma de México (0.128) <input checked="" type="checkbox"/> Instituto Nacional de Antropología e Historia (México) (0.093) <input checked="" type="checkbox"/> Time-Life books (0.2) <input checked="" type="checkbox"/> Create new topic			Jurisdiction		http://example.org/10022
21.	Maetzl, María de, 1882- <input checked="" type="checkbox"/> Maetzl Whitney, Maria de, 1882-1948 (73) <input checked="" type="checkbox"/> Maetzl Whitney, Maria de, 1882-1948. Essays. Selections (54) <input checked="" type="checkbox"/> Create new topic			Maetzl, María de, 1882- <input checked="" type="checkbox"/> Maetzl, Ramiro de, 1875-1936, (0.586) <input checked="" type="checkbox"/> Maetzl, Gustavo * de (1887-1947) (0.567) <input checked="" type="checkbox"/> Maetzl Whitney, Maria de, 1882-1948 (0.611) <input checked="" type="checkbox"/> Create new topic				Maetzl, María de, 1882- <input checked="" type="checkbox"/> Create new topic			Person		http://example.org/10011
22.	Bazant, Jan Choose new match	n85099113	100	Bazant, Jan <input checked="" type="checkbox"/> Create new topic				Bazant, Jan <input checked="" type="checkbox"/> Create new topic			Person		http://example.org/10011



# Process Times

Process	Time	Tool
Converting .marc to MARC/XML	7 - 8 mins	pymarc
Converting MARC/XML to BIBFRAME (and merging)	40 - 50 mins	Oxygen / bash
Extracting names (or subjects) from the bibframe file	Less than 2 mins	Oxygen
OpenRefine process	Few seconds	OpenRefine - GREL
Enriching names with URIs (from VIAF)	30 - 35 mins	OpenRefine + VIAF recon java client
Enriching names with URIs (from LC)	90 - 120 mins	OpenRefine + LC recon client
Enriching subjects with URIs (from LC)	70 - 90 mins	OpenRefine + LC recon client
OpenRefine process	Few seconds	OpenRefine - GREL
Ingesting (replacing example.org URIs)	60 - 70 mins (using Saxon EE on Oxygen) 100 - 120 mins (using Saxon HE on command-line)	Oxygen / Saxon command-line

# Parallel Processing



# Entity Matching

Source	Names		Subjects
	LC	VIAF	
<b>1985 Imprints</b>	92.41%	87.22%	55.98%
<b>2015 Imprints</b>	96.06%	86.33%	65.36%
<b>UA</b>	83.92%	79.84%	74.52%



# Overview of the Casalini SHARE VDE Project

An @Cult and Casalini Libri partnership

“ALIADA project, co-financed by the European Union in 2013-2015, originally applied the Linked Data paradigm using FRBRoo based ontologies.”<sup>1</sup>

“A prototype of a virtual discovery environment with a three BIBFRAME layer architecture (Person/Work, Instance, Item) has been established through the individual processes of analysis, entity identification and reconciliation, conversion and publication of data from MARC21 to RDF, within the context of libraries with different systems, habits and cataloguing traditions.”<sup>2</sup>

1. Casalini, Michele (2017). BIBFRAME and linked data practices for the stewardship of research knowledge. IFLA satellite meeting for Digital Humanities. Connecting Libraries and Research, Berlin.

2. Casalini Libri (2017). The SHARE-VDE Project. Retrieved from <http://share-vde.org/sharevde/clusters?!=en>



## Project participants:

- Stanford University
- University of California Berkeley
- Yale University
- Library of Congress
- University of Chicago
- University of Michigan Ann Arbor
- Harvard University
- Massachusetts Institute of Technology
- Duke University
- Cornell University
- Columbia University
- University of Pennsylvania
- Pennsylvania State University
- Texas A&M University
- University of Alberta / NEOS Library Consortium
- University of Toronto

