



EURIG Annual meeting
Fiesole 2017

local host



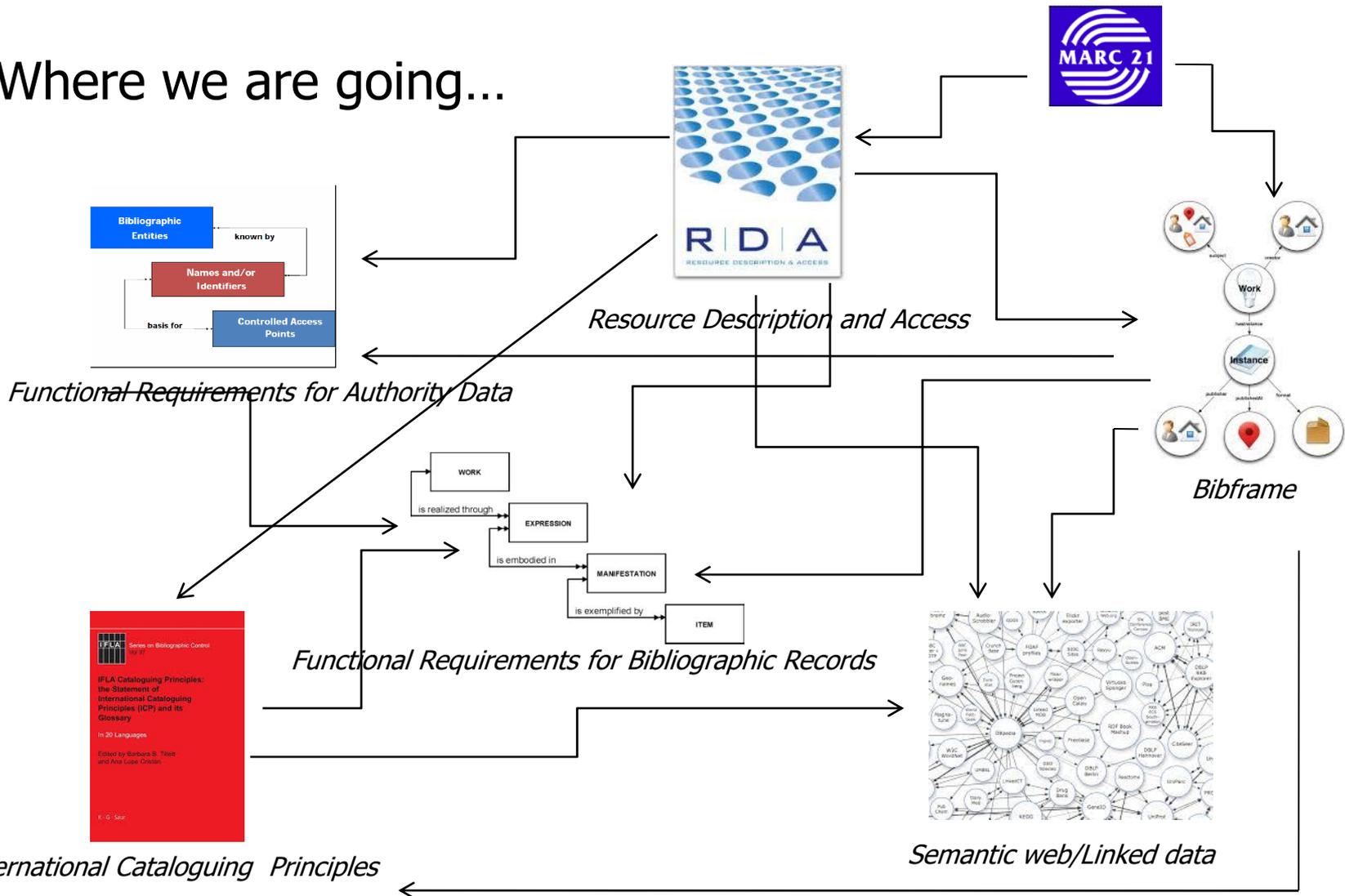
How RDA is essential in the reconciliation and conversion processes for quality linked data

Tiziana Possemato
Casalini Libri - @Cult

The theoretical context of SHARE-LOD projects

The theoretical context of the project

Where we are going...



