

IFLA-Satellite-Meeting 2017  
**Digital Humanities. Connecting Libraries and Research**  
Berlin, August 16th 2017

## **BIBFRAME and Linked Data practices for the stewardship of research knowledge**

Michele Casalini  
*Managing Director*  
Casalini Libri



# Index

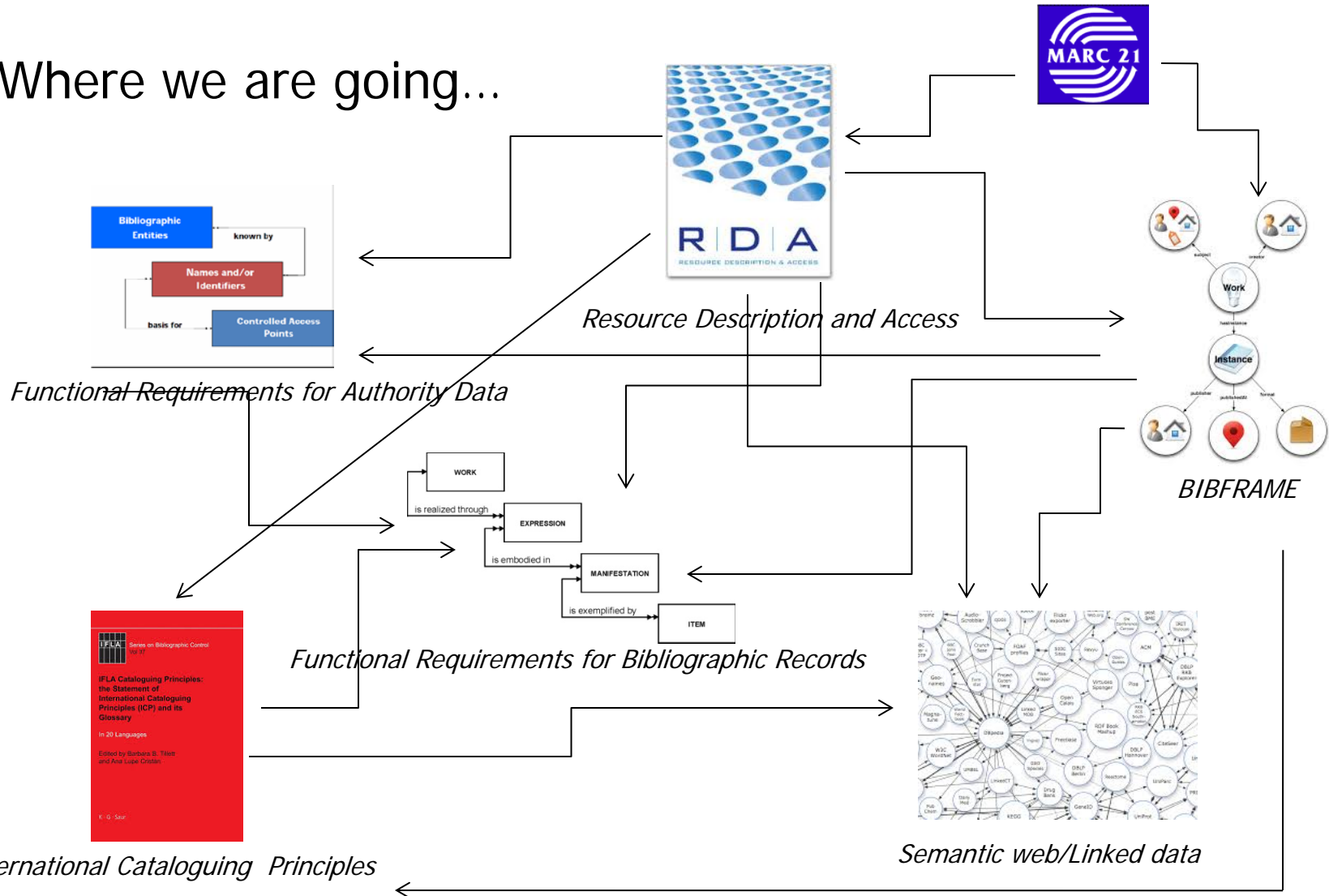
1. BIBFRAME and Linked Data
2. Experience
3. BIBFRAME and Linked Data practices: the SHARE-VDE project
  - 3.1 Project highlights
  - 3.2 Entity identification, reconciliation and data enrichment
  - 3.3 Enhanced MARC records with URIs
  - 3.4 Conversion of authority and bibliographical data in BIBFRAME
  - 3.5 Publication of a BIBFRAME three layered platform prototype
  - 3.6 Further connected topics
4. Conclusions
5. Resources and links