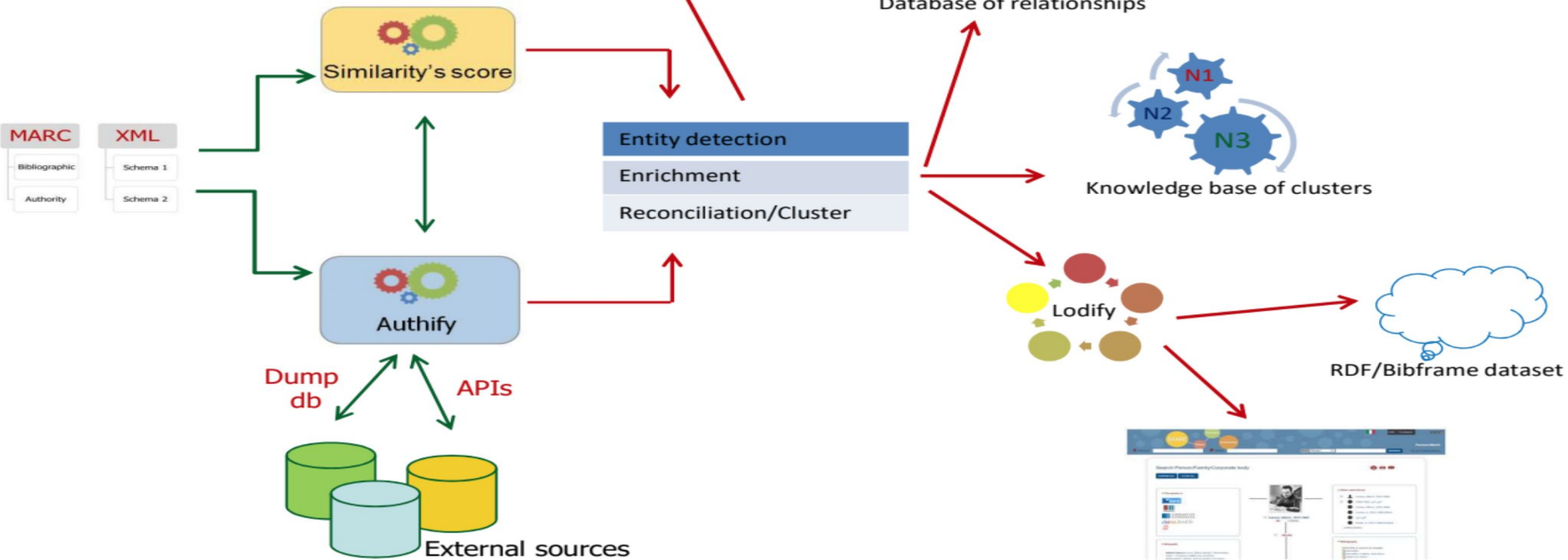
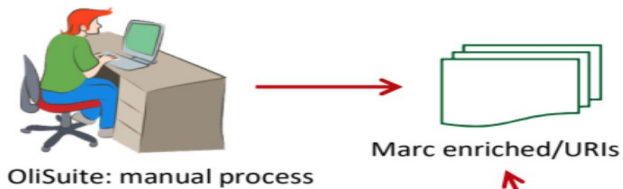
### Phase 1:

- MARC for 1985 and 2015 imprint data returned with URI enrichment
- Imprint data returned in bf:2.0
- Entity identification
- Reconciliation of data clustering
- Release of SHARE VDE with searchable imprint data
- Access to data through Blazegraph

### Phase 2:

- The catalogue of each library converted in to BIBFRAME 2.0 format. Entities are reconciled in the dataset and linked to SHARE-VDE project URIs of Deliverable 2 for identification.
- The SHARE-VDE project knowledge base of clusters in RDF format. D2 is common for all institutions as it includes data from all of the participants. Reconciled entities in the knowledge base are enriched with URIs from external sources. All variant forms are included coming by internal catalogues and by external sources.
- The dataset converted in BIBFRAME 2.0 with external URIs included. This dataset is produced in order to work autonomously from the project knowledge base D2, because it includes a certain number of relationships already present in the knowledge base. This makes the D3 'auto-consistent' and usable autonomously.
- The MARC21 version of D3. It includes all of the institution's records enriched with URIs.

