

European BIBFRAME Workshop  
Tutorial - 17 September 2018  
Firenze

# The mapping and the conversion workflow from MARC to BIBFRAME

Tiziana Possemato

Casalini Libri - @Cult



# Introduction and overall projects goals



# Our BIBFRAME projects overall goals

The main areas of our projects:

- **Enrichment** of MARC record with URIs
- **Conversion from MARC to RDF** using the BIBFRAME vocabulary (and other additional ontologies as needed)
- **Data publication** according to the **BIBFRAME** data model
- Batch/automated **data updating** procedures
- Batch/automated **data dissemination** to libraries
-

**The projects heart:**

**Entity identification, Reconciliation,  
Data enrichment  
and  
BIBFRAME Conversion**



# The new revolution: from record to entity

Full Record	MARC Tags
	<a href="#">Where to Request</a>
000	02384cam a2200445 a 4500
001	15789102
005	20100204124659.0
008	090624s2009 enka b 000 0 eng
010	__  a 2009025024
015	__  a GBA959347  2 bnb
016	7_  a 015291646  2 UK
020	__  a 9780521519748 (hbk.)
020	__  a 0521519748 (hbk.)
020	__  a 9780521732505 (pbk.)
020	__  a 0521732506 (pbk.)
035	__  a (OCoLC)ocn416716652
040	__  a DLC  c DLC  d YDXCP  d BWKUK  d BWK  d FDA  d UKM  d BTDTA  d CEF  d CDX  d DLC
050	00  a PR2803.A2  b H35 2009
082	00  a 822.3/3  2 22
100	1_  a Shakespeare, William,  d 1564-1616.
245	10  a <b>As you like it</b> /  c edited by Michael Hattaway.
250	__  a Updated ed.
260	__  a Cambridge, UK ;  a New York :  b Cambridge University Press,  c 2009.
300	__  a xv, 240 p. :  b ill. ;  c 24 cm.
490	1_  a New Cambridge Shakespeare
504	__  a Includes bibliographical references (p. 240).
505	0_  a Introduction. Journeys ; Plays within the play ; Theatrical genres ; Pastoral ; Counter-pastoral ; The condition of the country ; Politics ; 'Between you and the women the play must please' ; Gender ; Nuptials ; Sources ; Date and occasion ; Stage history ; Recent critical and stage interpretations -- Note on the text -- List of characters -- The play -- Textual analysis -- Appendixes : 1. An early court performance? ; 2. Extracts from Shakespeare's principal source, Lodge's Rosalind ; 3. The songs.
600	10  a Shakespeare, William,  d 1564-1616.  t <b>As you like it</b> .
650	_0  a Fathers and daughters  v Drama.
650	_0  a Exiles  v Drama.
655	_7  a Pastoral drama.  2 gsafd
655	_7  a Comedies.  2 gsafd
700	1_  a Hattaway, Michael.

As you like it [print]



Instance

As you like it [on-line]

Cambridge University Press



Publisher

Cambridge Univ. Press

Shakespeare, William, 1564-1616

Шекспир, У. 1564-1616 Уильям



Agent

Saixpēr, Gouilliam, 1564-1616

As you like it



Work

Come ti piace

Comme il vous plaira

Fathers and daughters



Subject

Patri e figlie

Pères et filles

# How reconciliation is obtained

Scope of these processes is to bring together and to make data available from different sources in a way that could be defined as *democratic* to better identify the entity in question.

Data reconciliation and enrichment is obtained by:

- *automated processes*
- *manual processes*

It is important to underline how the **relationship between the reconciliation and validation of the results** can differ greatly between the automated and manual processes:

- automated processes: a high level of reconciliation and clustering; a low level of result validation;
- manual processes: a low level of reconciliation and clustering; a high level of result validation.



# Entities in *cluster*: an example of collaboration and sharing



Vivaldi, Antonio, 1678-1741  
ID: 37154

▼ Questo autore in



▼ Altre forme del nome

- Vivaldi, Antonio, 1678-1741
- VIAF 1678-1741, אנטוניו, ויבאדי
- VIAF Vivaldi, Antonio, 1678-1741
- VIAF Vivaldi, Antonio
- VIAF Vivaldi, Antonio, sac., 1678-1741
- VIAF Вивальди, А. 1678-1741 Антонио
- VIAF Вивальди, Антонио, 1678-1741
- VIAF Vivaldi, Antonio, 1680-1741
- VIAF 1741-1678, فيدالدي, أنطونيو
- VIAF Antonio Vivaldi compositore e violinista italiano esponente di spicco del tardo barocco veneziano
- VIAF Vivaldi, Antonio, ca.1678-1741
- VIAF Vivaldi, Antonio (Italian composer and musician, 1678-1741)
- Prete rosso, 1678-1741
- Vivaldis, A., 1678-1741
- Vivaldi, A. (Antonio), 1678-1741
- Vivarudi, Antonio, 1678-1741
- Vivaldi, Antonio

...(altre forme)

The result of a reconciliation of the entity *Antonio Vivaldi* in the Share VDE project, with data from different sources and projects:

- the authorized form from a local authority file
- the variant forms originating from the references on the local authority records
- the variant forms originating from the VIAF
- the forms of the name used in the bibliographic records.

The cluster is completed and enriched with identifiers for the same entity, Antonio Vivaldi, from sources such as:

- Wikidata
- Library of Congress Name Authority File
- Data.bnf.fr
- VIAF

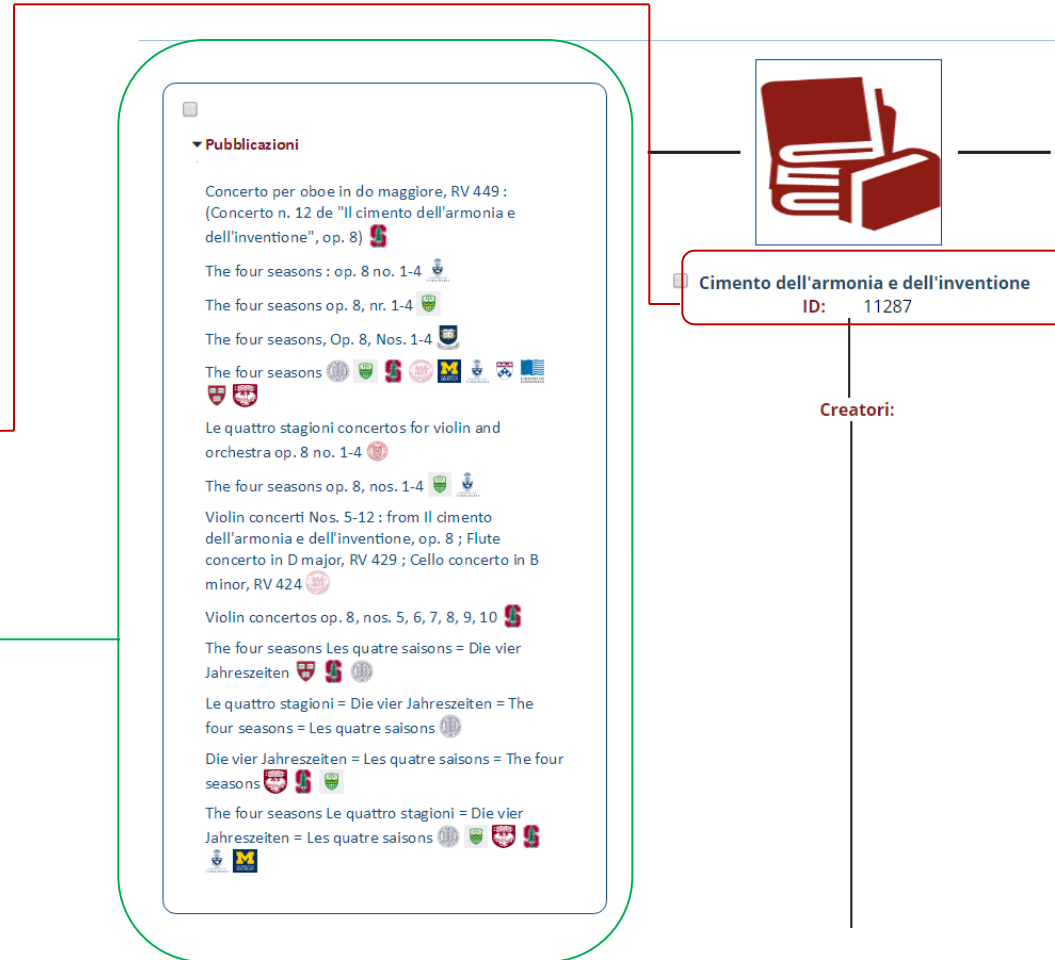


# An example of Work/Instances reconciliation

Grouping under a single work title of the many publication titles in the catalogue for *Cimento dell'armonia e dell'invenzione*

Single work title

Brings together different publications/resources present in different catalogues.