# 1. BIBFRAME and Linked Data

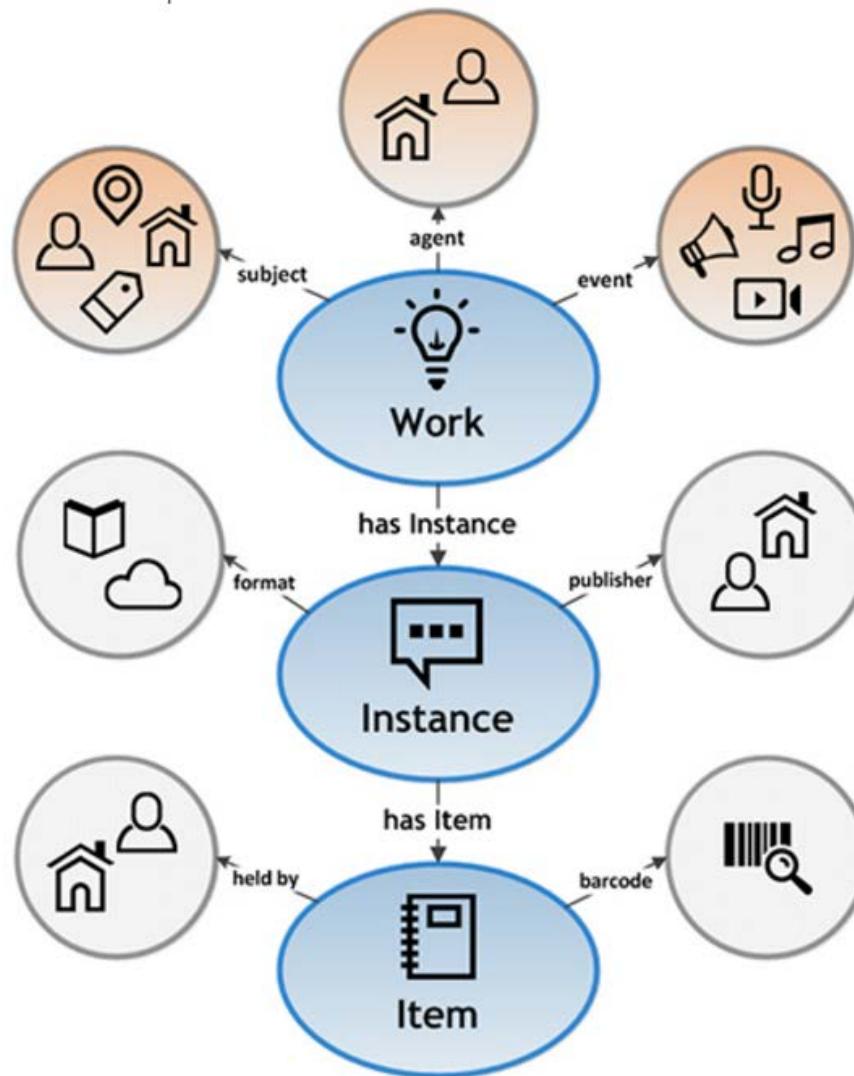


# Brief theoretical context

Where we are going...



# BIBFRAME 2.0 data model



## 2. Experience



# Experience

**Casalini Libri** produces, for publications from Romance language countries, more than 40,000 original bibliographic records in RDA as a member of the Program for Cooperative Cataloguing (PCC) with authority entries;

Bibliographic records are created using the @Cult **OLISuite WeCat** cataloguing modules;

**@Cult**, in addition to the LMS and Discovery tools field, is specialized in the development of software components and platforms to convert, enrich, reconcile and publish data of cultural institutions under the linked data paradigm;

**ALIADA** project, co-financed by the European Union in 2013-2015, originally applied the Linked Data paradigm using FRBRoo based ontologies.



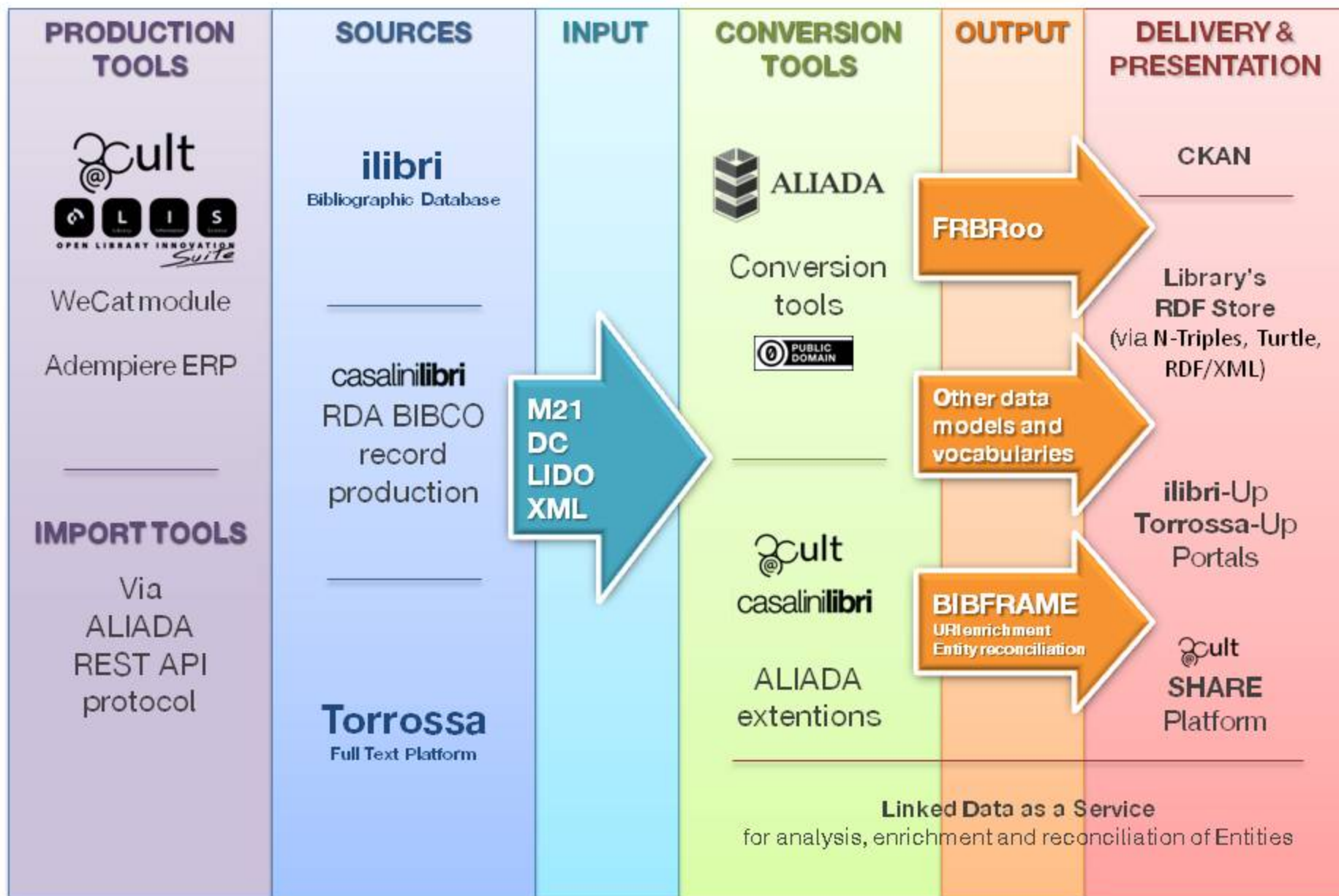
## 3.1 BIBFRAME and Linked Data practices: the SHARE-VDE project