# The SHARE-VDE processes overview



SHARE-VDE Portal

SHARE

Virtual

Discovery

Environment

## Person/Work



Person



Work



Person



BROWSE

## ▼ This person in



data.bnf.fr

VT  
AF Shakespeare, William, 1564-1616

ID: 10236

 Works

## ▼ Other name forms

- Shakespeare, William, 1564-1616
- Шекспир, Уильям, 1564-1616
- 1564-1616, ויליאם, ויליאם, ויליאם
- شكسبير، وليام
- Shakespeare, W. 1564-1616  
William
- Shakespeare, William, 1564-1616

...(other forms)

## ▼ Wikipedia

**William Shakespeare** (/ˈʃeɪkspiər/; 26 April 1564 (baptised) – 23 April 1616) was an English poet, playwright, and actor, widely regarded as the greatest writer in the English language and the world's pre-eminent dramatist. He is often called England's national poet, and the "Bard of Avon". His extant works, including collaborations, consist of approximately 38 plays, 154 sonnets, two long narrative poems, and a few other verses, some of uncertain authorship. His plays have been translated into many of the world's major languages and are still performed in many theatres.

## ▼ Bibliography

(Click title to search on Google)

- "Gamlet" Boris Pasternaka
- 1 Henry IV
- 12th nite
- 3 by Shakespeare

...(other titles)

# LC Meetings to develop Phase 3



Use cases for phase 3 are still being developed, but may include :

- publish all participants data in the SHARE VDE platform
- incorporate ability to batch update the record set with library exports,
- develop ability to return enhanced data back to libraries in an automated way,
- develop ability to edit the information in ShareVDE through cataloging tools,
- develop capacity for reports of some type and
- develop original cataloging workflows for SHAREVDE





# Data Through Conversion: Analysis based on BIBCO and CONSER Core

1. An examination of Casalini and LC bf:2.0 conversions bases on BIBCO and CONSER core elements
  - a. Do conversions give adequate coverage/treatment of core elements and in what ways?
  - b. How well are monographs and serials treated?
2. Comparing 1985 and 2015 imprint data
  - a. How well does bf: convert current and legacy MARC and encoding standards?
3. Pre vs post MARC to LD conversion URI enrichment efficacy
  - a. If URI enrichment of MARC data is to be done, what areas make the most sense?



## BSR<sup>1</sup> / CSR<sup>2</sup> to BIBFRAME Mappings:

Provided helpful reference tools for analysis of RDA core elements through conversion.

As these will have seen scrutiny by PCC already and we were looking at RDA Core, perhaps it wasn't surprising that all elements were represented fairly well. Still, a few interesting findings:

### Monographs and/or some general points of interest:

#### 1. Production, Publication, Distribution, Manufacture statements:

LC XSLT: Strips brackets and other marks of punctuation in mapping to place, agent and date

SHARE VDE: Maintains brackets and other punctuation but also clusters terms and mints associated URI

<<http://share-vde.org/sharevde/rdfBibframe2/ProvisionActivity/79dcbc23-f113-3c01-9159-9e359f0c994c>>

<<http://id.loc.gov/ontologies/bibframe/date>> "1958." .

<<http://share-vde.org/sharevde/rdfBibframe2/ProvisionActivity/79dcbc23-f113-3c01-9159-9e359f0c994c>>

<<http://id.loc.gov/ontologies/bibframe/date>> "[1958]" .

<<http://share-vde.org/sharevde/rdfBibframe2/ProvisionActivity/79dcbc23-f113-3c01-9159-9e359f0c994c>>

<<http://id.loc.gov/ontologies/bibframe/date>> "1958]" .

#### 2. Preferred title:

LC XSLT: Appropriately uses 130/240 or 245\$a in absence of them to generate preferred title of work

SHARE VDE: Uses URI to pull together title data for works and instances

<http://share-vde.org/sharevde/rdfBibframe2/title>

<http://share-vde.org/sharevde/rdfBibframe2/Title>

1. BIBCO Mapping BSR to BIBFRAME 2.0 Group (2017). BSR to BIBFRAME mapping. Retrieved from: <https://www.loc.gov/aba/pcc/bibframe/TaskGroups/BSR-PDF/BSRtoBIBFRAMEMapping.pdf>
2. Conser CSR to BIBFRAME Mapping Task Group (2017). CSR to BIBFRAME mapping. Retrieved from: <https://www.loc.gov/aba/pcc/bibframe/TaskGroups/CSR-PDF/CSRtoBIBFRAMEMapping.pdf>