[http://share-vde.org/sharevde/searchTitles?t\\_cluster\\_id=11287](http://share-vde.org/sharevde/searchTitles?t_cluster_id=11287)



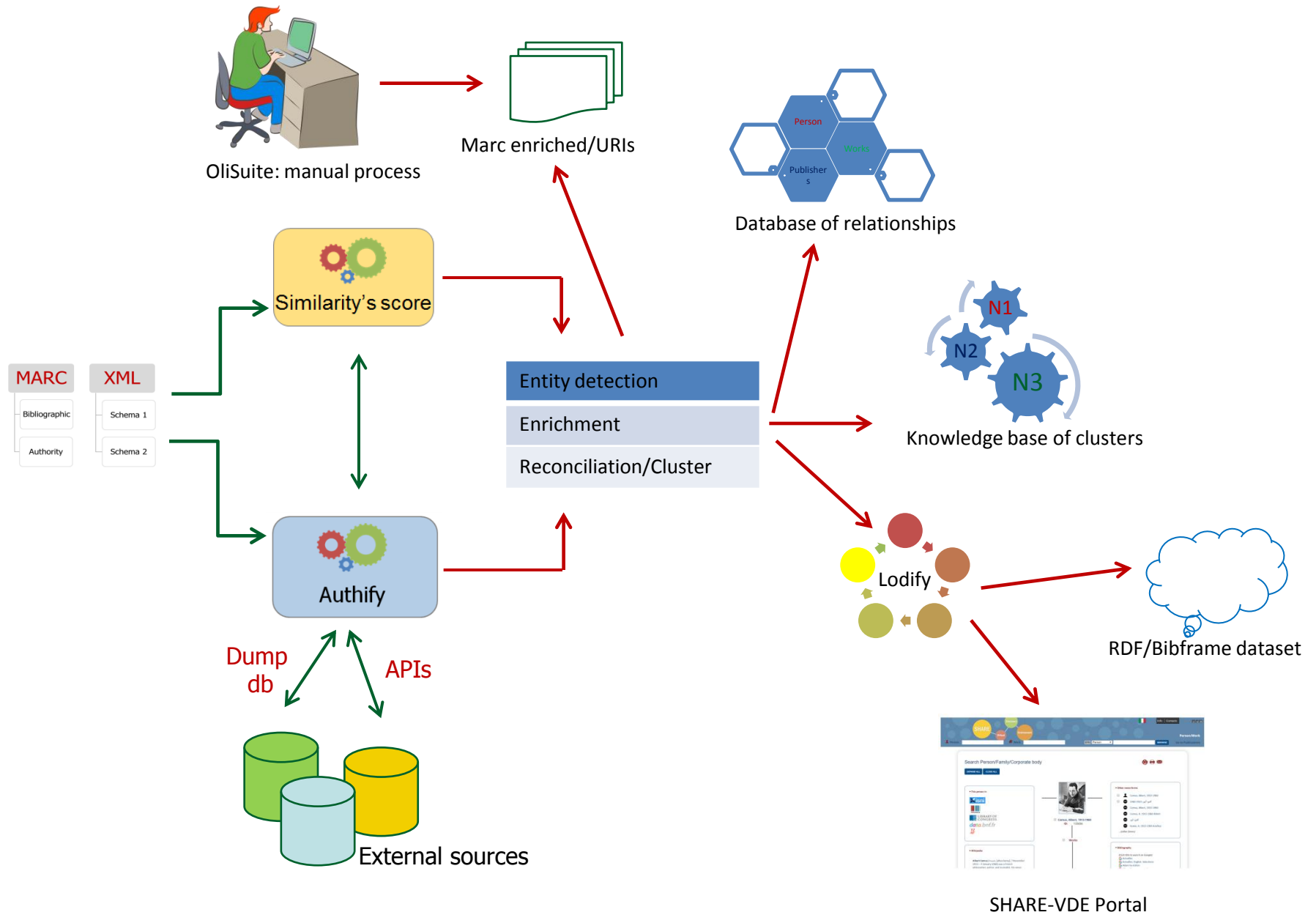


# BIBFRAME projects

## Process overview



# The SHARE-VDE processes

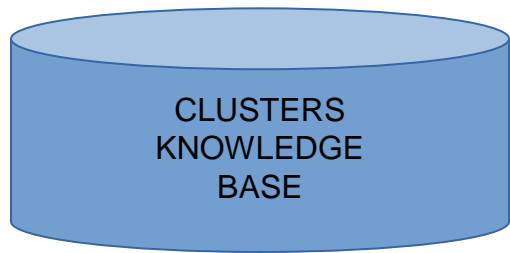
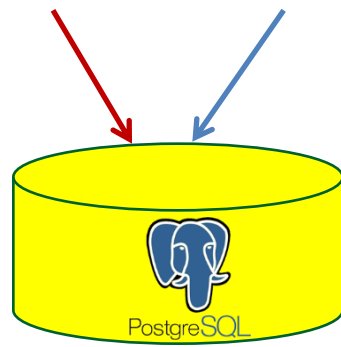
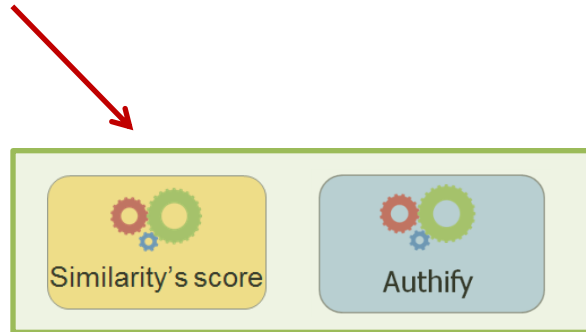


# Focus on processes 1/2

## Authority records



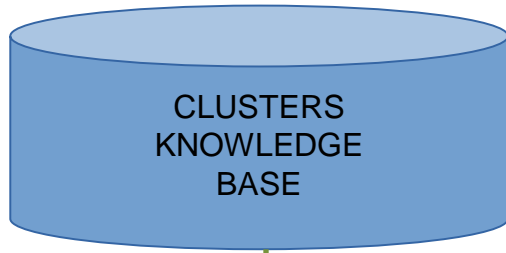
## Bibliographic records



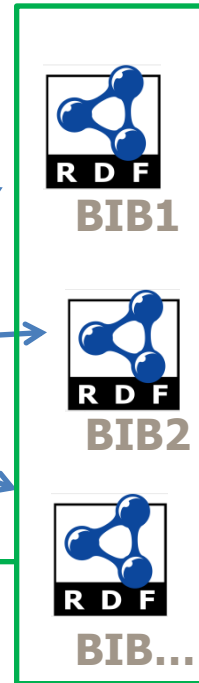
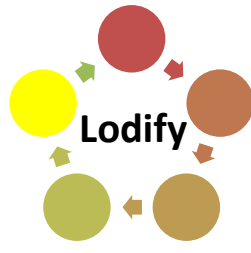
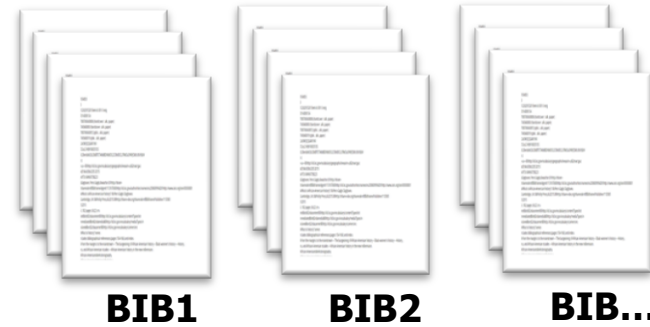
## Marc enriched (.pxml)



# Focus on processes 2/2

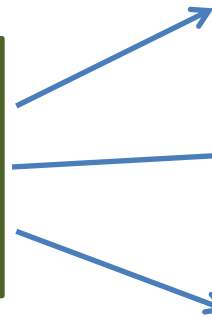


Marc enriched (.pxml)



Clusters Knowledge base

SHARE-VDE URIs External (VIAF) URIs



# The transformation processe step-by-step



# The clusterization process – Brief overview

The entity clusterization is pre-med to Marc-to-BF conversion process. It's realized through 3 main processes plus a data enrichment final step:

1. **Agent Authority data analysis and handling**, for a first pre-clusterization process.
2. **Bibliographic records handling**, with creation/feeding of Agent cluster and creation of Work title cluster (if a uniform title or preferred access point it exists).
3. Creation/feeding of **Work title cluster**, from proper title (tag 245).
4. **Marc 21 enrichment** (with URIs assigned in the previous three processes).