### Project highlights





# Casalini's BIBFRAME Conversion, Distribution & Publication Options



# SHARE-VDE approach

Among the guideline of the initiative:

- SHARE-VDE is a Research & Development project;
- The initiative is library community driven;
- The project aim is also to identify issues and problems related to new information management processes, and propose solutions;
- The components are designed to be available individually in as flexible as possible configuration for the library community's needs;
- Scalability will be tested up to the equivalent of over 100 million traditional bibliographic and authority records.

SHARE-VDE is a Casalini Libri and @Cult partnership project, under the direction of Tiziana Possemato.



# SHARE-VDE overall goals

The overall goals of the project are:

- To reach through short phases concrete results on which the library community can base further steps and decisions;
- To discuss, experiment and configure, together with the library community, the options for the future data creation, enhancement and sharing of all type of resources;
- To achieve a granularity and richness of relationships within library data and beyond, which is often hidden or unexpressed in a traditional catalogue;
- To prepare an environment which is useful for both library patrons - who can take advantage of more advanced discovery interfaces -, and librarians looking for cataloguing functions in native semantic web standards together with the integration of processes with the local systems and the implementation of tools in a shared collaborative environment.



# SHARE-VDE context

The development of this initiative started taking into account the following context:

- Libraries with different systems, habits and cataloguing rules and traditions;
- The emerging Linked Data paradigm;
- The broader library, archive and museum community expectations;
- The opportunity to provide more comprehensive access to researchers and students;
- The consciousness of cultural and language diversities, and the large differences among the needs of disciplines.



# SHARE Virtual Discovery Environment project

The project is divided into **three phases**. Each participant decides whether or not to take part in the subsequent phases.

**Phase 1** activities included: the analysis, enrichment, reconciliation and conversion into RDF of two sets of bibliographic data for each participating library (1985 and 2015 imprint titles); the publication of the first version of the platform applying the new data model; the release of test Marc records enriched with URIs and the BIBFRAME 1.0 datasets for each participating library.

A total of 2,308,204 bibliographical records and 3,601,327 authority records were converted into BIBFRAME 2.0 and published via the first release of the SHARE-VDE portal ([www.share-vde.org](http://www.share-vde.org)).

Phase 1: from October 2016 to January 2017.



# SHARE Virtual Discovery Environment project

**Phase 2** activities includes:

The entire library catalogue of each participating institution will be converted into BIBFRAME 2.0 and returned to each single institution (over 100 million records are processed).

The second version of the platform will be released focusing also on the Linked Data oriented record-less approach.

A relationship database that registers the relationships between entities (person, work, instances, subjects, publisher, etc...) will be established in order to assure a more precise identification rate of each entity to reach a higher quality of results without human intervention.

Refinement of data, e.g. for co-authors and editors, where there is a variety of ways in which they are identified in library records (Relator terms topic).



# SHARE Virtual Discovery Environment project

... cont. (**Phase 2** activities includes):

Export of data in Marc or RDF format filtering the library preferred URIs.

Inclusion of additional URI sources, e.g. specific sources for corporate bodies, subjects (LCSH, FAST, etc...) and RDA vocabularies.

Analysis for the creation of relationships among subject terms and strings in different languages.

Provenance declaration, update management and built-in instances will be addressed.

Phase 2: from March to September/December 2017.