### 3. Creators, contributors and relators:

LC XSLT: Agent and Role

SHARE VDE:

```
<http://share-vde.org/sharevde/rdfBibframe/Agent/78151> <http://www.w3.org/1999/02/22-rdf-syntax-ns#type>
<http://id.loc.gov/ontologies/bibframe/Agent> <http://share-vde.org> .
<http://share-vde.org/sharevde/rdfBibframe/Agent/78151> <http://www.w3.org/2000/01/rdf-schema#label>
"Guillaume,approximately 1300-1377." <http://share-vde.org> .
<http://share-vde.org/sharevde/rdfBibframe/Agent/78151> <http://www.loc.gov/mads/rdf/v1#isIdentifiedByAuthority>
<http://id.loc.gov/authorities/names/n50018452> <http://share-vde.org> .
<http://share-vde.org/sharevde/rdfBibframe/Agent/78151> <http://www.w3.org/2002/07/owl#sameAs>
<http://www.wikidata.org/entity/Q200580> <http://share-vde.org> .
<http://share-vde.org/sharevde/rdfBibframe/Agent/78151> <http://www.w3.org/2002/07/owl#sameAs>
<http://viaf.org/viaf/100181685/> <http://share-vde.org> .
```

Serials:

As noted in the *Final Report of the CONSER CSR to BIBFRAME Mapping Task Group*<sup>1</sup>, Numeric and/or alphabetic designation/Chronological designation of first issue or part of sequence (RDA 2.6.2/2.6.3) both map to firstIssue (similarly for lastIssue). The mapping works correctly in both conversions, but why would the data not be made more atomic?

The report by the CONSER CSR to BIBFRAME Mapping Task Group provides other information and is a good reference point.

1. Conser CSR to BIBFRAME Mapping Task Group (2017). Final report of the CSR to BIBFRAME Mapping Task Group. Retrieved from: <https://www.loc.gov/aba/pcc/bibframe/TaskGroups/CSR-PDF/FinalReportCONSERToPCCBIBFRAMETaskGroup.pdf>



# 1985 vs 2015 Imprint Data

Following the approach of the SHARE VDE to look at 1985 and imprint data to compare past and current standards through conversion found several interesting points.

Example:

1. LC XSLT: Earlier records lacking relationship designators only have “contributor” role assigned

```
<bf:role>
```

```
<bf:Role rdf:about="http://id.loc.gov/vocabulary/relators/ctb"/>
```

```
</bf:role>
```

SHARE VDE:

Relator Term Detection: “Starting from a Marc21 record (whatever is the specific dialect) the system analyses all (configured) tags that contain a name and, for each of them, tries to figure out (using the statements of responsibility of the input record or other parts of the record) what is the corresponding role within the work represented by the given record.” (Casalini Libri, 2017)

```
<http://share-vde.org/sharevde/rdfBibframe/Agent/3354732> <http://www.w3.org/2000/01/rdf-schema#label> "Winter-Hjelm,
```

```
Otto,1837-1931." <http://share-vde.org>
```

```
<http://share-vde.org/sharevde/rdfBibframe/Work/16985470> <http://id.loc.gov/vocabulary/relators/cre>
```

```
<http://share-vde.org/sharevde/rdfBibframe/Agent/3354732> <http://share-vde.org>
```



# URI Enrichment: Before or After?

## Guidance from the PCC Task Group on URI in MARC

June 2017: ALA MAC Proposals - <https://www.loc.gov/marc/mac/list-p.html#2017>

**\$0** - URIs that identify a 'Record' or 'Authority' entity describing a Thing (e.g. madsrdf:Authorities, SKOS Concepts for terms in controlled or standard vocabulary lists)

**\$1** - URIs that directly identify a Thing itself (sometimes referred to as a Real World Object or RWO, whether actual or conceptual)

**\$4** - Redefining Subfield \$4 to Encompass URIs for Relationships in the MARC 21 Authority and Bibliographic Formats

**758** - An identifier for a resource related to the resource described in the bibliographic record. Resources thus identified may include, but are not limited to, FRBR works, expressions, manifestations, and items. The field does not prescribe a particular content standard or data model.