# 1. Agent Authority data analysis and handling

We try to design the process step-by-step using a specific example (that will become the cluster ID 2660985). We have two Authority records coming from the same Institution (Harvard) and related to the same entity: **Chen, Hu, 1875-1922**

Authority record #1

=100 1\\${a}Chen, Hu,\$d1875-1922  
=400 1\\${a}Chen, Cangyu,\$d1875-1922  
=400 1\\${a}Chen, Zhenshan,\$d1875-1922  
=400 1\\${a}陳瑚,\$d1875-1922  
=667 \\\\${a}Machine-derived non-Latin script reference project.  
=667 \\\\${a}Non-Latin script reference not evaluated.  
=670 \\\\${a}Zhenshan shi chao, 1992:\$bt.p. (Chen Hu) biog. data (Chen Hu (1875-1922), t. Cangyu, h. Zhenshan, Miaoli Yuanli ren)

Authority record #2

=100 1\\${a}Chen, Hu  
=400 1\\${a}陳湖  
=400 1\\${a}陈虎  
=400 1\\${a}陈湖  
=400 1\\${a}



# 1. Agent Authority data analysis and handling

After the data analysis (applying the specific identification and clusterization logics [1]) the data are saved in a PostgreSQL database.

In this example, data are saved in two different tables:

- a. the **preferred forms** coming from Authority tag 1xx (saved in the **aut\_nme** table);
- b. the **variant forms** coming from Authority tag 4xx (saved in the **aut\_xref** table).

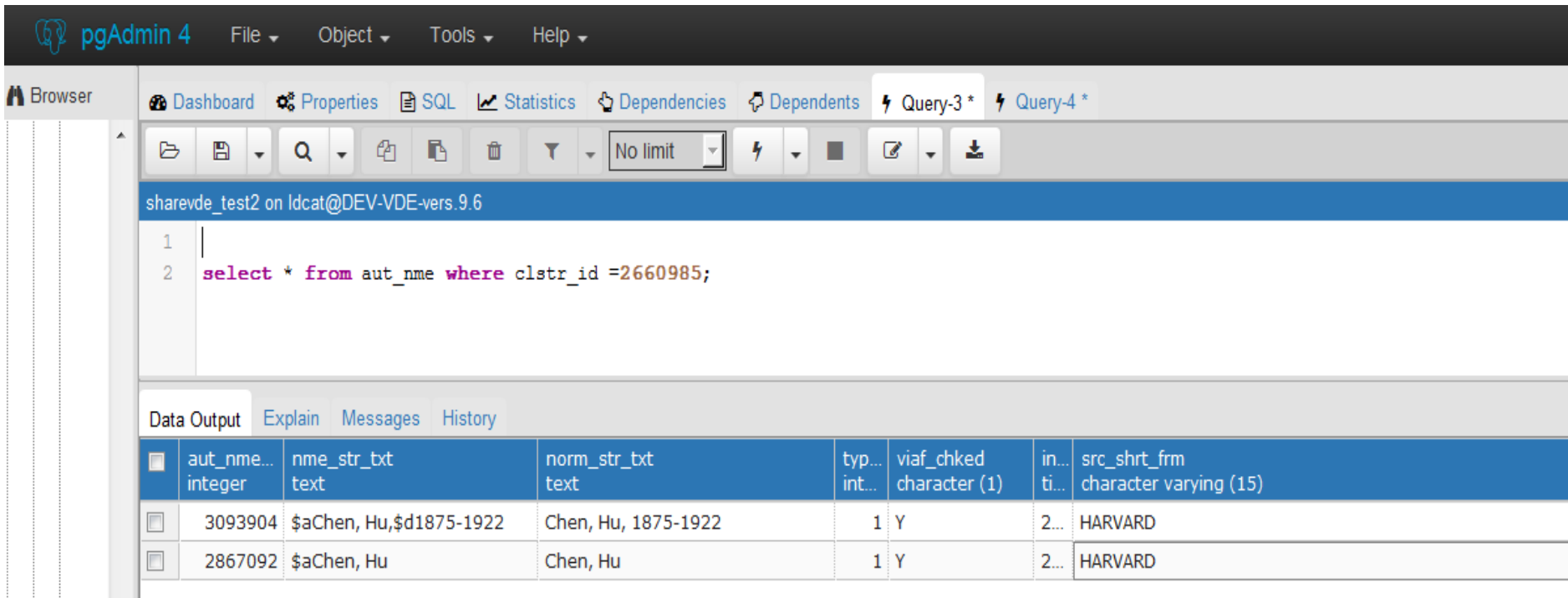
[1] the logic to search, analyse, identify and create a cluster starting from Marc data is not described here. We assume to apply a certain number of logicals and (related) machine processes to feed the PostgreSQL database that is part of the Clusterization architecture.





# 1. Agent Authority data analysis and handling

- a. the **preferred forms** in the `aut_nme` table:



The screenshot shows the pgAdmin 4 interface. The top menu bar includes 'File', 'Object', 'Tools', and 'Help'. The browser pane shows a query executed on the 'sharevde\_test2' database. The query is:

```
1 |
2 | select * from aut_nme where clstr_id =2660985;
```

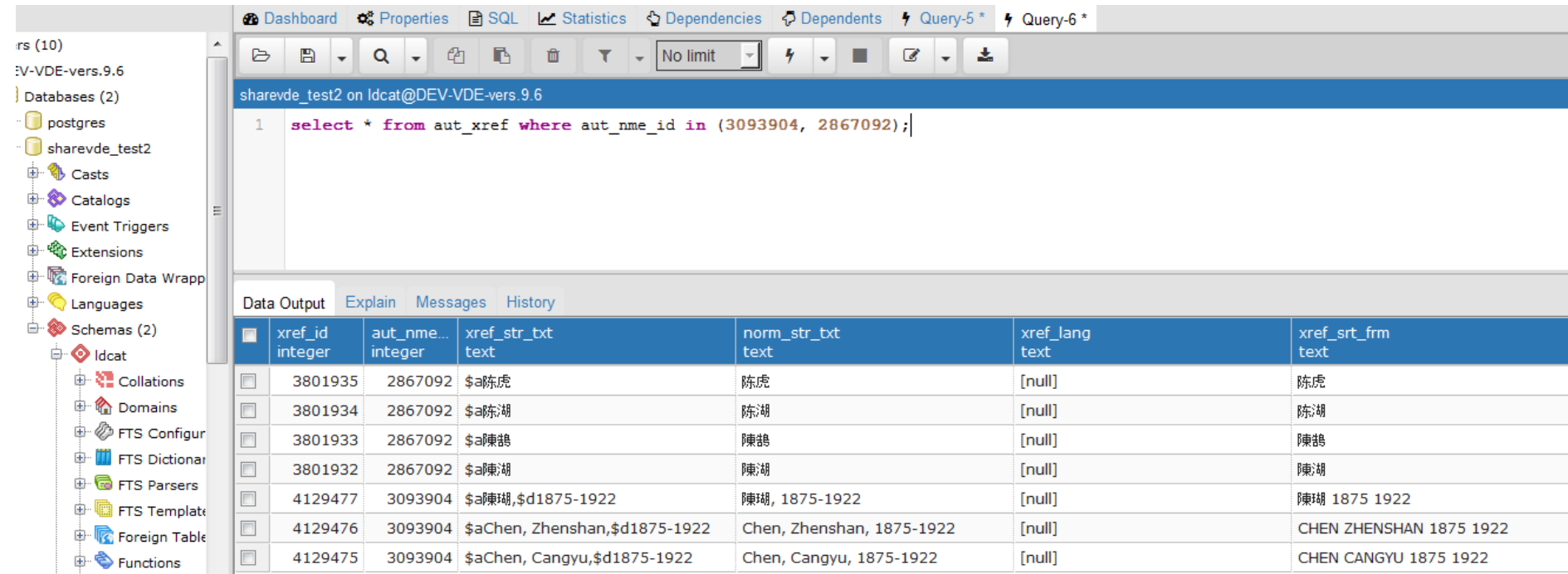
The results are displayed in a table with the following columns and data:

aut_nme...	nme_str_txt	norm_str_txt	typ...	viaf_chked	in...	src_shrt_frm
integer	text	text	int...	character (1)	ti...	character varying (15)
3093904	\$aChen, Hu,\$d1875-1922	Chen, Hu, 1875-1922	1	Y	2...	HARVARD
2867092	\$aChen, Hu	Chen, Hu	1	Y	2...	HARVARD



# 1. Agent Authority data analysis and handling

- b. the **variant forms** in the **aut\_xref** table:



The screenshot shows a database management interface with a SQL query window and a results table. The query is: `select * from aut_xref where aut_nme_id in (3093904, 2867092);`

xref_id integer	aut_nme... integer	xref_str_txt text	norm_str_txt text	xref_lang text	xref_srt_fm text
3801935	2867092	\$a陈虎	陈虎	[null]	陈虎
3801934	2867092	\$a陈湖	陈湖	[null]	陈湖
3801933	2867092	\$a陈鹤	陳鶴	[null]	陳鶴
3801932	2867092	\$a陳湖	陳湖	[null]	陳湖
4129477	3093904	\$a陳瑚,\$d1875-1922	陳瑚, 1875-1922	[null]	陳瑚 1875 1922
4129476	3093904	\$aChen, Zhenshan,\$d1875-1922	Chen, Zhenshan, 1875-1922	[null]	CHEN ZHENSHAN 1875 1922
4129475	3093904	\$aChen, Cangyu,\$d1875-1922	Chen, Cangyu, 1875-1922	[null]	CHEN CANGYU 1875 1922

At the end of this process the cluster for **Chen, Hu, 1875-1922** is created (cluster ID **2660985**). It's not used on the end users portal, until a Bibliographic data is not associated to it.