**Phase 3** will be a production phase shaped on the library community input.



# Participating libraries (1)

Phase 1	Phase 2	(in Country/State order):
X	X	Stanford University
X	X	University California Berkeley
X	X	Yale University
X	X	Library of Congress
X	X	University of Chicago
X	X	University of Michigan Ann Arbor
X	X	Harvard University
X		Massachusetts Institute of Technology
	X	Duke University
X		Cornell University
X		Columbia University
X	X	University of Pennsylvania





## Participating libraries (2)

Phase 1	Phase 2	(in Country/State order):
	x	Pennsylvania State University
x	x	Texas A&M University
	x	University of Alberta
x		University of Toronto



## 3.2. BIBFRAME and Linked Data practices: the SHARE-VDE project

Entity identification, reconciliation  
and data enrichment



# 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 BTSTA  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 [online]

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

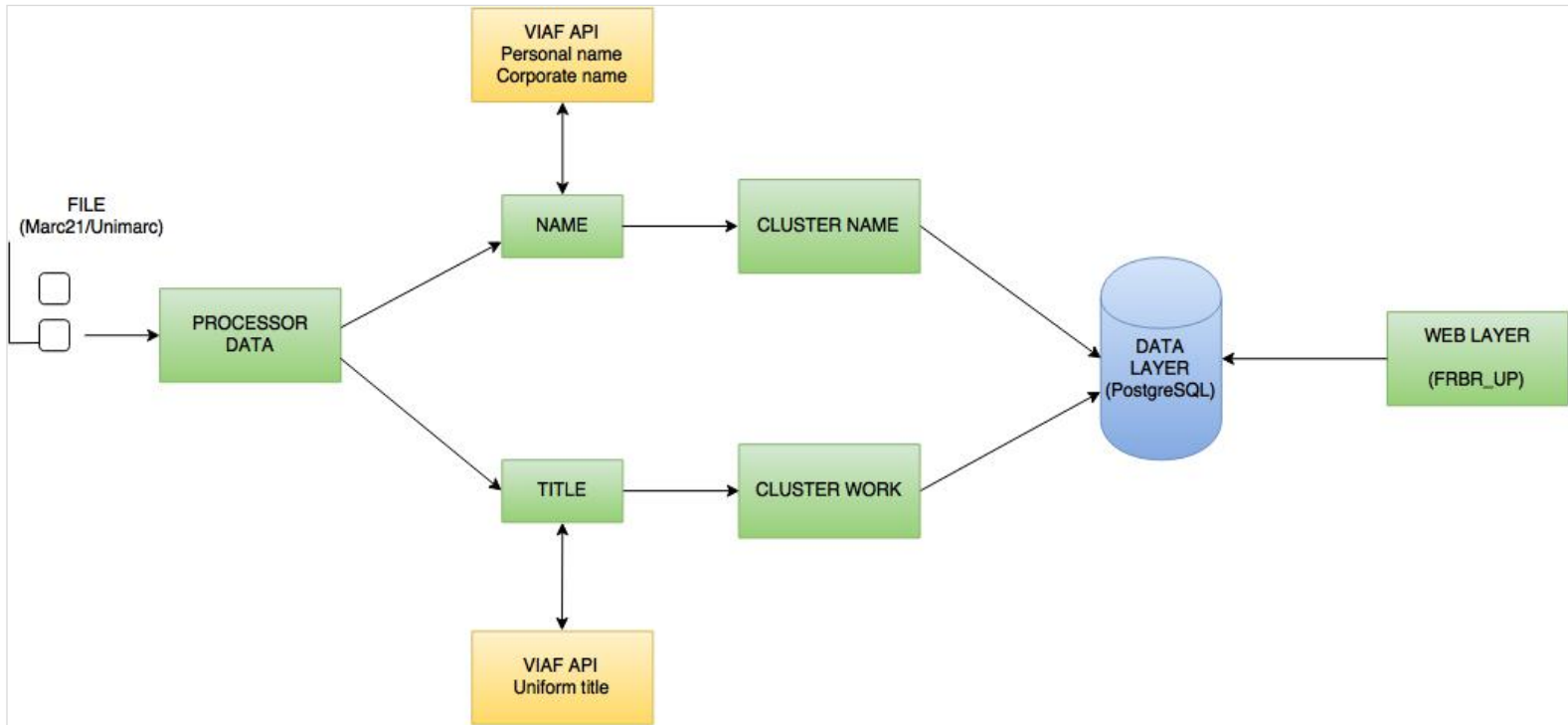


Subject

Padri e figlie

Pères et filles

# Authify - the tool to obtain precise URI retrieval



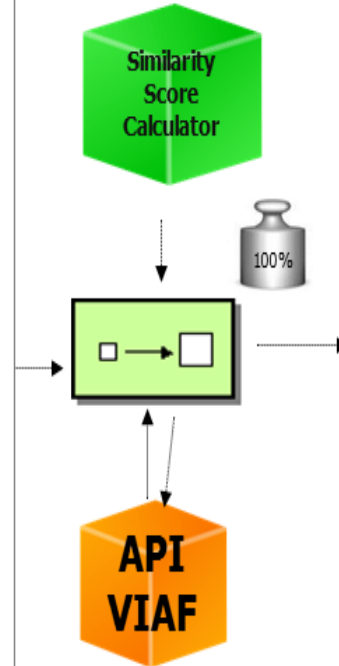
# Automated process of cluster creation example

001 00001  
200 \1 \$aEco,\$bUmberto\$f<1932- >.  
400 \0 \$aDedalus  
997 \\ \$aAUTHORITY

001 27283  
700 \1 \$aEco,\$bUmberto\$f<1932- >.  
997 \\ \$aUNINA

001 7258  
700 \1 \$aECO,\$bUmberto  
997 \\ \$aUNISA

001 7258  
700 \1 \$aEco,\$bUmberto  
997 \\ \$aUNIBAS



ID cluster: 1281829  
Author : Eco, Umberto, 1932-...  
VIAF ID: 108299403

## Other forms:














ECO, Umberto  
Eco, Umberto  
Dedalus  
1932- أمبيرتو-  
إكو، أمبرتو.  
Eco, Umberto (Italian writer,  
architect, and critic, born 1932)  
Umberto Eco semiologo,  
filosofo e scrittore italiano  
1932, אומברטו, אקו-