# Bf:2.0 LC from Enriched MARC Data

```
<bf:contribution>
  <bf:Contribution>
    <bf:agent>
      <bf:Agent rdf:about="http://id.loc.gov/authorities/names/n2005058924">
        <rdf:type rdf:resource="http://id.loc.gov/ontologies/bibframe/Person"/>
        <bflc:name00MatchKey>Sengupta, Ashis,</bflc:name00MatchKey>
        <bflc:name00MarcKey>7001 $aSengupta,
Ashis,$eeditor.$0http://id.loc.gov/authorities/names/n2005058924$0http://viaf.org/viaf/24010261</bflc:name00MarcKey>
        <rdfs:label>Sengupta, Ashis,</rdfs:label>
        <bf:identifiedBy>
          <bf:Identifier>
            <rdf:value rdf:resource="http://viaf.org/viaf/24010261"/>
          </bf:Identifier>
        </bf:identifiedBy>
      </bf:Agent>
    </bf:agent>
    <bf:role>
      <bf:Role>
        <rdfs:label>editor.</rdfs:label>
        <bflc:relatorMatchKey>editor</bflc:relatorMatchKey>
      </bf:Role>
    </bf:role>
  </bf:Contribution>
```



# Bf:2.0 LC Enriched Post Conversion

```
<bf:Contribution>
  <bf:agent>
    <bf:Agent rdf:about="http://id.loc.gov/authorities/names/n2005058924">
      <rdf:type rdf:resource="http://id.loc.gov/ontologies/bibframe/Person"/>
      <bflc:name00MatchKey>Sengupta, Ashis,</bflc:name00MatchKey>
      <bflc:name00MarcKey>7001 $aSengupta, Ashis,$editor.</bflc:name00MarcKey>
      <rdfs:label>Sengupta, Ashis,</rdfs:label>
      <bf:identifiedBy>
        <bf:Identifier>
          <rdf:value rdf:resource="http://viaf.org/viaf/24010261"/>
        </bf:Identifier>
      </bf:identifiedBy>
    </bf:Agent>
  </bf:agent>
  <bf:role>
    <bf:Role>
      <rdfs:label>editor.</rdfs:label>
      <bflc:relatorMatchKey>editor</bflc:relatorMatchKey>
    </bf:Role>
  </bf:role>
</bf:Contribution>
```



# MARC from Casalini Enriched with URI

=100 1\$aTchaikovsky, Peter  
lich,\$d1840-1893.\$1<http://share-vde.org/sharevde/rdfBibframe/Agent/25011>\$0<http://id.loc.gov/authorities/names/n79072979>\$1<http://data.bnf.fr/13900329>\$1<http://www.wikidata.org/entity/Q7315>\$1<http://viaf.org/viaf/99258155/>

=100 1\$aLeCompte, Margaret Diane,\$eauthor.\$1<http://isni.org/isni/0000000116573223>

=758 \\\$4<http://rdaregistry.info/Elements/m/P30004>\$0<http://www.worldcat.org/oclc/900194099>

=758 \\\$4<http://rdaregistry.info/Elements/m/P30139>\$1<http://worldcat.org/entity/work/id/2267517995>

=758 \\\$4<http://rdaregistry.info/Elements/m/P30135>\$iHas work manifested:  
\$1<http://worldcat.org/entity/work/id/2267517995>





# URI in Casalini bf:2.0 Data (D2)

<<http://share-vde.org/sharevde/rdfBibframe/Agent/3354732>>

<<http://www.loc.gov/mads/rdf/v1#isIdentifiedByAuthority>> <<http://id.loc.gov/authorities/names/n87114204>>

<<http://share-vde.org>> .

<<http://share-vde.org/sharevde/rdfBibframe/Agent/3354732>> <<http://www.w3.org/2002/07/owl#sameAs>>

<<http://isni.org/isni/0000000080160020>> <<http://share-vde.org>> .

<<http://share-vde.org/sharevde/rdfBibframe/Agent/3354732>> <<http://www.w3.org/2002/07/owl#sameAs>>

<<http://www.wikidata.org/entity/Q2975366>> <<http://share-vde.org>> .

