Will you be my bf: forever?

...

Analysis Techniques for Conversion to BIBFRAME at the University of Alberta

Ian Bigelow, Sharon Farnel and Danoosh Davoodi

UNIVERSITY OF ALBERTA
LIBRARIES

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?

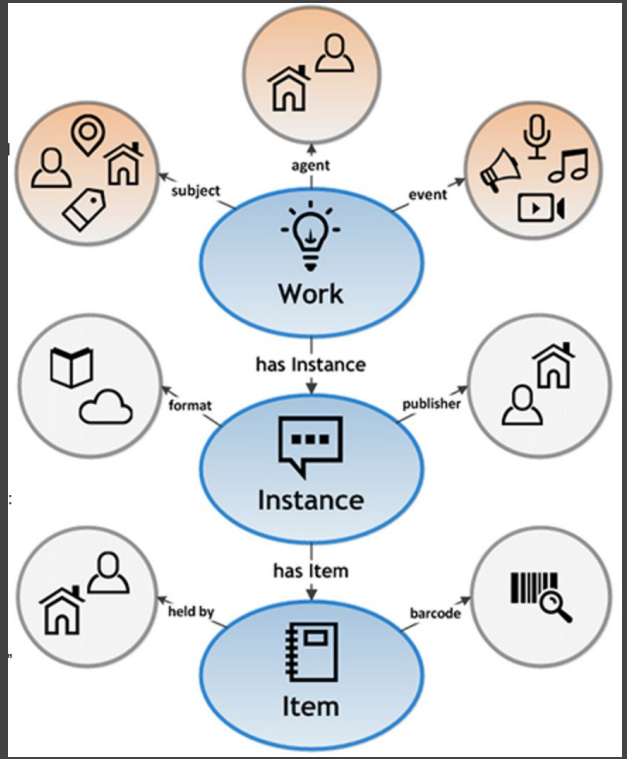
BIBFRAME
Casalini & @Cult

BIBFRAME
Library of Congress

UNIVERSITY OF ALBERTA
LIBRARIES

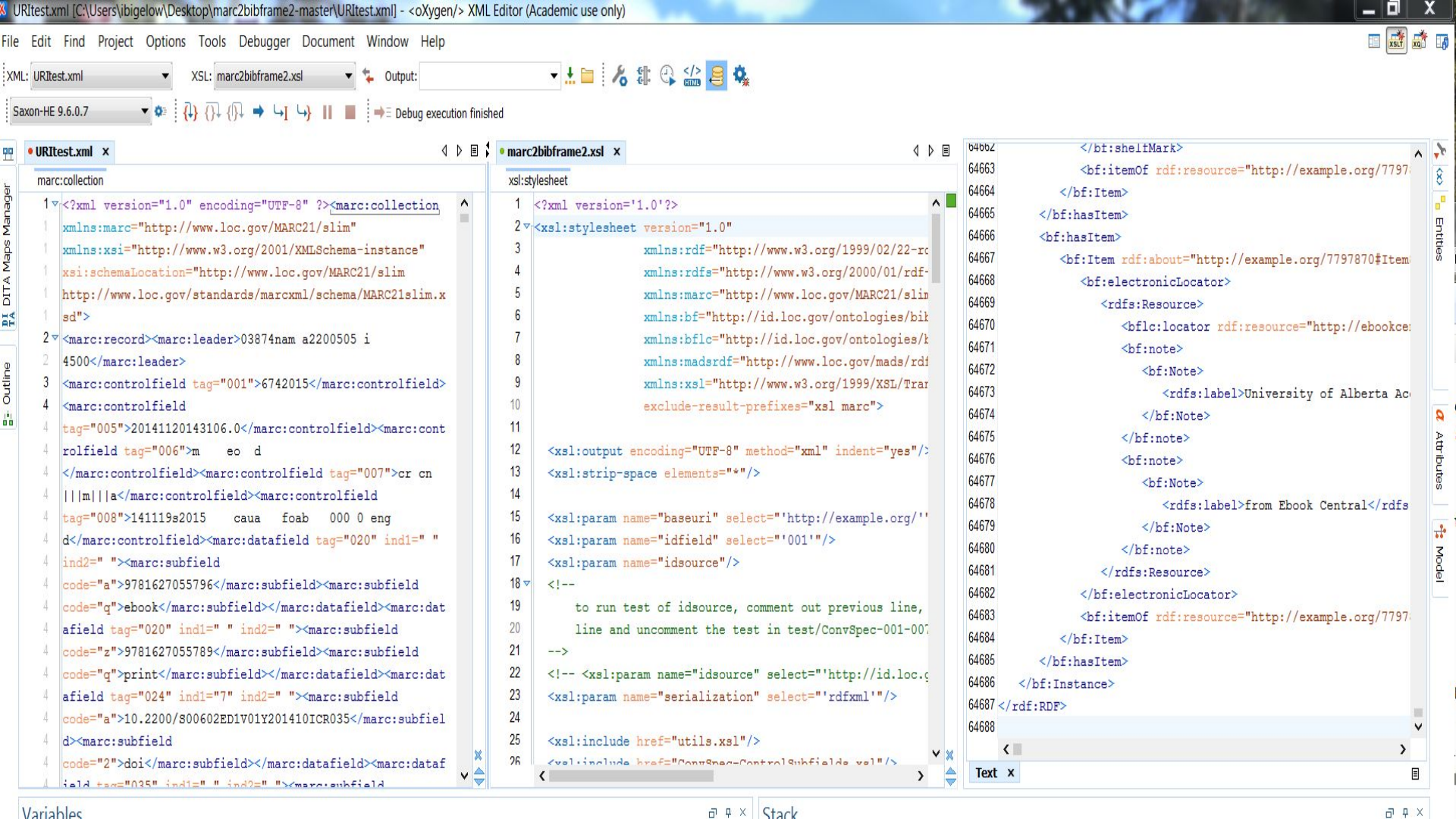
Overview of the LC BIBFRAME Converter

- 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

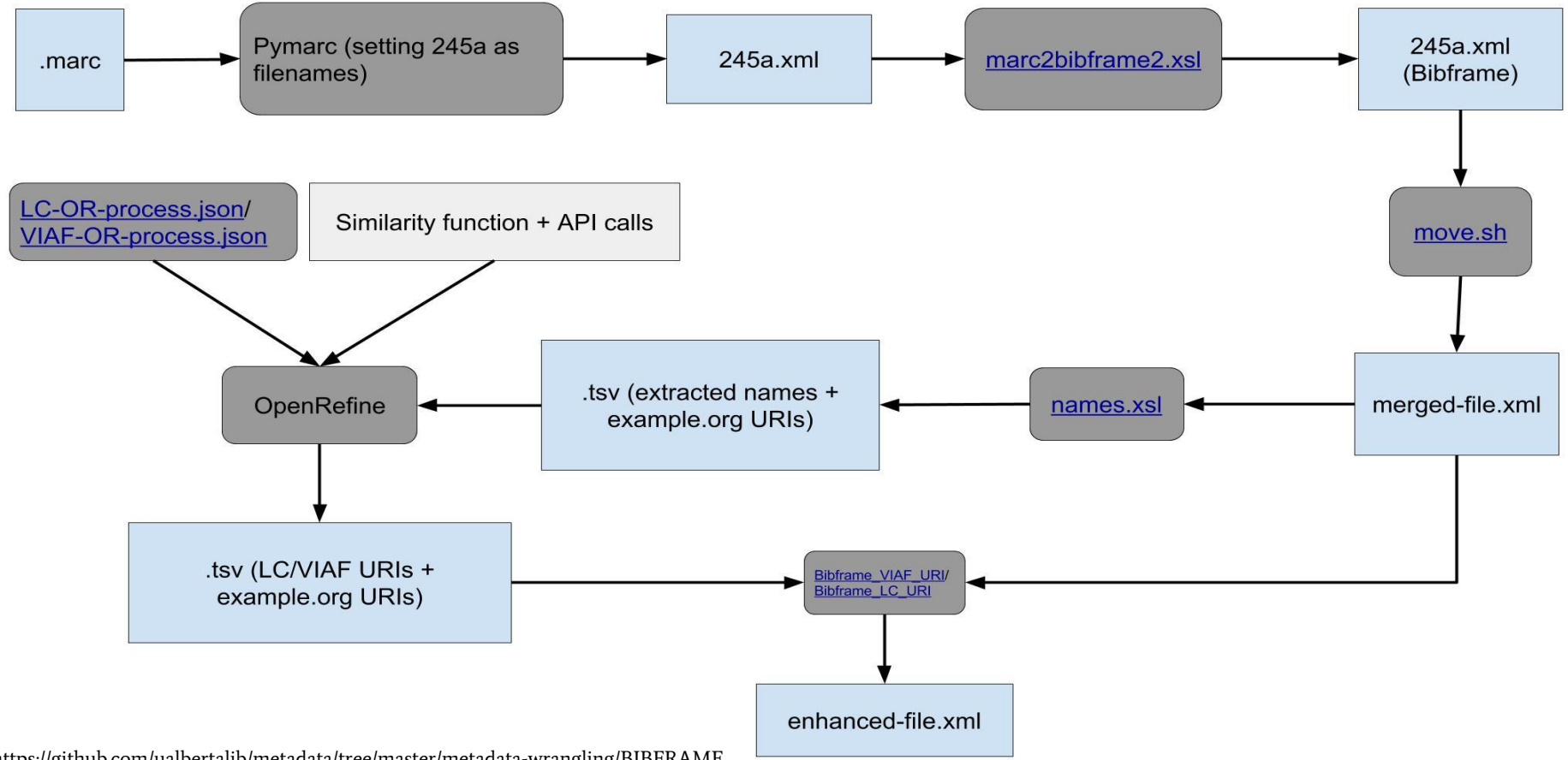


http://bibframe.org/

UNIVERSITY OF ALBERTA LIBRARIES


URItest.xml [C:\Users\ibigelow\Desktop\marc2bibframe2-master\URItest.xml] - <oXygen/> XML Editor (Academic use only)
File Edit Find Project Options Tools Debugger Document Window Help
XML: URItest.xml
XSL: marc2bibframe2.xsl
Output:
Saxon-HE 9.6.0.7
Debug execution finished
URItest.xml
marc:collection
<?xml version="1.0" encoding="UTF-8" ?><marc:collection
xmlns:marc="http://www.loc.gov/MARC21/slim"
xmlns:xsi="http://www.w3.org/2001/XMLSchema-instance"
xsi:schemaLocation="http://www.loc.gov/MARC21/slim
http://www.loc.gov/standards/marcxml/schema/MARC21slim.x
sd">