# The Cluster Knowledge Base

## An example of clusterization algorithms

Clusterization of "forename" heading type

Example:

"\$aBridget,\$cof Sweden, Saint,\$dapproximately 1303-1373"

- 1) selections of interesting subfield
- 2) normalization of string text without diacritics, accents: Bridget of Sweden Saint approximately 1303-1373.
- 3) translate all in uppercase and search string into db variant forms and cross references:  
BRIDGET OF SWEDEN SAINT APPROXIMATELY 1303-1373
- 4) if no cluster found, subfields will be analyzed
  - 4.1 comparing \$a with other existing forms
  - 4.2 comparing only the numeric part of \$d (having same \$a): \$dapproximately 1303-1373 => \$d1303-1373
  - 4.3 comparing \$c for "saint" or "santa" or other forms (having same \$a)

## 2. Bibliographic records handling

This pre-clusterization process ends with the enrichment of URIs coming from external sources (such as VIAF, LCNAF, ISNI etc.), through Authify [2].

---

Now, the process **2. Bibliographic records handling** can start.

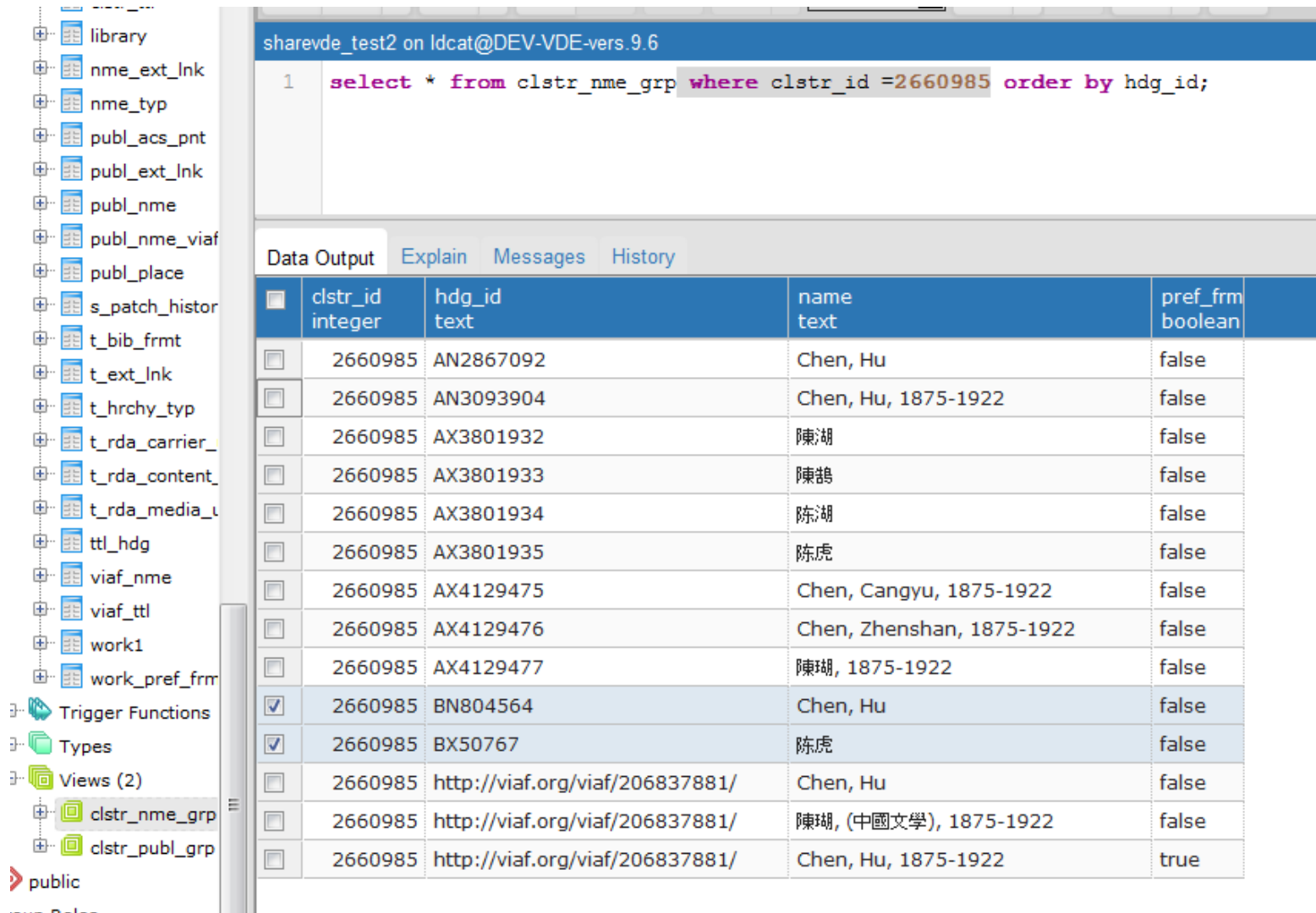
The Bibliographic process is similar (in term of logic) to the previous one, and it feeds the same PostgreSQL database.

[2] The logics applied in Authify to enrich headings are not described in this document.



## 2. Bibliographic records handling

The final cluster ID 2660985, with all preferred and variant forms, coming both from Authority and from Bibliographic records:



The screenshot displays a database interface with a query window and a results table. The query window shows the following SQL statement:

```
1 select * from clstr_nme_grp where clstr_id =2660985 order by hdg_id;
```

The results table, titled "Data Output", contains the following data:

<input type="checkbox"/>	clstr_id integer	hdg_id text	name text	pref_frm boolean
<input type="checkbox"/>	2660985	AN2867092	Chen, Hu	false
<input type="checkbox"/>	2660985	AN3093904	Chen, Hu, 1875-1922	false
<input type="checkbox"/>	2660985	AX3801932	陳湖	false
<input type="checkbox"/>	2660985	AX3801933	陳鵠	false
<input type="checkbox"/>	2660985	AX3801934	陈湖	false
<input type="checkbox"/>	2660985	AX3801935	陈虎	false
<input type="checkbox"/>	2660985	AX4129475	Chen, Cangyu, 1875-1922	false
<input type="checkbox"/>	2660985	AX4129476	Chen, Zhenshan, 1875-1922	false
<input type="checkbox"/>	2660985	AX4129477	陳瑚, 1875-1922	false
<input checked="" type="checkbox"/>	2660985	BN804564	Chen, Hu	false
<input checked="" type="checkbox"/>	2660985	BX50767	陈虎	false
<input type="checkbox"/>	2660985	<a href="http://viaf.org/viaf/206837881/">http://viaf.org/viaf/206837881/</a>	Chen, Hu	false
<input type="checkbox"/>	2660985	<a href="http://viaf.org/viaf/206837881/">http://viaf.org/viaf/206837881/</a>	陳瑚, (中國文學), 1875-1922	false
<input type="checkbox"/>	2660985	<a href="http://viaf.org/viaf/206837881/">http://viaf.org/viaf/206837881/</a>	Chen, Hu, 1875-1922	true

# How the record is 'used'

LEADER 01956nam a2200265Ma 4500

001 014538309-1

005 20151124104409.0

008 150528s2015 cc 000 0 chi d

020 \$a9787513034630

020 \$a751303463X

035 0 \$aocn910288776

040 \$aBCBTC\$beng\$cBCBTC

100 1 \$aChen, Hu.

**100 1 \$601\$a陈虎.**

245 10\$602\$aSi xing an jian zheng ming biao zhun yan jiu /\$cChen Hu zhu.

245 10\$602\$a死刑案件证明标准研究 /\$c陈虎著.

250 \$603\$aDi 1 ban.

250 \$603\$a第1版.

260 \$604\$aBeijing :\$bZhi shi chan quan chu ban she,\$c2015.

260 \$604\$a北京 :\$b知识产权出版社,\$c2015.

300 \$a1 v.

520 \$600\$a我国法定证明标准十分严格,但是在实践中却经常出现留有余地的判决,证明标准的表达与实践之间呈现出一种悖反关系.[cut]

In this case the underlined/red tag is used to retrieve a variant form of the name used as preferred.