<<http://share-vde.org/sharevde/rdfBibframe/Agent/3354732>> <<http://www.w3.org/2002/07/owl#sameAs>>

<<http://viaf.org/viaf/54413320/>> <<http://share-vde.org>> .





# SHARE VDE as an Identifier/Authority Resource

CODE	SOURCE	INSTITUTION	LINK TO PROJECT PAGE	URI
NAF (Names)	Library of Congress Names	Library of Congress	<a href="http://id.loc.gov/authorities/names.html">http://id.loc.gov/authorities/names.html</a>	<a href="http://id.loc.gov/authorities/names">http://id.loc.gov/authorities/names</a>
NAF (NameTitle + Titles)	Library of Congress Title / NameTitle	Library of Congress	<a href="http://id.loc.gov/authorities/names.html">http://id.loc.gov/authorities/names.html</a>	<a href="http://id.loc.gov/authorities/names">http://id.loc.gov/authorities/names</a>
LCSH	Library of Congress Subject Headings	Library of Congress	<a href="http://id.loc.gov/authorities/subjects.html">http://id.loc.gov/authorities/subjects.html</a>	<a href="http://id.loc.gov/authorities/subjects">http://id.loc.gov/authorities/subjects</a> <a href="http://id.loc.gov/authorities#conceptscheme">http://id.loc.gov/authorities#conceptscheme</a>
LC Classification	Library of Congress Classifications	Library of Congress	<a href="http://id.loc.gov/authorities/classification.html">http://id.loc.gov/authorities/classification.html</a>	<a href="http://id.loc.gov/authorities/classification">http://id.loc.gov/authorities/classification</a>
LC Identifiers	Library of Congress Standard Identifiers Scheme	Library of Congress	<a href="http://id.loc.gov/vocabulary/identifiers.html">http://id.loc.gov/vocabulary/identifiers.html</a>	<a href="http://id.loc.gov/vocabulary/identifiers">http://id.loc.gov/vocabulary/identifiers</a>
Geographic Areas	MARC List for Geographic Areas	Library of Congress	<a href="http://id.loc.gov/vocabulary/geographicAreas.html">http://id.loc.gov/vocabulary/geographicAreas.html</a>	<a href="http://id.loc.gov/vocabulary/geographicAreas">http://id.loc.gov/vocabulary/geographicAreas</a>
Relators	MARC Code List for Relators	Library of Congress	<a href="http://id.loc.gov/vocabulary/relators.html">http://id.loc.gov/vocabulary/relators.html</a>	<a href="http://id.loc.gov/vocabulary/relators">http://id.loc.gov/vocabulary/relators</a>
Languages	MARC List for Languages	Library of Congress	<a href="http://id.loc.gov/vocabulary/languages.html">http://id.loc.gov/vocabulary/languages.html</a>	<a href="http://id.loc.gov/vocabulary/languages">http://id.loc.gov/vocabulary/languages</a>
Countries	MARC List for Countries	Library of Congress	<a href="http://id.loc.gov/vocabulary/countries.html">http://id.loc.gov/vocabulary/countries.html</a>	<a href="http://id.loc.gov/vocabulary/countries">http://id.loc.gov/vocabulary/countries</a>
ISNI	ISNI - International Standard Name Identifier	OCLC	<a href="http://isni.org/search">http://isni.org/search</a>	<a href="http://isni.org/isni/[ID]">http://isni.org/isni/[ID]</a>

# Making bf: Updates Work: The importance of Work URI

How do we cluster work information through conversion and ongoing updates to BIBFRAME data?

MARC 758 was created to give a place to record primary resource identifiers, including work URI. Until we can move past the need for our MARC data, it will be important to include work URI in MARC to support ongoing conversions to bf: data.

- <http://worldcat.org/entity/work/id/>
- <http://share-vde.org/sharevde/rdfBibframe2/Work/>



# The SHARE Vocabulary question

By now it is likely clear that modeling and vocabularies used by the Casalini SHARE VDE represent a different flavor from LC bf:

<<http://share-vde.org/sharevde/rdfBibframe2/Title/1b8aba79-fbdb-392f-bc2a-9e061bb63db5>> <<http://www.w3.org/2000/01/rdf-schema#label>> " Will you **please** be quiet, **please**" <<http://share-vde.org>>

<<http://share-vde.org/sharevde/rdfBibframe2/Work/285922>>

<<http://www.w3.org/1999/02/22-rdf-syntax-ns#type>>

<<http://id.loc.gov/ontologies/bibframe/Work>> <<http://share-vde.org>> .