The theoretical context of our SHARE projects

New standards, models and technologies as ways to approach entity **identification** and the **relationships** between entities, recognized as the key element in the construction of new **entity detection** and **entity identification** processes:

- **RDA – *Resource Description and Access***, the international guidelines to manage resources
- **Linked Open Data** philosophy and technology
- **BIBFRAME**: one of more interesting models to convert and publish data. This model is considered ‘the core’ ontology, completed with the ontologies for specific domains, that libraries will suggest

RDA Toolkit: Identify and Relationship

The structure of the RDA Toolkit clearly expresses the importance given by the standard to concepts of **identification** and **relationship**:

- Section 1: Recording Attributes of Manifestations & Items
- Section 2: Recording Attributes of Works & Expressions
- Section 3: Recording Attributes of Agents
- Section 4: Recording Attributes of Concepts, Objects, Events & Places

IDENTIFY

RDA Toolkit: Identify and Relationship

- Section 5: Recording Primary Relationships between Works, Expressions, Manifestations & Items
- Section 6: Recording Relationships to Agents
- Section 7: Recording Relationships with Concepts, Objects, Events & Places
- Section 8: Recording Relationships between Works, Expressions, Manifestations & Items
- Section 9: Recording Relationships between Agents
- Section 10: Recording Relationships between Concepts, Objects, Events & Places

RELATIONSHIPS

The 4 rules for Linked Data creation

by *Sir Tim Berners-Lee*

1. Use URIs as names for things: **give unique names to things**;
2. Use HTTP URIs so that people can look up those names: **the names assigned to things must also be machine readable**;
3. When someone looks up a URI, provide useful information, using the standards (RDF, SPARQL): **things must be self-explanatory (dereferencing)**;
4. Include links to other URIs so that they can discover more things: **create links with other objects** (any object can become the subject of a new statement).

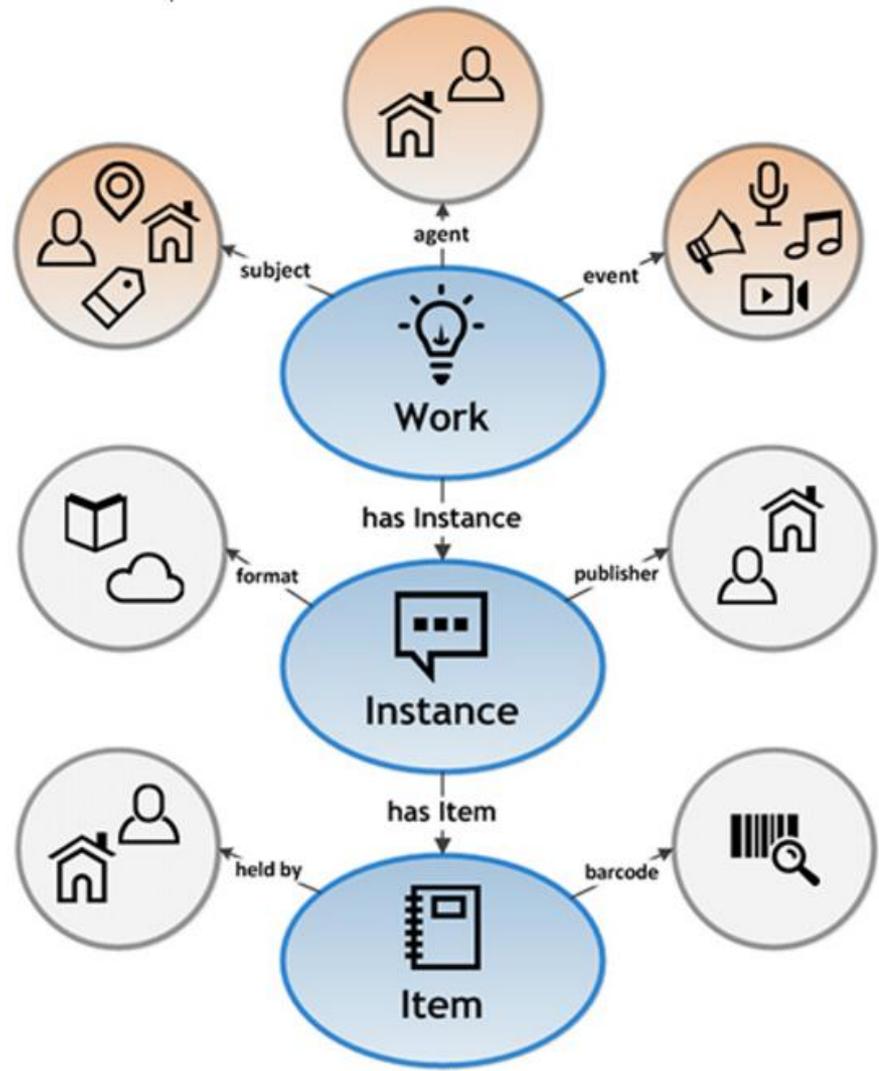
BIBFRAME – Bibliographic Framework Initiative