## 2. Bibliographic records handling

The same process active for Agent cluster is executed for the pre-clusterization of Work title, using preferred access points such as tag 130, 240, 730 and so on. [3]

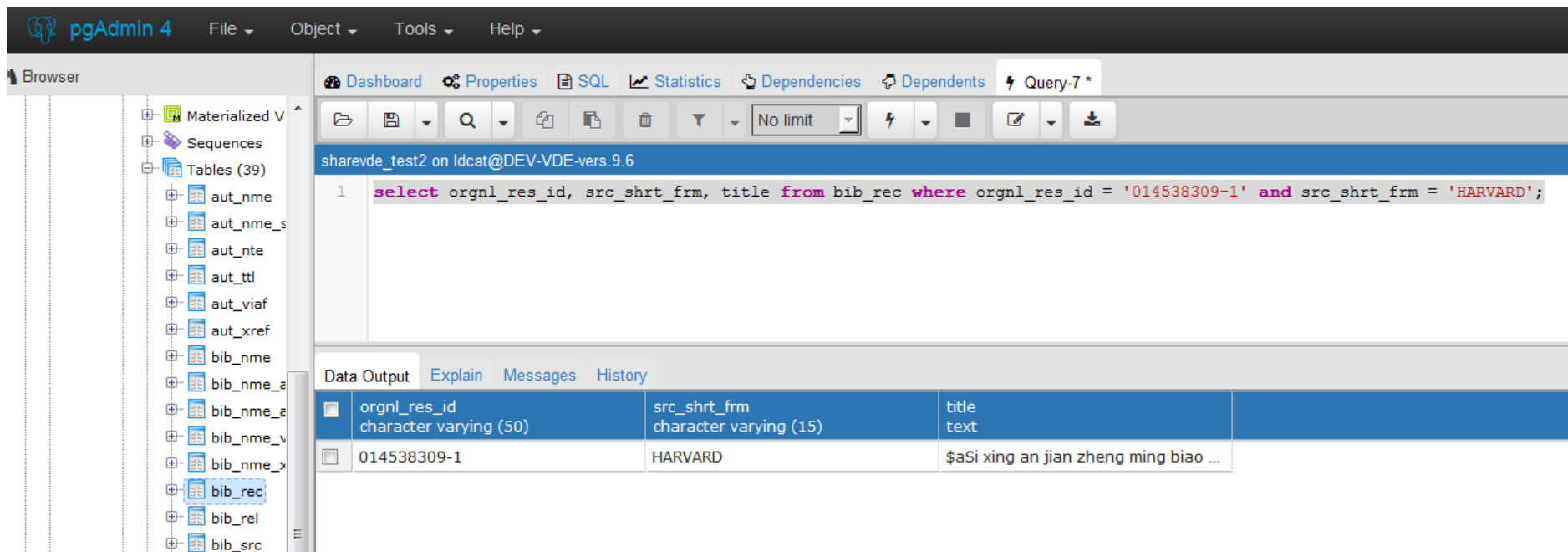
A new Work cluster (if it doesn't exist) is created, and it's ready to be fed with data coming from tag 245 (title of Instance).

[3] As in the previous clusterization process, the logic to select, analyse and create a cluster for a Work is out of scope of this document.



# 3. Creation/feeding of the Work title cluster

In our example, we'll use the title recorded as Instance or Manifestation record in tag 245: **Si xing an jian zheng ming biao zhun yan jiu**. It will produce the Work cluster ID 1196479



The screenshot shows the pgAdmin 4 interface. The left sidebar displays a tree view of database objects, with the 'bib\_rec' table selected. The main window shows a SQL query being executed in a query editor. The query is:

```
1 select orgnl_res_id, src_shrt_frm, title from bib_rec where orgnl_res_id = '014538309-1' and src_shrt_frm = 'HARVARD';
```

The results are displayed in a table with the following columns and data:

orgnl_res_id	src_shrt_frm	title
014538309-1	HARVARD	\$aSi xing an jian zheng ming biao ...

The title coming from tag 245 and recorded in PostgreSQL database





# 3. Creation/feeding of the Work title cluster

The cluster of Work with ID 1196479 in PostgreSQL

The screenshot shows the pgAdmin 4 interface. On the left, a tree view lists database objects, with 'tt1\_hdg' selected. The main pane displays a SQL query: `select * from tt1_hdg where clstr_id = 1196479;`. Below the query, the 'Data Output' tab shows the results of the query in a table format.

tt1_hdg_id	tt1_str_txt	clstr_id	hrchy_lv	match...	viaf_id	orgnl_res_id
integer	text	integer	character varying (1)	boolean	text	character varying
3983744	Si xing an jian zheng ming biao z...	1196479	2	false	[null]	014538309-1
3983743	Si xing an jian zheng ming biao z...	1196479	1	true	[null]	[null]



## 4. Marc 21 enrichment

At the end of the clusterization processes, the files containing bibliographic records must be enriched: it will be used to search for the entities on the end-users Portal (in the Instances layer) and to simplify the conversion into RDF through Lodify.

All records are enriched with the URIs obtained from the previous steps.

URIs are added to the specific tags identifying the entities (of type Agent, Work and the others we have not mentioned in this document)



# 4. Marc 21 enrichment

LEADER 01956nam a2200265Ma 4500

001 014538309-1

005 20151124104409.0

008 150528s2015 cc 000 0 chi d

020 \$a9787513034630

020 \$a751303463X

035 0 \$aocn910288776

040 \$aBCBTC\$beng\$cBCBTC

100 1 \$601\$aChen, Hu.\$1<http://share->

[vde.org/sharevde/rdfBibframe/Agent/2660985\\$0http://id.loc.gov/authorities/names/nr2001039650](http://sharevde.org/sharevde/rdfBibframe/Agent/2660985$0http://id.loc.gov/authorities/names/nr2001039650)

[\\$1http://www.isni.org/isni/0000000082319701](http://www.isni.org/isni/0000000082319701)[\\$1http://viaf.org/viaf/206837881/](http://viaf.org/viaf/206837881/)

100 1 \$601\$a陈虎.

245 10\$602\$aSi xing an jian zheng ming biao zhun yan jiu /\$cChen Hu zhu.

245 10\$602\$a死刑案件证明标准研究 /\$c陈虎著.

250 \$603\$aDi 1 ban.

250 \$603\$a第1版.

260 \$604\$aBeijing :\$bZhi shi chan quan chu ban she,\$c2015.\$9<http://share->

[vde.org/sharevde/rdfBibframe/Publisher/113862](http://sharevde.org/sharevde/rdfBibframe/Publisher/113862)

260 \$604\$a北京 :\$b知识产权出版社,\$c2015.

300 \$a1 v.

520 \$600\$a我国法定证明标准十分严格,但是在实践中却经常出现留有余地的判决,证明标准的表达与实践之间呈现出一种悖反关系.[cut]

912 10\$a014538309-1

996 10\$aSi xing an jian zheng ming biao zhun yan jiu\$9<http://share->

[vde.org/sharevde/rdfBibframe/Work/1196479](http://sharevde.org/sharevde/rdfBibframe/Work/1196479)

**The MARC 21 record enriched with URIs:  
the Marc record is ready now  
to be converted in BIBFRAME**

# The final result on the end users Portal

## The Agent cluster with ID 2660985

The screenshot displays the SHARE VDE portal interface. At the top, there is a navigation bar with the text "Person/Work" and a language selector set to "Italiano". Below this is a search bar with fields for "Work", "Publishers", and "Subject".

The main content area features a search result for the agent cluster with ID 2660985. The search criteria are "Person/Family/Corporate body". The result is displayed in a central panel with a red silhouette icon and a book icon. Below the icon, the name "Chen, Hu, 1875-1922" is listed with the ID "2660985". A "Works" section is also visible below the name.

On the left side, a box titled "This person in" lists various identifiers and logos, including ISNI, LIBRARY OF CONGRESS, WorldCat identities, and VIAF. On the right side, a box titled "Other name forms" lists several alternative names and identifiers, including "Chen, Hu", "Chen, Hu, 1875-1922", "陳湖, (中國文學), 1875-1922", and "陳湖".

At the bottom of the search result panel, there are two buttons: "SELECT ALL" and "UNCHECKED ALL". Below these buttons, there are two search filters: "Chinese 101a" and "Si xing an jian zheng ming biao zhun yan jiu".

[http://share-vde.org/sharevde/searchNames?n\\_cluster\\_id=2660985](http://share-vde.org/sharevde/searchNames?n_cluster_id=2660985)

# The final result on the end users Portal

The Work cluster with ID 1196479

The screenshot displays the SHARE VDE portal interface. At the top, there is a navigation bar with a logo featuring the word "SHARE" in a yellow circle, and other circles labeled "Discovery", "Virtual", and "Environment". The text "Person/Work" is written in a script font. A language selector "Italiano" is in the top right. Below the navigation bar are three search input fields labeled "Work", "Publishers", and "Subject".