# Manual entity enrichment example

(screenshot from the URI Management System in OLSuite)

Selected heading: Kafka, Franz, 1883-1924

New

Source	Http Uri		Validated	Options
NAF	<a href="http://id.loc.gov/authorities/names/n81063091">http://id.loc.gov/authorities/names/n81063091</a>	<input type="checkbox"/>	✗	   
ISNI	<a href="http://isni.org/isni/0000000012280370X">http://isni.org/isni/0000000012280370X</a>	<input type="checkbox"/>	✗	   
VIAF	<a href="http://viaf.org/viaf/56611857/">http://viaf.org/viaf/56611857/</a>	<input type="checkbox"/>	✗	   

Delete

Cataloguers can check, modify, delete or add other identifiers to the same heading



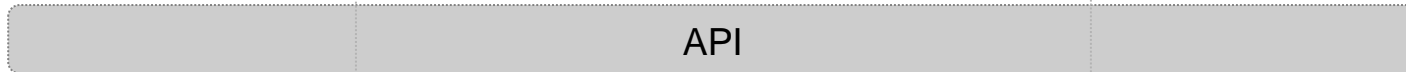
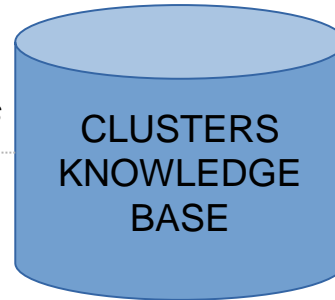


*Injection services (massive)*

*Injection services (massive)*

*Cluster search services*

*Injection services (single cluster)*



GET

PUT

/names

/corporates

/works

/people

/relatorTerms

/cluster/new



### **3.3. BIBFRAME and Linked Data practices: the SHARE-VDE project**

**Enhanced MARC records with URIs**





# Enhanced MARC records with URIs

In the recent months the following important steps were achieved by the PCC Task Group on URISs in MARC:

- Redefining Subfield \$4 to Encompass URIs for Relationships in the MARC 21 Authority and Bibliographic Formats
- Use of Subfields \$0 and \$1 to Capture Uniform Resource Identifiers (URIs) in the MARC 21 Formats
- Defining Field 758 (Related Work Identifier) in the MARC 21 Authority and Bibliographic Formats

For details please refer to

<https://www.loc.gov/aba/pcc/bibframe/TaskGroups/URI-TaskGroup.html>



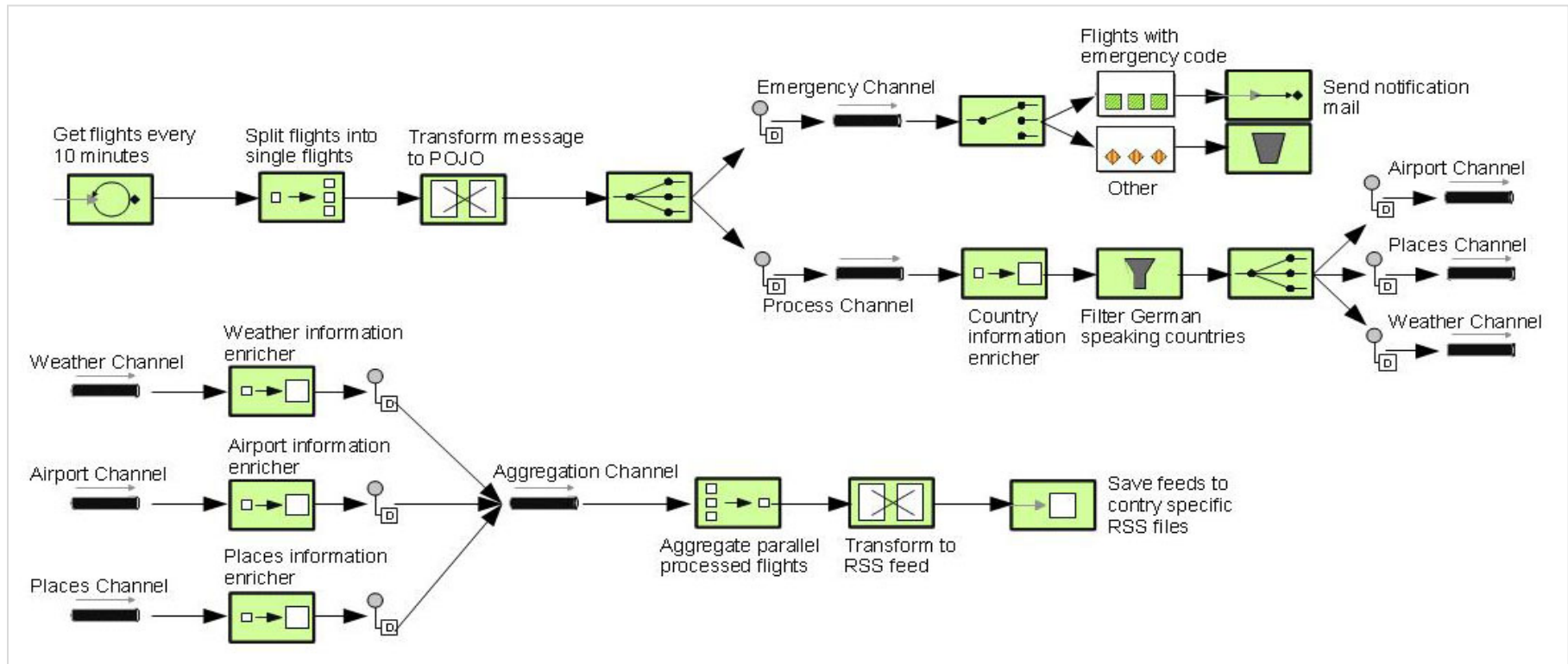
## **3.4 BIBFRAME and Linked Data practices: the SHARE-VDE project**