<marc:record><marc:leader>03874nam a2200505 i
4500</marc:leader>
<marc:controlfield tag="001">6742015</marc:controlfield>
<marc:controlfield
tag="005">20141120143106.0</marc:controlfield><marc:cont
rolfield tag="006">m eo d
</marc:controlfield><marc:controlfield tag="007">cr cn
|||m|||a</marc:controlfield><marc:controlfield
tag="008">141119s2015 caua foab 000 0 eng
d</marc:controlfield><marc:datafield tag="020" ind1=" "
ind2=" "><marc:subfield
code="a">9781627055796</marc:subfield><marc:subfield
code="q">ebook</marc:subfield></marc:datafield><marc:dat
afield tag="020" ind1=" " ind2=" "><marc:subfield
code="z">9781627055789</marc:subfield><marc:subfield
code="q">print</marc:subfield></marc:datafield><marc:dat
afield tag="024" ind1="7" ind2=" "><marc:subfield
code="a">10.2200/S00602ED1V01Y201410ICR035</marc:subfiel
d><marc:subfield
code="2">doi</marc:subfield></marc:datafield><marc:dataf
marc2bibframe2.xsl
xsl:stylesheet
<?xml version='1.0'?>
<xsl:stylesheet version="1.0"
xmlns:rdf="http://www.w3.org/1999/02/22-rd
xmlns:rdfs="http://www.w3.org/2000/01/rdf-
xmlns:marc="http://www.loc.gov/MARC21/slim
xmlns:bf="http://id.loc.gov/ontologies/bib
xmlns:bflc="http://id.loc.gov/ontologies/b