The main content area is titled "Search Work/Publications" and includes "EXPAND ALL" and "CLOSE ALL" buttons. On the right, there are icons for RDF, print, and email. A search result is shown in a box with a "Publications" dropdown and the text "Si xing an jian zheng ming biao zhun yan jiu" with a small crest icon.

Below this, a larger box contains a stack of books icon, the text "Si xing an jian zheng ming biao zhun yan jiu", and "ID: 1196479". A vertical line labeled "Creators:" points down to a search box containing "(2660985) Chen, Hu, 1875-1922".

[http://share-vde.org/sharevde/searchTitles?t\\_cluster\\_id=1196479](http://share-vde.org/sharevde/searchTitles?t_cluster_id=1196479)

# From Marc 21 to BIBFRAME: the conversion process



# The starting point: the LOC Marc-to-BF mapping

The conversion from Marc 21 to BIBFRAME starts from the Library of Congress mapping rules. This official source is enlarged/modified to manage some special requirements of different projects.

Here, as example, the MARC tag 300 – Physical description – processing:

	A	B	C	D	E
1	Fields 3XX - Physical Description, etc. - R1, 06/07/2017	Conversion 1			
2					
3	300 - PHYSICAL DESCRIPTION (R)				
4	Indicators				
5	Subfield Codes				
6	\$a - Extent (R)	I - extent - Extent - Make a string of all a, f, g and including all repeats of subfields. Keep in order as in field, insert blank for each subfield code then put this string in rdfs:label			
7	\$b - Other physical details (NR)	I - note - Note - noteType "Physical details" ; rdfs:label "\$b content"			
8	\$c - Dimensions (R)	I - dimensions - literal			
9	\$e - Accompanying material (NR)	I - note - Note - noteType "Accompanying material" ; rdfs:label "\$e content"			
10	\$f - Type of unit (R)	See \$a			
11	\$g - Size of unit (R)	See \$a			
12	\$3 - Materials specified (NR)	See Subfield \$3 spec			
13	\$6 - Linkage (NR)	See Subfield \$6 spec			
14					
15	306 - PLAYING TIME (NR)	I - duration			
16	Indicators				
17	Subfield Codes				
18	\$a - Playing time (R)	## - literal			
19	\$6 - Linkage (NR)	See Subfield \$6 spec			
20					

# The starting point: the LOC Marc-to-BF mapping

And here, as example, the MARC tag 100 – Main Entry-Personal Name – processing:

## 0.3) Basic RDF Patterns for Names, Titles, and Relationships

### 0.3.1) RDF for names

```
<resource>  bf:contribution [ a      bf:Contribution ;
                    bf:agent  [ a      bf: Person, Organization, etc.
                                rdfs:label      "label from Process 1.3";
                                identifiedBy [a      Identifier ..... ];  see Subfield $0 spec
                                bflc:nameXXMatchKey  "string from Process 1.1";
                                bflc:nameXXMarckKey   "string from Process 1.2" ] ;
                    bf:role  [ a  bf:Role
                                [rdfs:label  "..."] ;                               see Process 1.4
                                bf:code      "..."] ].                             see Process 1.4
```

If URI from ID for role, then instead:

```
bf:role      URI for role
```

If name is from 1XX:

Use bflc:PrimaryContribution instead of bf:Contribution

Also add (needed?)

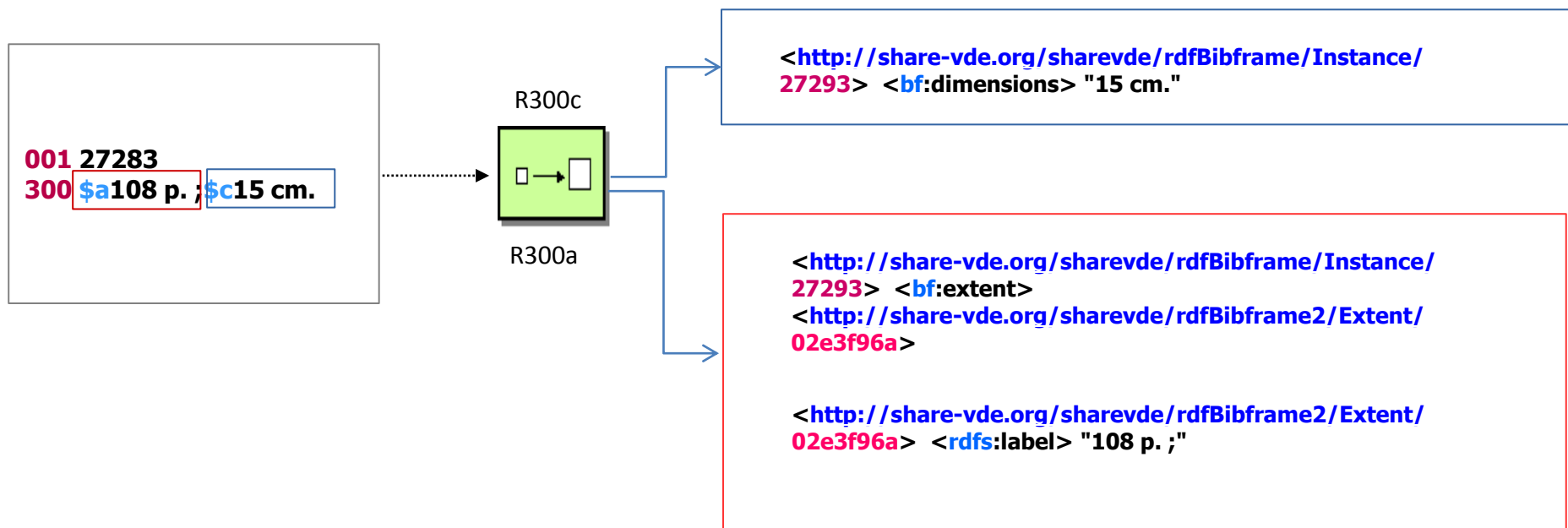
```
<resource>  bflc:primaryContributorNameXXMatchKey "string from Process 1.1"
```



# Lodify - Conversion templates

**Lodify** converts each incoming record by means of Conversion templates. Each template associates:

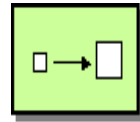
- a MARC record belonging to the incoming data-stream
- with a set of (conversion) rules associated with BIBFRAME vocabulary (here, two rules for tag 300a and tag 300c)



# Lodify - Conversion templates

**100 10\$a**Castro, Juan Antonio.**\$1**<http://share-vde.org/sharevde/rdfBibframe/Agent/1801277>**\$0**<http://id.loc.gov/authorities/names/n88638722>**\$1**<http://www.isni.org/isni/0000000059264386>**\$1**<http://www.wikidata.org/entity/Q21480742>**\$1**<http://viaf.org/viaf/63125720/>

(here, the rule for tag 100a subfields 0/1)



R100a01

**<**<http://share-vde.org/sharevde/rdfBibframe/Agent/1801277>**>** **<**<http://www.w3.org/1999/02/22-rdf-syntax-ns#type>**>**  
**<**bf:Person**>** .

**<**<http://share-vde.org/sharevde/rdfBibframe/Agent/1801277>**>** **<**rdfs:label**>** "Castro, Juan Antonio." .

**<**<http://share-vde.org/sharevde/rdfBibframe/Agent/1801277>**>** **<**bfcl:name00MatchKey**>** "Castro, Juan Antonio." .

**<**<http://share-vde.org/sharevde/rdfBibframe/Agent/1801277>**>** **<**bfcl:name00MarcKey**>** "10010\$aCastro, Juan Antonio.  
\$1<http://share-vde.org/sharevde/rdfBibframe/Agent/1801277> \$0<http://id.loc.gov/authorities/names/n88638722>  
\$1<http://www.isni.org/isni/0000000059264386> \$1<http://www.wikidata.org/entity/Q21480742>  
\$1<http://viaf.org/viaf/63125720/>" .

**<**<http://share-vde.org/sharevde/rdfBibframe/Agent/1801277>**>** **<**bfcl:primaryContributorName00MatchKey**>** "Castro,  
Juan Antonio." .

## Lodify – Marc-to-BF logic

Here we show (with two examples) how SHARE-VDE converts MARC21 bibliographic tags into set of BIBFRAME triples.

The first section of each tag shows pseudo-triples (to be more readable as possible for human eye).

The second section shows concrete examples.