### **Conversion of authority and bibliographical data in BIBFRAME**



# Lodify - the RDF (BIBFRAME) conversion tool

The process is split into atomic pieces (processors), each of these responsible for a small part of the overall task. Each processor can act as a splitter or aggregator and can achieve content manipulation of the incoming message.

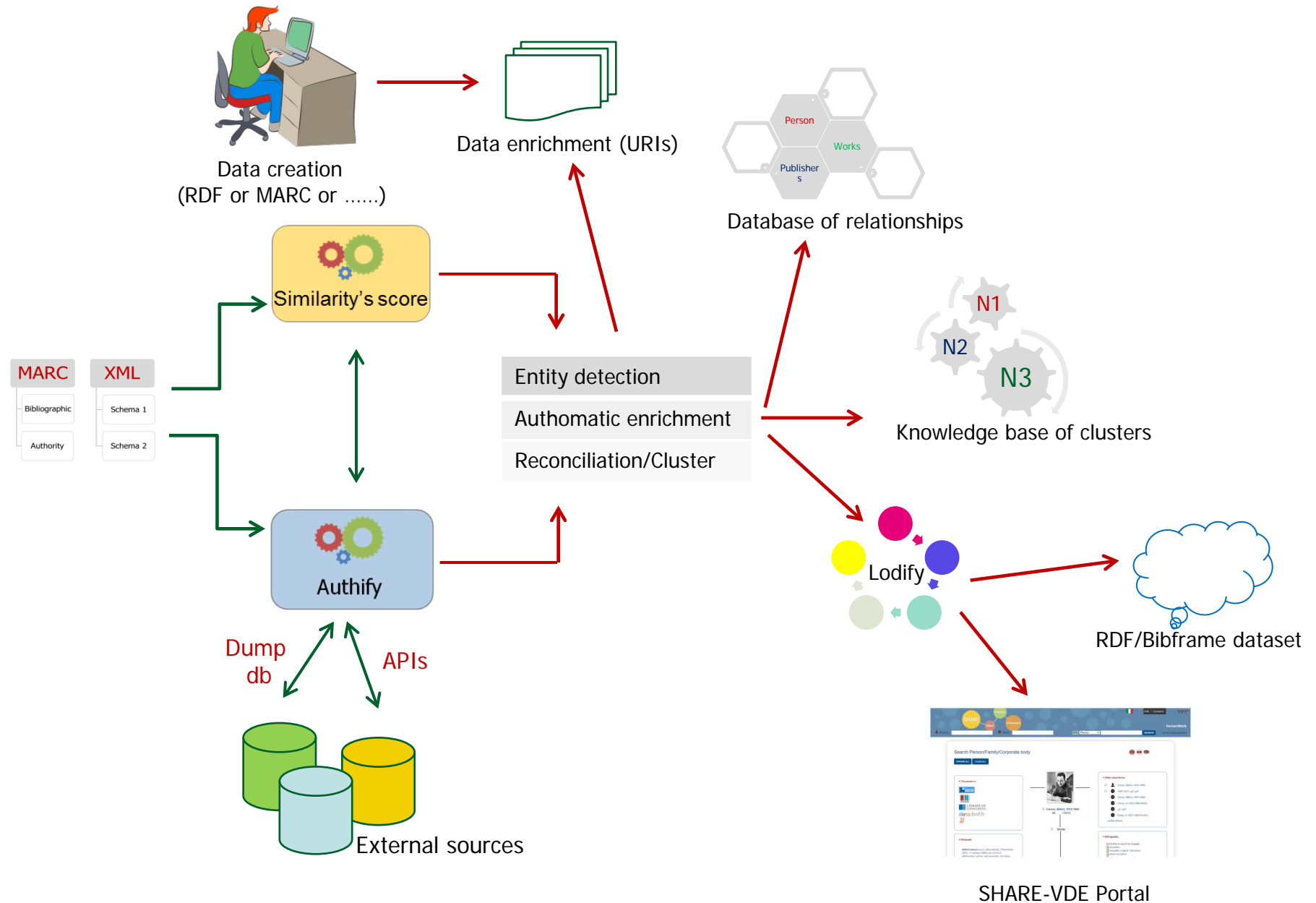


## **3.5 BIBFRAME and Linked Data practices: the SHARE-VDE project**

**Publication of a BIBFRAME three layered  
platform prototype**



# The overall SHARE-VDE processes



# Albert Camus on the SHARE-VDE platform

SHARE Virtual Environment Discovery Person/Work English Info Contattaci A A A Vai a Pubblicazioni