xmlns:madsrdf="http://www.loc.gov/mads/rdf
xmlns:xsl="http://www.w3.org/1999/XSL/Tran
exclude-result-prefixes="xsl marc">
<xsl:output encoding="UTF-8" method="xml" indent="yes"/>
<xsl:strip-space elements="*"/>
<xsl:param name="baseuri" select="'http://example.org/''
<xsl:param name="idfield" select="'001'"/>
<xsl:param name="idsource"/>
<!--
to run test of idsource, comment out previous line,
line and uncomment the test in test/ConvSpec-001-007
-->
<!-- <xsl:param name="idsource" select="'http://id.loc.g
<xsl:param name="serialization" select="'rdfxml'"/>
<xsl:include href="utils.xsl"/>
</bf:shelfMark>
<bf:itemOf rdf:resource="http://example.org/7797
</bf:Item>
</bf:hasItem>
<bf:hasItem>
<bf:Item rdf:about="http://example.org/7797870#Item
<bf:electronicLocator>
<rdfs:Resource>
<bflc:locator rdf:resource="http://ebookce
<bf:note>
<bf:Note>
<rdfs:label>University of Alberta Ac
</bf:Note>
</bf:note>
<bf:note>
<bf:Note>
<rdfs:label>from Ebook Central</rdfs
</bf:Note>
</bf:note>
</rdfs:Resource>
</bf:electronicLocator>
<bf:itemOf rdf:resource="http://example.org/7797
</bf:Item>
</bf:hasItem>
</bf:Instance>
</rdf:RDF>
Text
Variables
Stack
DITA Maps Manager
Outline
Entities
Attributes
Model