# Determining Workflow

Transitioning to linked data was never going to be easy. Several reasons for this include:

- MARC has been around for a very long time and we have lots of it
- The ILS and discovery systems we use are built around MARC data
- Staff have also been using MARC for a very long time
- Currently our entire workflow for resource description, from LC, OCLC and other vendors through to copy and original cataloguing all depends on MARC



It has been 15 years since Roy Tennant wrote MARC Must Die<sup>1</sup>. We now find ourselves in a position where non-MARC data alternatives exist.

Let us assume that either approach (Casalini SHARE VDE or In house processes through LC Converter) allows us to fully convert our data, establish updates and work through copy and original cataloguing workflows (There are plenty of challenges here, but success seems in sight).

We now have a more general infrastructure challenge:

- UAL and NEOS have a lot invested in our current ILS and support mechanisms
- While UAL utilized BlackLight, significant work is still being put into its development for use with our current systems. A shift for development to work with bf:2.0 data would be major
- Our MARC data does not represent our entire collection and these would all need to be connected

1. Tennant, R. (2002). MARC must die. (Digital Libraries). *Library Journal*, (17). 26.

# Strategic Planning

UAL (and other libraries) will need to develop clear strategic direction across units to tackle some of these issues.

Even so, experience from the Canadian Linked Data Initiative has highlighted that:

1. There are limited resources in any given institution for working towards implementation
2. Central planning for large scale projects across institutions is challenging

With this in mind, we wonder if libraries really could use an ally to bring about critical mass for change





# Possible Workflows

In house (LC bf:2.0 XSLT):

- Continue refining the balance of MARC vs. bf enrichment
- Automate processing to finish full collection
- Establish a process for ongoing updates
- Data into triplestore
- Discovery: BlackLight, Jupiter ... something else?

Casalini SHARE VDE:

- Continue working with phase 3 use cases to develop
  - Full collection access and updates
  - Original and copy cataloguing workflows
- Discovery already in place as proof of concept and/or production environment
- Even if only for short/medium term use, this would give a lot of breathing room for infrastructure change and learning through process improvement



We're going to need a lot of coffee!