Persona  Opera  Persona  SCORRI

## Persona/Ente/Famiglia



ESPANDI

RIDUCI

### ▼ Questo autore in



data.bnf.fr

VI  
AF



Camus, Albert, 1913-1960  
ID: 133656

Opere

### ▼ Altre forme del nome

- Camus, Albert, 1913-1960
- 1960-1913 كامو، ألبير
- Camus, Albert, 1913-1960
- Camus, A. 1913-1960 Albert
- كامو، ألبير
- Камю, А. 1913-1960 Альбер

...(altre forme)

### ▼ Bibliografia

(Clicca per cercare su google)

- Actuelles
  - Actuelles. English. Selections
  - Adam ha-rishon
  - Albert Camus vous parle
- ...(altri titoli)

[http://share-vde.org/sharevde/searchNames?n\\_cluster\\_id=133656](http://share-vde.org/sharevde/searchNames?n_cluster_id=133656)

# 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
- 1678-1741, ויבאלי, אנטוניו
- Vivaldi, Antonio, 1678-1741
- Vivaldi, Antonio
- Vivaldi, Antonio, sac., 1678-1741
- Вивальди, А. 1678-1741 Антонио
- Вивальди, Антонио, 1678-1741
- Vivaldi, Antonio, 1680-1741
- 1741-1678, فيبالي، أنطونيو
- Antonio Vivaldi compositore e violinista italiano esponente di spicco del tardo barocco veneziano
- Vivaldi, Antonio, ca.1678-1741
- 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

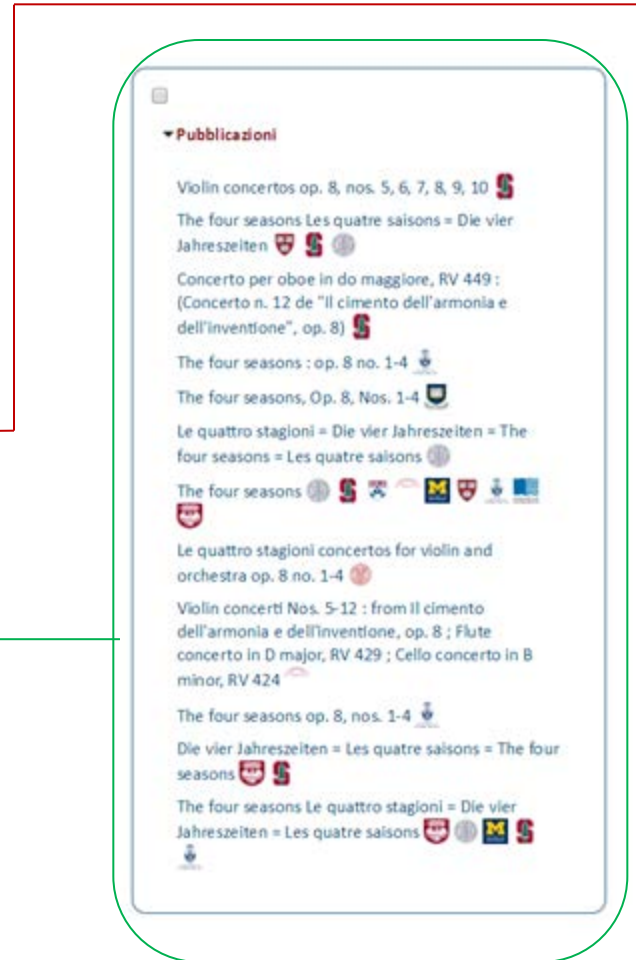


# An example of Work/Instances reconciliation

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.



▼ Pubblicazioni

Violin concertos op. 8, nos. 5, 6, 7, 8, 9, 10

The four seasons Les quatre saisons = Die vier Jahreszeiten

Concerto per oboe in do maggiore, RV 449 : (Concerto n. 12 de "Il cimento dell'armonia e dell'invenzione", op. 8)

The four seasons : op. 8 no. 1-4

The four seasons, Op. 8, Nos. 1-4

Le quattro stagioni = Die vier Jahreszeiten = The four seasons = Les quatre saisons

The four seasons

Le quattro stagioni concertos for violin and orchestra op. 8 no. 1-4

Violin concerti Nos. 5-12 : from Il cimento dell'armonia e dell'invenzione, op. 8 ; Flute concerto in D major, RV 429 ; Cello concerto in B minor, RV 424

The four seasons op. 8, nos. 1-4

Die vier Jahreszeiten = Les quatre saisons = The four seasons

The four seasons Le quattro stagioni = Die vier Jahreszeiten = Les quatre saisons



Cimento dell'armonia e dell'invenzione  
ID: 11287

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



# Example of same Instances present in different libraries

The header features a navigation menu with 'Info' and 'Contacts' links, a language selector set to 'Italiano', and a 'Go to Person/Work' link. A search bar is present with a magnifying glass icon. Below the search bar, there are options for 'Advanced search', a location dropdown set to 'Everywhere', a 'show 10' results per page dropdown, and an 'ordered by Score' dropdown. On the left, there are three circular icons labeled 'SHARE', 'Virtual', and 'Environment', with 'Discovery' written above them.

**Info**

Select one or more items in order to activate some functionalities like: print, send by email, download in available formats.

**Library**

- University of California Berkeley (2)
- Columbia University (1)
- Cornell University (1)
- Harvard University (1)
- Library of Congress (1)
- Stanford University (1)
- Texas A&M University (1)
- University of Chicago (1)

More...