For the SHARE-VDE project, all the mapping and conversion topics are analyzed and discussed in one of the

**SHARE-VDE Working Group**, the **Transformation Council**

# Lodify – Marc-to-BF logic

## How the tag 500 – General note, is mapped and managed

### a. Tag 500 (General Note)

Connected to instance record:

\$a

```
<Instance>  bf:note      <_node>
<_node>     rdf:type     bf:Note
<_node>     rdfs:label  "$a"
```

500 \$a Közrem. Simon Zoltán, Fülöp András

```
<http://share-vde.org/sharevde/rdfBibframe2/Instance/YALE56456> bf:note
```

```
<http://share-vde.org/sharevde/rdfBibframe2/Note/487cb467-d631-37bf-a94e-f5f7243b46ee>.
```

```
<http://share-vde.org/sharevde/rdfBibframe2/Note/487cb467-d631-37bf-a94e-f5f7243b46ee> rdf:type bf:Note .
```

```
<http://share-vde.org/sharevde/rdfBibframe2/Note/487cb467-d631-37bf-a94e-f5f7243b46ee> rdfs:label "Közrem. Simon Zoltán, Fülöp András".
```

# Lodify – Marc-to-BF logic

## How the tag 020 – ISBN, is mapped and managed

b. Tag 020 (international Standard book number ISBN)

Connected to instance record:

\$a

```
<Instance> bf:identifiedBy <_node>  
<_node> rdf:type bf:isbn  
<_node> rdf:value "$a"
```

\$c

```
<Instance> bf:acquisitionTerms "$c"
```

020 [\\$a](#)978-963-88816-2-8 [\\$c](#)3000,-Ft

```
<http://share-vde.org/sharevde/rdfBibframe2/Instance/YALE43570> bf:identifiedBy
```

```
<http://share-vde.org/sharevde/rdfBibframe2/Isbn/d859f1c4-e626-3a42-bd2d-8e78a84a50bb>
```

```
<http://share-vde.org/sharevde/rdfBibframe2/Isbn/d859f1c4-e626-3a42-bd2d-8e78a84a50bb> rdf:type bf:ISBN .
```

```
<http://share-vde.org/sharevde/rdfBibframe2/Isbn/d859f1c4-e626-3a42-bd2d-8e78a84a50bb> rdf:value "978-963-88816-2-8" .
```

```
<http://share-vde.org/sharevde/rdfBibframe2/Instance/YALE43570> bf:acquisitionTerms "3000,-Ft" .
```

# MARC-to-BF conversion process Troubleshooting



# Marc-to-BF troubleshooting

## Transliterations: how to manage them (1/6)

Record example: <http://id.loc.gov/tools/bibframe/compare-id/full-ttl?find=12069992>

```
MARC
01468cam a2200349 a 4500
001 12069992
005 20091016083942.0
008 000621s1998 is b 000 0 heb d
010 $a 00015646
035 $(C$TRLIN)DCLH00-B1634
040 $aN$N$C$N$J$S$dCtY$dDLC-R$dDLC
042 $alccopycat
050 00 $aBM720.T4$bF73 1998
066 $c2
100 1 $6880-01$aFraiman, Hayim.
240 10 $aKitsur dine terumot u-ma'ašerot
245 10 $6880-02$aSefer Kitsur dine terumot u-ma'ašerot /$cme-et Hayim ben Yišra'el
iman.
246 14 $6880-03$aKitsur dine terumot u-ma'aserot
260 $6880-04$aBene-Beraḳ :$bMishor,$c759 [1998 or 1999]
300 $a146 p. ;$c24 cm.
504 $aIncludes bibliographical references.
650 0 $aTerumah.
650 0 $aTithes (Jewish law)
880 1 $6100-01/(2/r$aמאיר בן ישראל חיים בן פריימן.
880 10 $6245-02/(2/r$aמאת חיים בן ישראל מאיר פריימן/$cספר קיצור דיני תרומות ומעשרות
880 14 $6246-03/(2/r$aקיצור דיני תרומות ומעשרות
880 0 $6246-04/(2/r$aמאת חיים בן ישראל מאיר פריימן/$cמאת חיים בן ישראל מאיר פריימן
```

```
http://bibframe.example.org/12069992#Agent100-14> a bf:Agent,
bf:Person ;
rdfs:label "Fraiman, Ḥayim." ;
bflc:name00MarcKey "1001 $6880-01$aFraiman, Ḥayim." ;
bflc:name00MatchKey "Fraiman, Ḥayim." ;
bflc:primaryContributorName00MatchKey "Fraiman, Ḥayim." .
```

```
<http://id.loc.gov/authorities/names/n88237680> a bf:Agent,
bf:Person ;
rdfs:label "פריימן, חיים בן ישראל מאיר." ;
bflc:name00MarcKey "8801 $6100-01/(2/r$aמאיר בן ישראל חיים בן ישראל מאיר פריימן." ;
bflc:name00MatchKey "פריימן, חיים בן ישראל מאיר." ;
bflc:primaryContributorName00MatchKey "פריימן, חיים בן ישראל מאיר." .
```

# Marc-to-BF troubleshooting

## Transliterations: how to manage them (2/6)

Record example: <http://id.loc.gov/tools/bibframe/compare-id/full-ttl?find=12069992>

```
http://bibframe.example.org/12069992#Agent100-14> a bf:Agent,  
    bf:Person ;  
rdfs:label "Fraiman, Ḥayim." ;  
bflc:name00MarcKey "1001 $6880-01$aFraiman, Ḥayim." ;  
bflc:name00MatchKey "Fraiman, Ḥayim." ;  
bflc:primaryContributorName00MatchKey "Fraiman, Ḥayim." .
```

```
<http://id.loc.gov/authorities/names/n88237680> a bf:Agent,  
    bf:Person ;  
rdfs:label "פריימן, חיים בן ישראל מאיר." ;  
bflc:name00MarcKey "8801 $6100-01/(2/r$aמאיר בן ישראל חיים בן ישראל מאיר." ;  
bflc:name00MatchKey "פריימן, חיים בן ישראל מאיר." ;  
bflc:primaryContributorName00MatchKey "פריימן, חיים בן ישראל מאיר." .
```

The record has a tag 100 and its transliteration in the 880.

The LoC conversion tool translates the two tags separately. A NAF URI is attributed to the 880 section of code; a specifically created URI is attributed to the 100 section.

The two entities are not reconciled but they are the same Person.



# Marc-to-BF troubleshooting

## Transliterations: how to manage them (3/6)

Record example: <http://id.loc.gov/tools/bibframe/compare-id/full-ttl?find=12069992>

```
<http://id.loc.gov/authorities/names/n88237680> a bf:Agent,  
    bf:Person ;  
    rdfs:label "פריימן, מאיר בן ישראל חיים." ;  
    bflc:name00MarcKey "8801 $6100-01/(2/r$aפריימן, מאיר בן ישראל חיים." ;  
    bflc:name00MatchKey "פריימן, מאיר בן ישראל חיים." ;  
    bflc:primaryContributorName00MatchKey "פריימן, מאיר בן ישראל חיים." .
```

```
<http://bibframe.example.org/12069992#Agent100-14> a bf:Agent,  
    bf:Person ;  
    rdfs:label "Fraiman, Ḥayim." ;  
    bflc:name00MarcKey "1001 $6880-01$aFraiman, Ḥayim." ;  
    bflc:name00MatchKey "Fraiman, Ḥayim." ;  
    bflc:primaryContributorName00MatchKey "Fraiman, Ḥayim." .
```

# Marc-to-BF troubleshooting

## Transliterations: how to manage them (4/6)

Record example: <http://id.loc.gov/tools/bibframe/compare-id/full-ttl?find=12069992>

```
<http://id.loc.gov/authorities/names/n88237680> a bf:Agent,  
    bf:Person ;  
    rdfs:label "חיים בן ישראל מאיר, פריימן." ;  
    bflc:name00MarcKey "8801 $6100-01/(2/r$aפריימן, מאיר, חיים בן ישראל." ;  
    bflc:name00MatchKey "פריימן, מאיר, חיים בן ישראל." ;  
    bflc:primaryContributorName00MatchKey "חיים בן ישראל, מאיר, פריימן." .
```

```
<http://id.loc.gov/authorities/names/n88237680> a bf:Agent,  
    bf:Person ;  
    rdfs:label "Fraiman, Ḥayim." ;  
    bflc:name00MarcKey "1001 $6880-01$aFraiman, Ḥayim." ;  
    bflc:name00MatchKey "Fraiman, Ḥayim." ;  
    bflc:primaryContributorName00MatchKey "Fraiman, Ḥayim." .
```

Hypothesis 1: to reconcile the two sections using the same NAF URI.  
In this way it is not possible to link the two labels that have been generated to their original MARC tag.  
In a conversion process from BF to MARC21 it will not be possible to distinguish which tag generated which label.