The **Bibliographic Framework as a Web of Data: Linked Data Model and Supporting Services** document published by the Library of Congress on November 21, 2012, sets out a new data model designed as an evolution, in linked open data, of the Marc 21 format.

The reflections on the new cataloguing rules focus on some specific points, including:

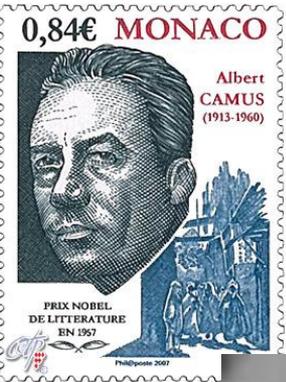
- a greater level of identification and analysis of the data;
- greater attention to controlled vocabularies;
- more widespread use of terms instead of codes;
- emphasis on relationships;
- greater flexibility in controlled items.

BIBFRAME – Data model v. 2.0



Who's Who?

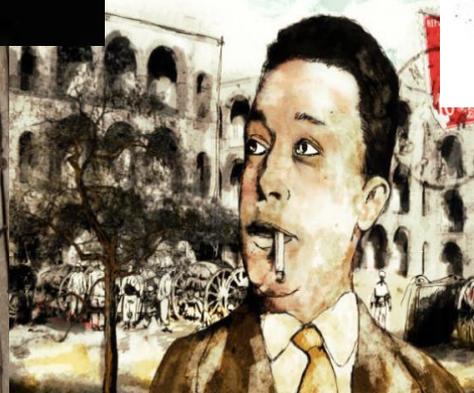
The question at hand:
how to identify an entity?



I don't care, and you find that attractive. But I also don't care about the



Happy Valentine's Day ♡
Love, _____ and
Albert Camus
benkling.tumblr



Albert Camus



The importance of identification in the catalographic tradition (and not only!)



Entity identification: it has traditionally been considered a highly important aspect of cataloguing.

But, the use of attributes to identify an entity has not been widely used

* *Both pictures are taken at the City Lights Bookstore, in San Francisco*

New cooperative scenarios

New context: new ways of cooperating between institutions and corporations, further removed from a complex *reductio ad unum* approach and physical merging.

The new generation of Authority control and discovery tools: cross-institutional processes of cooperation, integration and virtualization.

New data enrichment opportunities absolutely not possible in the past.