Tutto x +

**MARC** **Tabular view** **Short view**

**The storm and other things / Eugenio Montale ; translated, with preface and commentary, by William Arrowsmith (9 documenti)**

Montale, Eugenio, 1896-1981

Publisher New York : Norton, c1985

Extent 219 p. ; 25 cm.

Format Language material

<input type="checkbox"/>	ID	CORNELL1138554 (Click for details)
-----		
<input type="checkbox"/>	ID	UCBERKELEYUCb125195606 (Click for details)
-----		
<input type="checkbox"/>	ID	UMICH000663201 (Click for details)
-----		
<input type="checkbox"/>	ID	COLUMBIA444093 (Click for details)
-----		
<input type="checkbox"/>	ID	LOC2859531 (Click for details)
-----		
<input type="checkbox"/>	ID	YALE440351 (Click for details)
-----		
<input type="checkbox"/>	ID	HARVARD000456337-9 (Click for details)

**Format**

Language material (13)

**Bibliographic level**

Monographs (13)

**Author (Person)**

Montale, Eugenio, 1896-1981 (13)

Arrowsmith, William, 1924-1992 (9)

Arrowsmith, William, 1924- (4)

**Author (Corporate)**

**Author (Conference)**

**Works**

Bufera e altro (13)

**Publication**

Norton (13)

**Language**

English (13)

Italian (3)

**Date**

5 results found

[http://www.sharevde.org/sharevde/search?t\\_cluster\\_id=7961;Bufera%20e%20altro&v=ll&dls=true](http://www.sharevde.org/sharevde/search?t_cluster_id=7961;Bufera%20e%20altro&v=ll&dls=true)

## 3.6 Further connected topics



## Further connected topics

- User interface platform enhancements
  - For both Person/Work and Instance layers : search-box for relationship presentation (e.g. author to subject, author to publisher);
  - First edition identification for the chronological positioning of Entities;
  - “Work of” or “Work concerning” a Person.
- Additions of ontologies other than BIBFRAME.
- Subject URI enrichment; content, media, carrier enrichment.
- Analysis of relationships among subject terms in different languages.
- Provenance declaration.
- Update management and URI Registry.



## 4. Conclusions



# Conclusions

- All research, the Humanities in particular, needs more than ever to be visible, accessible and innovative.
- Perceptions of research fields not foremost in the mind of the public at large can be changed and their importance for society acknowledged, reducing the risk of the Humanities field being marginalised.
- Key to achieving this is collaboration among all stakeholders of the information chain.
- Linked Data practices give vital support to the stewardship of research and introduce an invaluable opportunity that can contribute to taking forward cultural heritage for future generations.



## 5. Resources and links



# Resources and links

Bibliographic Framework Initiative homepage: <http://www.loc.gov/bibframe>

RDA Steering Committee (RSC): <http://www.rda-rsc.org>

PCC Task Group on URIs in MARC: <http://www.loc.gov/aba/pcc/bibframe/TaskGroups/URI-TaskGroup.html>

Linked Data for Production (LD4P):

<http://wiki.duraspace.org/pages/viewpage.action?pageId=74515029>

IMLS Shareable Authorities Forum:

<http://confluence.cornell.edu/display/sharedauth/IMLS+Shareable+Authorities+Forum+Home>

Linked Open Data in Libraries Archives and Museums (LODLAM): <http://www.lodlam.net>

European RDA Interest Group - 2017 meeting conference material: <http://www.casalini.it/eurig2017>

SHARE-Virtual Discovery Environment and the Casalini experience and roadmap for supplying BIBFRAME data. [Presentation delivered at the Program for Cooperative Cataloguing Operations Committee Meeting, Library of Congress, May 5th 2017]

[http://www.loc.gov/aba/pcc/documents/OpCo-2017/PCC-OpCo-2017\\_SHARE-VDE\\_Casalini-Possemato.pdf](http://www.loc.gov/aba/pcc/documents/OpCo-2017/PCC-OpCo-2017_SHARE-VDE_Casalini-Possemato.pdf)



# Examples on the SHARE-VDE platform

(with enhanced MARC records and BIBFRAME datasets)

[www.share-vde.org](http://www.share-vde.org)

Emily Bronte:

[http://share-vde.org/sharevde/searchNames?n\\_cluster\\_id=318705](http://share-vde.org/sharevde/searchNames?n_cluster_id=318705)

and this Work Wuthering Heights:

<http://share-vde.org/sharevde/resource?uri=LOC18843460&v=I&dcnr=1>

Frankenstein:

<http://share-vde.org/sharevde/resource?uri=LOC18789412&v=I&dcnr=8>

Eugenio Montale:

[http://share-vde.org/sharevde/searchNames?n\\_cluster\\_id=166369](http://share-vde.org/sharevde/searchNames?n_cluster_id=166369)

and his Works:

<http://share-vde.org/sharevde/resource?uri=UCBERKELEYUCb232697760&dir=1&v=I>

Reconciliation of the same instances present in different catalogues of his Works:

[http://www.share-vde.org/sharevde/search?t\\_cluster\\_id=8665;Stones%20of%20Venice&v=II&dls=true](http://www.share-vde.org/sharevde/search?t_cluster_id=8665;Stones%20of%20Venice&v=II&dls=true)





IFLA-Satellite-Meeting 2017  
**Digital Humanities. Connecting Libraries and Research**  
Berlin, August 16th 2017

**Thank you.**

**Any feedback will be precious.**

Michele Casalini  
[michele@casalini.it](mailto:michele@casalini.it)  
[www.casalini.it](http://www.casalini.it)