# Marc-to-BF troubleshooting

## Transliterations: how to manage them (5/6)

Record example: <http://id.loc.gov/tools/bibframe/compare-id/full-ttl?find=12069992>

```
<http://bibframe.example.org/12069992#Agent880-14> a bf:Agent,  
    bf:Person ;  
    rdfs:label "פריימן, מאיר בן ישראל חיים." ;  
    bflc:name00MarcKey "8801 $6100-01/(2/r$aפריימן, מאיר בן ישראל חיים." ;  
    bflc:name00MatchKey "פריימן, מאיר בן ישראל חיים." ;  
    bflc:primaryContributorName00MatchKey "פריימן, מאיר בן ישראל חיים." .  
    owl:sameAs <http://id.loc.gov/authorities/names/n88237680>
```

```
<http://bibframe.example.org/12069992#Agent100-14> a bf:Agent,  
    bf:Person ;  
    rdfs:label "Fraiman, Ḥayim." ;  
    bflc:name00MarcKey "1001 $6880-01$aFraiman, Ḥayim." ;  
    bflc:name00MatchKey "Fraiman, Ḥayim." ;  
    bflc:primaryContributorName00MatchKey "Fraiman, Ḥayim." .  
    owl:sameAs <http://id.loc.gov/authorities/names/n88237680>
```

Hypothesis 2: each section has a specifically created URI. The two URIs will be linked to the NAF URI with a <same as>

# Marc-to-BF troubleshooting

## Transliterations: how to manage them (6/6)

Record example: <http://id.loc.gov/tools/bibframe/compare-id/full-ttl?find=12069992>

```
<http://id.loc.gov/authorities/names/n88237680>
  bf:nameMatch [ a bf:NameMatch ;
    rdfs:label "פריימן, מאיר בן ישראל חיים." ;
  bflc:name00MarcKey "8801 $6100-01/(2/r$aפריימן, מאיר בן ישראל חיים." ;
  bflc:name00MatchKey "פריימן, מאיר בן ישראל חיים." ;
  bflc:primaryContributorName00MatchKey "פריימן, מאיר בן ישראל חיים." ] .

  bf:nameMatch [ a bf:NameMatch ;
    rdfs:label "Fraiman, Ḥayim." ;
  bflc:name00MarcKey "1001 $6880-01$aFraiman, Ḥayim." ;
  bflc:name00MatchKey "Fraiman, Ḥayim." ;
  bflc:primaryContributorName00MatchKey "Fraiman, Ḥayim." ] .
```

Hypothesis 3: to use a set of classes and properties specific to this situation (not in the vocabulary at the present)

# Marc-to-BF troubleshooting

## Relationships of Person with Work or with Instance (tag 700/710/711) (1/2)

Record example: <http://id.loc.gov/tools/bibframe/compare-ccn/full-ttl?find=2011029677>

```
<http://bibframe.example.org/16901812#Work> a bf:Text,  
bf:Work ;
```

```
rdfs:label "Hippolytus" ;
```

```
bf:contribution
```

```
[ a bf:Contribution ;
```

```
bf:agent
```

```
<http://id.loc.gov/authorities/names/nr90011474> ;
```

```
bf:role <http://id.loc.gov/vocabulary/relators/edt> ,
```

```
<http://id.loc.gov/vocabulary/relators/trl> ] ;
```

MARC

```
010 0a 2011029677  
020 $a9783110188172 (hardcover : alk. paper)  
020 $z9783110215397 (e-ISBN: PDF)  
020 $z9783110368321 (e-ISBN: EPUB)  
040 $aDLC$beng$cDLC$erda$dDLC  
041 1 $ager$agrc$hgrc  
042 $apcc  
050 00 $aPA3976.G5$bH57 2015  
082 00 $a882/.01$223  
100 0 $aEuripides,$author.  
240 10 $aHippolytus  
245 10 $aHippolytos /$cEuripides ; herausgegeben, übersetzt und kon  
264 1 $aBerlin ;$aBoston :$bDe Gruyter,$c[2015]  
300 $ax, 392 pages ;$c24 cm.  
336 $atext$2rdacontent  
337 $aunmediated$2rdamedia  
338 $avolume$2rdacarrier  
490 0 $aGriechische Dramen  
504 $aIncludes bibliographical references (pages 387-392), appen  
600 00 $aEuripides$vTranslations into German.  
600 00 $aHippolytus$c(Mythological character)$vDrama.  
700 1 $aRoth, Peter,$d1957-$eeditor,$etranslator.  
700 0 $aEuripides.$tHippolytus.$lGerman.  
906 $a7$bcbc$corianew$d1$eeciio$f20$av-aencatla
```

# Marc-to-BF troubleshooting

## Relationships of Person with Work or with Instance (tag 700/710/711) (2/2)

Record example: <http://id.loc.gov/tools/bibframe/compare-ccn/full-ttl?find=2011029677>

The translator should indicate a contribution to the Instance and not the Work, but it appears that is attributed to the Work by default.  
We would expect something like this:

```
<http://bibframe.example.org/16901812#Instance>
```

```
  bf:Instance ;
```

```
  bf:contribution
```

```
    [ a bf:Contribution ;
```

```
      bf:agent <http://id.loc.gov/authorities/names/nr90011474> ;
```

```
      bf:role <http://id.loc.gov/vocabulary/relators/edt> ,
```

```
        <http://id.loc.gov/vocabulary/relators/trl> ] ;
```

# Marc-to-BF troubleshooting

## Source for tags 336/337/338

Record example: <http://id.loc.gov/tools/bibframe/compare-id/full-ttl?find=18345151>

### MARC

```
01769cam a2200409 i 4500
001 18345151
005 20170417074850.0
008 141023t20152009meu d 000 j eng
010 $a 2014041792
020 $a9781628994476 (large print : hardcover : alk. paper)
020 $a9781628994520 (large print : pbk. : alk. paper)
040 $aDLC$beng$erda$cDLC
042 $apcc
050 10 $aPS3569.A826$bA6 2015
082 00 $a813/.54$223
100 1 $aSavage, Les,$author.
240 10 $aShort stories.$kSelections
245 10 $aLong gun :$ba western quartet /$cLes Savage, Jr.
250 $aCenter Point Large Print edition. Large print edition.
264 1 $aThorndike, Maine :$bCenter Point Large Print,$c2015.
264 4 $c@2009
300 $a303 pages (large print) ;$c23 cm.
336 $atext$btxt$2rdacontent
337 $aunmediated$bnc$2rdamedia
338 $avolume$bnc$2rdacarrier
490 0 $aCenter Point Large Print edition western
505 0 $aLong Gun -- Traitor town -- Blood star over Santa Fe --
```

### LC BIBFRAME TOOL:

```
<http://id.loc.gov/vocabulary/carriers/nc> a
bf:Carrier ;
  rdfs:label "volume" ;
  bf:source [ a bf:Source ;
    rdfs:label "rdacarrier" ] .
```

### LODIFY:

```
<http://rdaregistry.info/termList/RDACarrierType
/1049> a bf:Carrier ;
  rdfs:label "volume" .
```

```
<http://id.loc.gov/vocabulary/carriers/nc> a
bf:Carrier ;
  rdfs:label "volume" .
```

# Marc-to-BF troubleshooting

## How to enrich a N/T heading (tag 100 \$a \$t)

Record example: <http://share-vde.org/sharevde/resource?uri=NLM589892&&dcnr=9&v=m21>

LEADER 01689cam a2200385 a 4500

001 589892

005 20170701013644.0

008 840712s1985 gw |||| |||| 00||0|eng

010 \$a84-14041

020 \$a9780387132570 (U.S.)

020 \$a0387132570 (U.S.)

035 \$98408228

040 \$aDNLM\$cDNLM

041 09\$aeng

044 \$9Germany

060 00\$aWE 860\$bM297b 1985

100 1 \$aMaquet, Paul G. J.,\$d1928-\$tBiomécanique du genou.\$lEnglish.

245 10\$aBiomechanics of the hip :\$bas applied to osteoarthritis and related conditions /\$cPaul G.J. Maquet ; with a foreword by William H. Harris.

260 \$aBerlin ; New York :\$bSpringer-Verlag,\$c1985.\$9<http://share-vde.org/sharevde/rdfBibframe/Publisher/115460>

300 \$axiii, 309 p. :\$bill.

How can a tag 100 containing a \$t be enriched unequivocally ?

This tag may generate two URIs: one for the name and one for the name/title.

How can each be distinguished from the other?

The same problem is encountered when creating RDF from an enriched MARC21



# SHARE-VDE: Phase 2 deliverables overview



BIB1



BIB2



BIB...

**Deliverable 1:** The datasets in BIBFRAME 2.0 of the entire catalogue of each institution with the "tuples" derived directly from the MARC records, delivered both as triples and as quadruples with the addition of the provenance and with Share-VDE URIs



Clusters Knowledge base

**Deliverable 2:** The knowledge base of clusters accessible in RDF



BIB1



BIB2



BIB...

**Deliverable 3:** The datasets in BIBFRAME 2.0 for each institution with the triples that include the URIs from the external sources

External (VIAF) URIs

SHARE-VDE URIs



Marc enriched  
(Binary)  
(one for LIB)

**Deliverable 4:** The MARC21 records for each institution enriched with URIs



European BIBFRAME Workshop  
Tutorial - 17 September 2018  
Firenze

And now ..... questions!!!

Tiziana Possemato  
Casalini Libri - @Cult