Focus on identifying entities and discovering their relationships with other entities.

Data reconciliation, enrichment and conversion

With the on-line presence of different catalogues and authority files available in various formats and, where possible, in open way, also the concept of authority control and of union catalogue has evolved into the **grouping of an entity's identifying attributes** from different sources.

The process is best known as ***reconciliation*** and consists in **creating a cluster of data that all refer to the same entity**.

The new revolution: from record to entity

```

Full Record | MARC Tags | Where to Request
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 As you like it / |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 As you like it.
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

Padri e figlie

Pères et filles

EXPAND ALL CLOSE ALL

This person in



Guicciardini, Francesco, 1483-1540
ID: 8569

Works

Other name forms

- Guicciardini Francesco
- Guicciardini, Francesco <1483-1540>
- Guicciardini, Francesco
- Guichardin, François, 1483-1540
- Guicciardini, Francesco (1482-1540)
- Francesco Guicciardini scrittore, storico e politico italiano
- Guicciardini, Francesco, 1483-1540
- Guicciardini, Francesco
- GUICCIARDINI, Francesco
- GUICCARDINI, Francesco <1483-1540>

Wikipedia

Francesco Guicciardini (Italian: [franˈtʃesko ɡwiʧʧiˈdʒarˈdiːni]; 6 March 1483 – 22 May 1540) was an Italian historian and statesman. A friend and critic of Niccolò Machiavelli, he is considered one of the major political writers of the Italian Renaissance. In his masterpiece, *The History of Italy*, Guicciardini paved the way for a new style in historiography with his use of government sources to support arguments and the realistic analysis of the people and events of his time.

Wikipedia

Biography and activity

Storico, uomo politico

SELECT ALL UNCHECKED ALL

EXPAND ALL CLOSE ALL

A-C

- | | | |
|---|---|--|
| <input type="checkbox"/> Antimachiavelli | <input type="checkbox"/> Autodifesa di un politico | <input type="checkbox"/> Avertissements politiques |
| <input type="checkbox"/> Carteggi | <input type="checkbox"/> Carteggi di Francesco Guicciardini | <input type="checkbox"/> Cassa Nazionale di previdenza per la invalidita e la vecchiaia degli operai |
| <input type="checkbox"/> Cento giorni alla consulta | <input type="checkbox"/> Compendio della Cronica di Froissart | <input type="checkbox"/> Considerations a propos des discours de Machiavel sur la premiere decade de Tite Live ... |

The identification of entity goes through several roads...

...or it doesn't go...





Guicciardini, Francesco, 1483-1540