Lots of data in BIBFRAME 2



https://github.com/ualbertalib/metadata/tree/master/metadata-wrangling/BIBFRAME

```
<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:madsrdf="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	Waseda 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	Yourman, 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

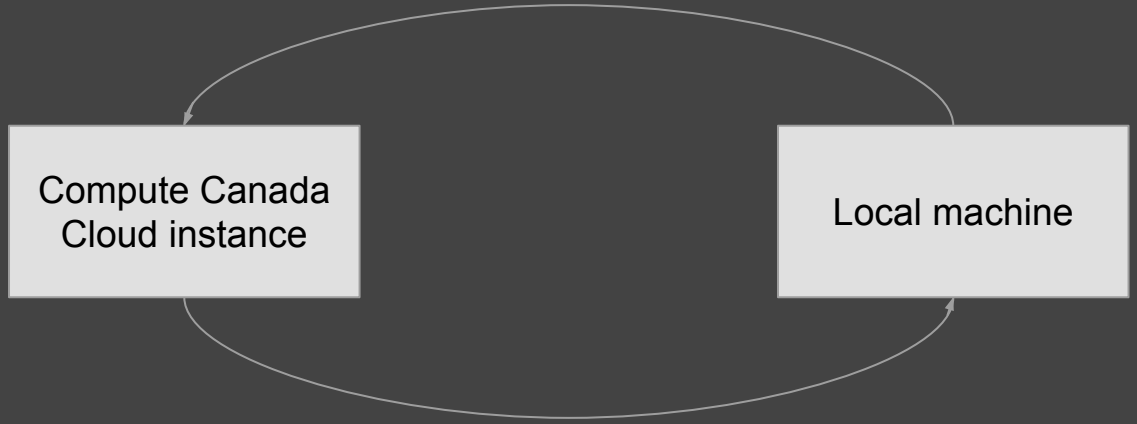
how as: **rows** records Show: 5 10 **25** 50 rows « first ‹ previous **1 - 25** next › last

All	Names	match-ID-F	score-lc	Names1	LCP	match-ID	score	Cor	LCC	score-c	Corp/personal	Column 3	Ingest Key
	Search for match												
11.	Conservatorio Superior de Música Conservatorio Superior de Música de Sevilla (84) Real Conservatorio Superior de Música de Madrid (79) Conservatorio Superior de Música Santa Cruz de Tenerife (72) Create new topic Search for match	n84186607	84	Conservatorio Superior de Música Carmona, José Carlos, 1963- (0.188) Pons, Jordi, 1960- (0.219) Vidal Corella, Vicente (0.125) Create new topic				Conservatorio Superior de Música Conservatoire national supérieur de musique et de danse de Paris (0.406) Real Escuela Superior de Arte Dramático (Spain) (0.313) Real Conservatorio Superior de Música de Madrid (0.681) Create new topic			Organization		http://example.org/108 18
12.	Puiggrós, Rodolfo, 1906-1980 Choose new match	n88232712	98	Puiggrós, Rodolfo, 1906-1980 Frank, André Gunder, 1929-2005 (0.167) Puiggrós, Rodolfo, 1906-1980 (1) Create new topic	n88232712	n88232712	1	Puiggrós, Rodolfo, 1906-1980 Create new topic			Person		http://example.org/104 11
13.	Silva Neto, Serafim da, 1917-1960 Choose new match	n80152925	100	Silva Neto, Serafim da, 1917-1960 Silva Neto, Serafim ★da★ 1917-1960 (0.941) Create new topic	n80152925	n80152925	0.9411764705882353	Silva Neto, Serafim da, 1917-1960 Create new topic			Person		http://example.org/100 10
14.	Font y Sagué, Norbert, 1873-1910 Create new topic Search for match			Font y Sagué, Norbert, 1873-1910 Create new topic				Font y Sagué, Norbert, 1873-1910 Create new topic			Person		http://example.org/100 10