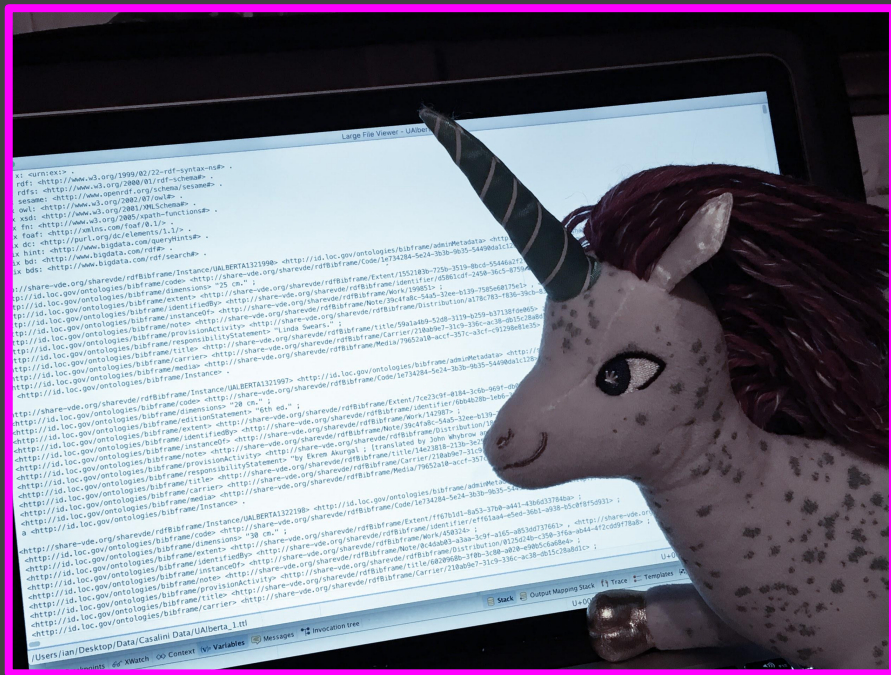
10 Sa Plays. Sk Selections  
10 Sa Three plays / Sc A  
0# Sa Ma Rainey's black bottom -- Fences -- Joe Turner's come and gone.  
8 ## S4 <http://rdaregistry.info/Elements/m/P30135> Si Has work manifested: Sa Wilson, August. | Plays. Selection  
8 ## S4 <http://rdaregistry.info/Elements/m/P30135> Si Has work manifested: Sa Three plays \$0 <http://worldcat.org>  
8 ## S4 <http://rdaregistry.info/Elements/w/P10147> Si Is container of (work): Sa Ma Rainey's Black Bottom \$0  
8 ## S4 <http://rdaregistry.info/Elements/w/P10147> Si Is container of (work): Sa Wilson, August. | Ma Rainey's b  
8 ## S4 <http://rdaregistry.info/Elements/w/P10147> Si Is container of (work): Sa Fences \$0 [http://www.wikid](http://www.wikidata.org)  
8 ## S4 <http://rdaregistry.info/Elements/w/P10147> Si Is container of (work): Sa Wilson, August. | Fences \$0  
8 ## S4 <http://rdaregistry.info/Elements/w/P10147> Si Is container of (work): Sa Joe Turner's Come and Gone \$0  
eld 758 with URIs and labels, materials specified and primary and other relationships explicitly given; application  
examples in this section \$1 is used for the wikidata and ja.dbpedia. Because these sources model their entities  
ads:Authority. Examples with \$1 can be parsed as assertions about the R while examples with \$0 would  
e when converting to RDF.  
ard, Henry,  
100 1# Sa Beard, Henry. Lord of the rings : Sb a parody of J.R.R. Tolkien's The Lord of the Rings / Sc by Henry N. Beard and Dou  
245 10 Sa Bored of the Rings / Sc by Henry N. Beard and Douglas  
758 ## S4 <http://rdaregistry.info/Elements/m/P30135> Si Has work manifested: Sa Bored of the Rings \$1 [http://www.w](http://www.wikidata.org)  
758 ## S4 <http://rdaregistry.info/Elements/w/P10197> Si Parody of (work): Sa Lord of the rings (Tolkien, J. R. R.) \$0  
100 1# Sa Dyck, Andrew R. Sq (Andrew Roy), Sd 1947-  
245 12 Sa A commentary on Cicero, De legibus / Sc Andrew R. Dyck.  
600 10 Sa Cicero, Marcus Tullius. St De legibus.  
758 ## S4 <http://rdaregistry.info/Elements/w/P10187> Si Commentary on (work): Sa De Legibus \$1 [http://www.wikida](http://www.wikidata.org)  
130 0# Sa Shichinin no samurai (Motion picture)  
245 10 Sa Seven samurai / Sc Janus Films : Toho Co., Ltd. : produced by Sojiro Motoki : screenplay by Akira Kurosawa  
758 ## S4 <http://rdaregistry.info/Elements/m/P30135> Si Has work manifested: Sa Seven Samurai \$1 [http://www.wikida](http://www.wikidata.org)  
758 ## S4 <http://rdaregistry.info/Elements/m/P30135> Si Has work manifested: Sa 七人の侍 \$1 [ja.dbpedia.org/resource](http://ja.dbpedia.org/resource)

No single version of bf will be forever, it will evolve and develop with usage. We need to flip the switch to implementation planning in order to:

- Work on incremental improvement and change through experimentation
- Enlist the expertise of the wider library community - This work cannot be completed by a select few experts
- Begin creating the necessary infrastructure to fully transition

Changing to an implementation mindset does not mean things will happen overnight, but might actually allow libraries to move on.

Many thanks to all the magical individuals working with us!



From Bibliographic Services and those across UAL to everywhere else

# THANK YOU!

QUESTIONS? COMMENTS?

[bigelow@ualberta.ca](mailto:bigelow@ualberta.ca)

[sharon.farnel@ualberta.ca](mailto:sharon.farnel@ualberta.ca)

[danoosh@ualberta.ca](mailto:danoosh@ualberta.ca)