ID: 8569

Works

Year of publication: 1901



Subject: Previdenza sociale



Guicciardini, Francesco, 1851-1915

Author: Guicciardini, Francesco

Title: La Cassa Nazionale di previdenza per la invalidita e la vecchiaia degli operai / Francesco Guicciardini

Publisher: Firenze : ..., 1901

Physical description: 32 p. ; 23 cm

Dewey: 360

Uncontrolled subject: Previdenza sociale

Preferred title for the work: Cassa Nazionale di previdenza per la invalidita e la vecchiaia degli operai

Format: Language material

Bibliographic level: Monograph

Language: Italian

Record Nr.: 000180165

You will find it: Univ. Federico II

Location code: 60 DONO COMES 5/22

Opac: Check copies here

Identify a Person

Identify a Work

000 00680nz 2200181 4500
001 000000000201
003 CaOOAMICUS
005 20170505115045.0
008 170505 n annnnabbn n2ann und u
040 a Casalini
b ita
e rda
046 f 1775-12-16
g 1817-07-18
2 dtf
100 1 a Austen, Jane
d 1775-1817
370 a Steventon (Hampshire, England)
b Winchester (England)
c England
c Great Britain
e Bath (England)
2 naf
372 a Fiction
2 lcsh
374 a Authors
a Novelists
2 lcsh
375 a female
377 a eng
670 a Wikipedia, via WWW, Feb. 8, 2012

000 00904nz 2200217 4500
001 000000000202
003 CaOOAMICUS
005 20170505120143.0
008 170505 n annnnabbn n2ann und u
040 a PS
b ita
046 k 1815
2 edtf
100 1 a Austen, Jane
d 1775-1817
t Emma
370 g England.
g Great Britain.
2 naf
380 a Novels.
a Domestic fiction.
a Humorous fiction.
2 lcgft
386 n nat.
a English.
a Britons.
2 lcdgt
386 n gdr.
a Women.
2 lcdgt
500 1 w nax@
i Dramatized as (work):
a Asher, Sandy.
t Emma
500 1 w nax@
i Dramatized as (work):
a Bloom, Michael
d 1950-
t Emma
530 0 w nax@
i Adapted as motion picture (work):
a Emma (Motion picture.; 1996)
670 a Encyc. Britannica online, January 9, 2015
b (Emma, novel by Jane Austen, published in three volumes in 1815)
670 a Austen, J. Emma; an authoritative text, backgrounds, reviews, and criticism, 1972.

How reconciliation is obtained

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 profoundly between the automated and manual processes:

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

Starting by the end

An example of reconciliation: Albert Camus in SHARE-VDE project

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

Persona Opera Persona SCORRI

Persona/Ente/Famiglia 📄 🖨️ ✉️

ESPANDI RIDUCI

▼ Questo autore in

- 
- 
- 
- 
- 

▼ Wikipedia

Albert Camus (IPA: [al'beɪ ka'my]) (Mondovi, 7 novembre 1913 – Villeblevin, 4 gennaio 1960) è stato uno scrittore, filosofo, saggista, drammaturgo e attivista francese.

Con la sua multiforme opera è stato in grado di descrivere e comprendere la tragicità di una delle