15.	Eyzaguirre, Jaime Choose new match	n80090270	100	Eyzaguirre, Jaime Eyzaguirre, Jaime. (0.944) Barros van Buren, Mario (0.261) Eyzaguirre, Jaime, 1935- (0.708) Create new topic	n80090270	n80090270	0.9444444444444444	Eyzaguirre, Jaime Universidad Católica de Chile. Oficina Coordinadora de Investigaciones "Profesor Jaime Eyzaguirre" (0.111) Create new topic			Person		http://example.org/103 12
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	1	Valdeavellano, Luis G. de (Luis García de), 1904- Create new topic			Person		http://example.org/106 11
17.	Tudebodus, Petrus Tudebodus, Petrus, active 1096-1100 (65) Tudebodus, Petrus, active 1096-1100 (65) Create new topic Search for match			Tudebodus, Petrus Tudebodus, Petrus, active 1096-1100 (0.486) Create new topic				Tudebodus, Petrus Create new topic			Person		http://example.org/106 12
18.	Palavicini, Félix Fulgencio, 1881- Choose new match	n82032394	98	Palavicini, Félix Fulgencio, 1881- Choose new match	n82032394	n82032394	1	Palavicini, Félix Fulgencio, 1881- Create new topic			Person		http://example.org/104 11
19.	Mexico. Congreso Constituyente (1916-1917) Choose new match	n83120024	100	Mexico. Congreso Constituyente (1916-1917) Create new topic		n83120024		Mexico. Congreso Constituyente (1916-1917) Choose new match	n83120024	1	Jurisdiction		http://example.org/104 20
20.	Mexico Choose new match	n81013960	100	Mexico Eisenstein, Sergei, 1898-1948. (0.067) Cortés, Hernán, 1485-1547. (0.038) Maximilian, Emperor of Mexico, 1832-1867 (0.15) Create new topic				Mexico Universidad nacional autónoma de México (0.128) Instituto Nacional de Antropología e Historia (México) (0.093) Time-Life books (0.2) Create new topic			Jurisdiction		http://example.org/104 22
21.	Maeztu, María de, 1882- Maeztu Whitney, María de, 1882-1948 (73) Maeztu Whitney, María de, 1882-1948. Essays. Selections (54) Create new topic Search for match			Maeztu, María de, 1882- Maeztu, Ramiro de, 1875-1936. (0.586) Maeztu, Gustavo ★de★ 1887-1947 (0.567) Maeztu Whitney, María de, 1882-1948 (0.611) Create new topic				Maeztu, María de, 1882- Create new topic			Person		http://example.org/100 11
22.	Bazant, Jan Choose new match	n85099113	100	Bazant, Jan Create new topic				Bazant, Jan Create new topic			Person		http://example.org/102 11

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	
1985 Imprints	92.41%	87.22%	55.98%
2015 Imprints	96.06%	86.33%	65.36%
UA	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.”[1]

“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.”[2]

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?l=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
- Pennsylviania 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

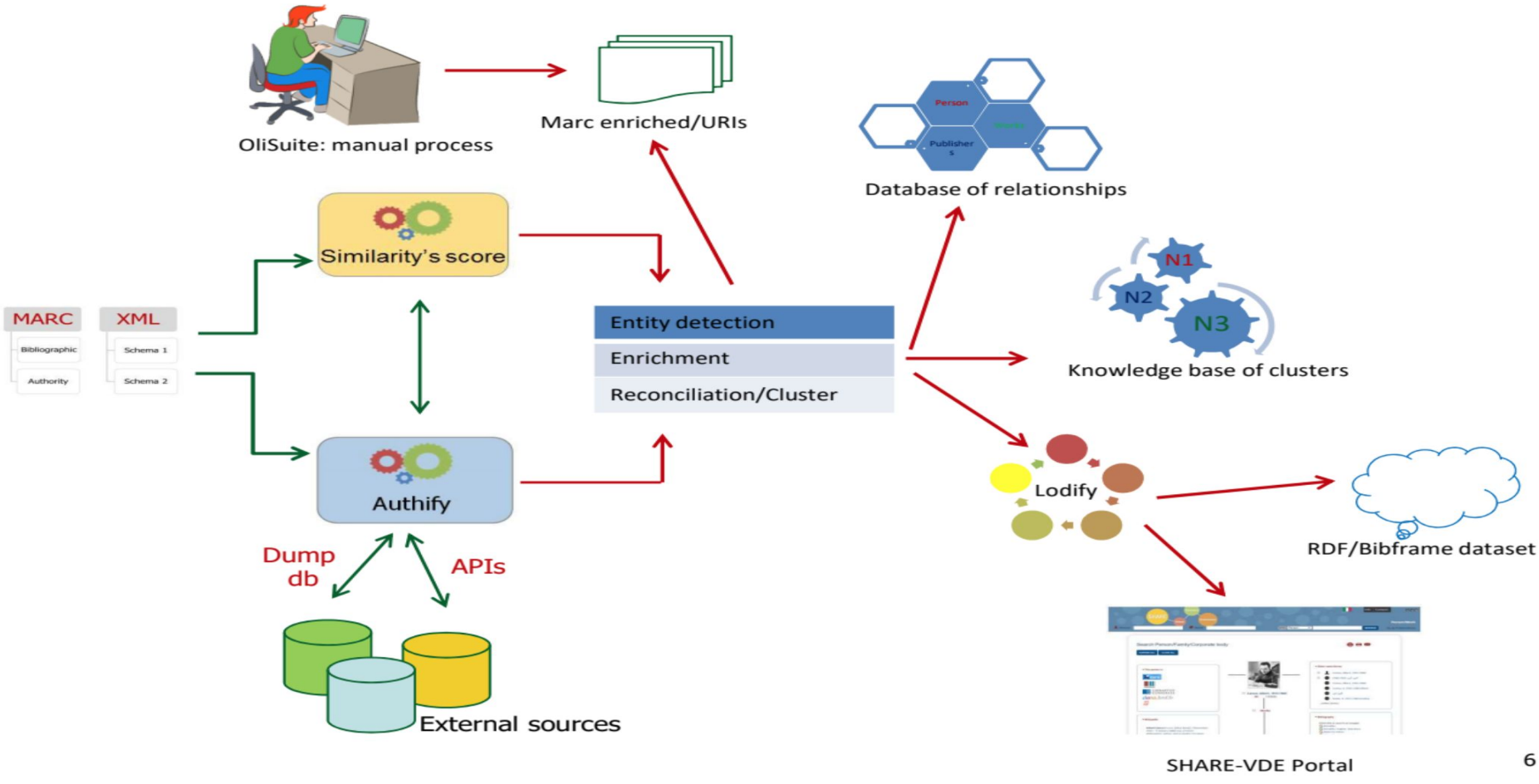
Phase 2:

- Creation of relationship database to support entity identification
- Improvements of processes from phase 1 for MARC and bf:2.0 data
- Basefile to be returned in bf:2.0 and enriched MARC
- Web discoverability through application of other ontologies such as schema.org

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

UNIVERSITY OF ALBERTA LIBRARIES

The SHARE-VDE processes overview



▾ This person in

WIKIDATA
LIBRARY OF CONGRESS
data.bnf.fr
VIAF

Shakespeare, William, 1564-1616
ID: 10236

Works

▾ Other name forms

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

...(other forms)

▾ Wikipedia

William Shakespeare (/ˈʃeɪkspɪər/; 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

▾ Bibliography

(Click title to search on Google)
"Gamlet" Borisa 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

UNIVERSITY OF ALBERTA
LIBRARIES

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[1] / CSR[2] 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 BIBRAME 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*[1], 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>
```

UNIVERSITY OF ALBERTA LIBRARIES

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 $aSengupta, Ashis,$eeditor.</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>
```

UNIVERSITY OF ALBERTA LIBRARIES

MARC from Casalini Enriched with URI

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

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

=758 \\$4http://rdaregistry.info/Elements/m/P30004$0http://www.worldcat.org/oclc/900194099
=758 \\$4http://rdaregistry.info/Elements/m/P30139$1http://worldcat.org/entity/work/id/2267517995

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

UNIVERSITY OF ALBERTA LIBRARIES

URI in Casalini bf:2.0 Data

<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> .

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[1]. 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

UNIVERSITY OF ALBERTA LIBRARIES

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

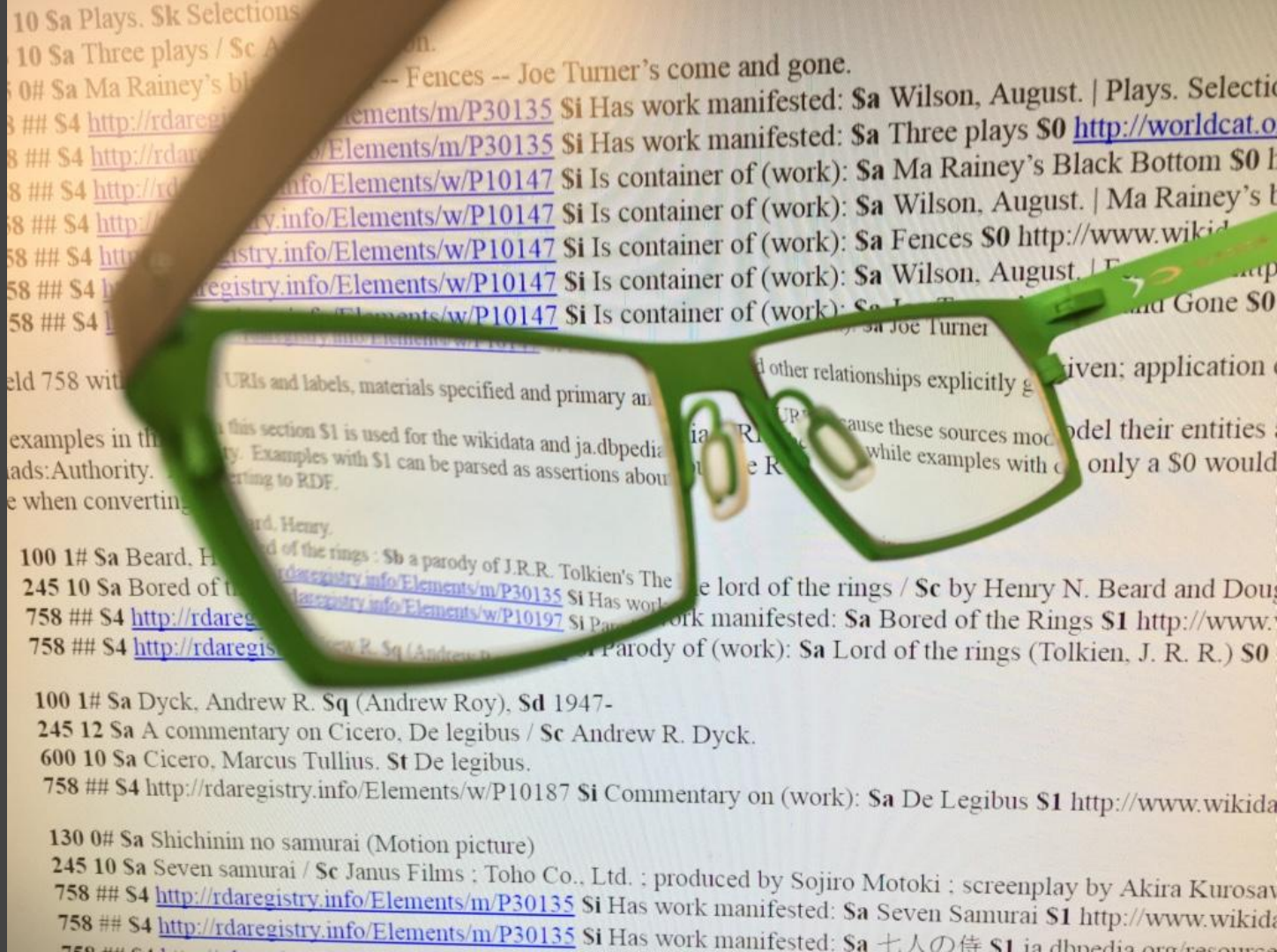
In house (LC bf:2.0 XSLT):

- 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?

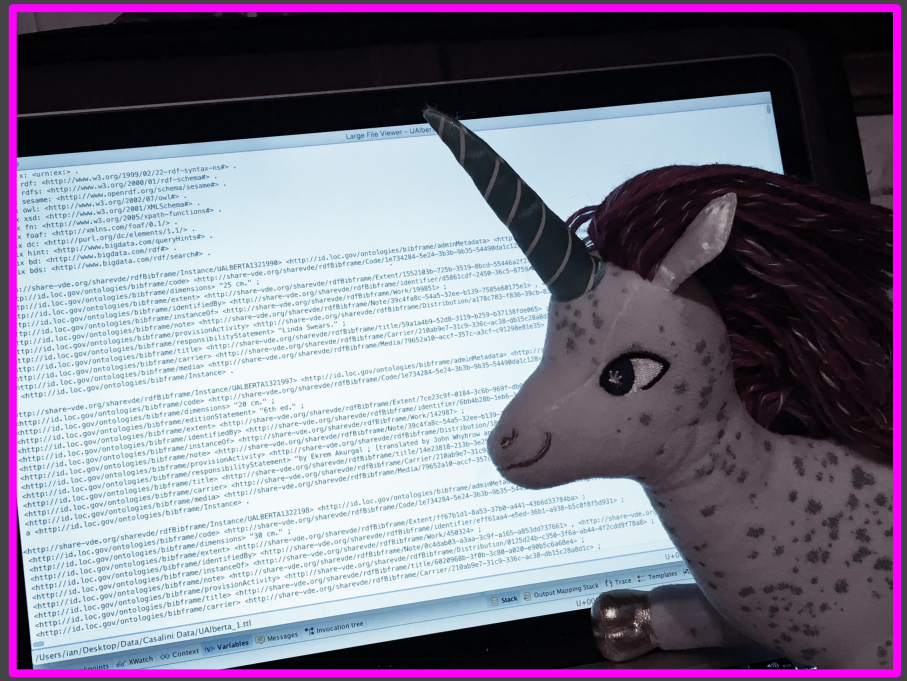
Casalini SHARE VDE:

- 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!



Many thanks to all the magical individuals working with us!



From Bibliographic Services and those across UAL to everywhere else

UNIVERSITY OF ALBERTA LIBRARIES

THANK YOU!

QUESTIONS? COMMENTS?

bigelow@ualberta.ca

sharon.farnel@ualberta.ca

danoosh@ualberta.ca

UNIVERSITY OF ALBERTA
LIBRARIES