Camus, Albert, 1913-1960
ID: 133656

Opere

▼ Altre forme del nome

-  Camus, Albert, 1913-1960
-  1960-1913 كامو، ألبيير
-  Camus, Albert, 1913-1960
-  Camus, A. 1913-1960 Albert
-  كامو، ألبيير
-  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

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

Vivaldi, Antonio, 1678-1741
ID: 37154

▼ 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, فيفالي، 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)

▼ Questo autore in

- WIKIDATA
- LIBRARY OF CONGRESS
- data.bnf.fr
- VIAF

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'amore e dell'inventione*

One work title

Brings together different publications 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'inventione", 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'inventione, 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'inventione
ID: 11287

http://share-vde.org/sharevde/searchTitles?t_cluster_id=11287

Data entification, reconciliation, enrichment and publication

Bring together and make available data from different sources in a way that could be defined as *democratic* to better identify the entity in question.

Even wider reconciliation and enrichment processes form the basis of a number of projects that convert and publish bibliographic catalogues as linked open data, such as:

- **Share Catalogue:** <http://catalogo.share-cat.unina.it/sharecat/clusters?l=en> (an @Cult project for Italian Universities Libraries)
- **Share VDE – Share Virtual Discovery Environment:** www.share-vde.org (in partnership between Casalini Libri and @Cult)

SHARE-VDE brief project overview

Threefold goals:

- Conversion, supply and management of authority and bibliographical data in BIBFRAME taking into account the complexity of the long and heterogeneous transition time;
- Development of detection services for entity identification including relator terms, and creation of a common knowledge base of clusters of reconciliated results for names and works;
- Publication of a FRBR/BIBFRAME three layered platform with build-in instances techniques.

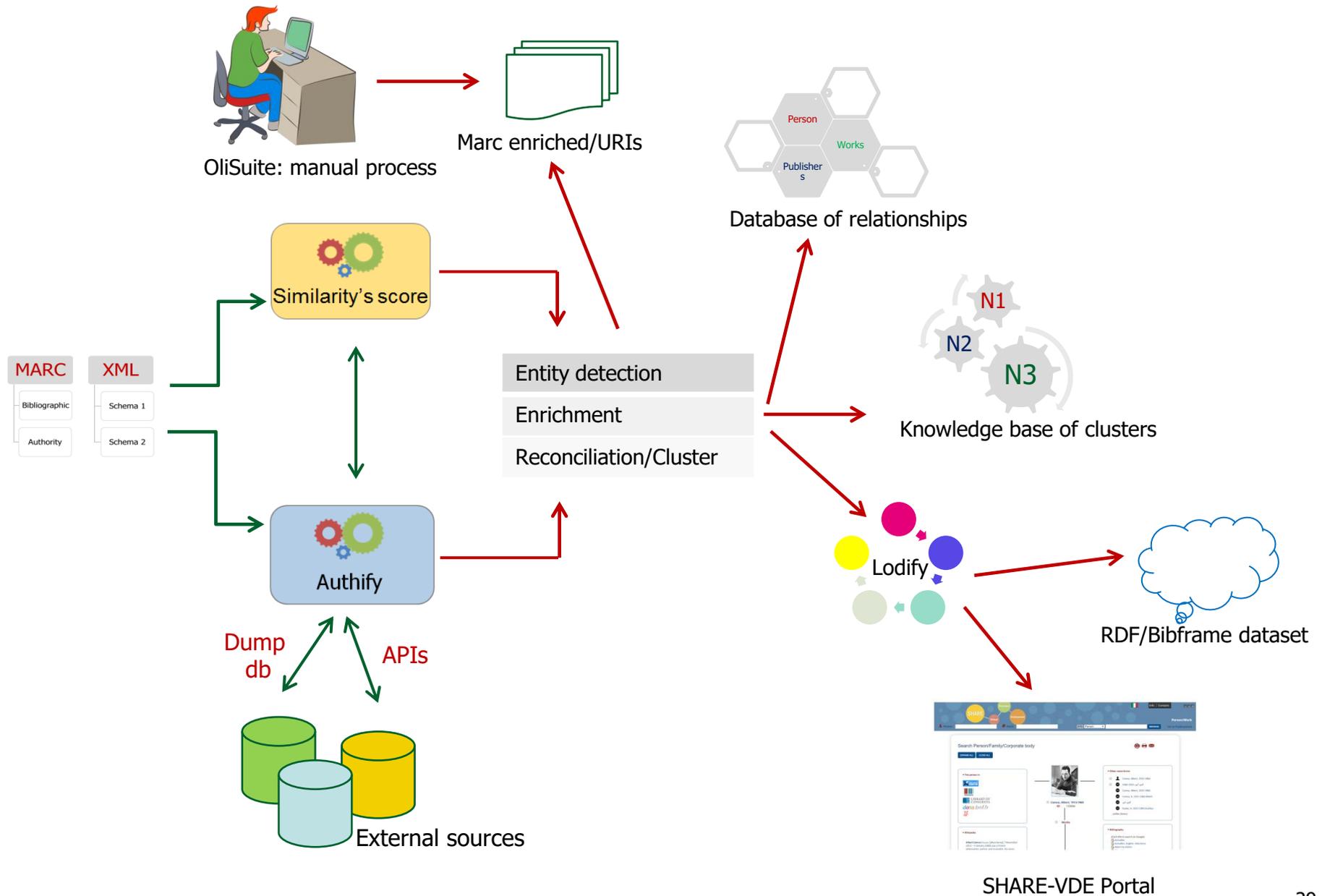
Participants libraries

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

Participants libraries

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

The SHARE-VDE processes overview



Conclusions: the sharing and reuse of information resources

All energy and effort made to facilitate the **sharing and reuse of data, assets, and tools** produced by libraries, archives, museums and other institutions, and to guarantee their availability to a wider public, enriching the World Wide Web with information that would otherwise remain mostly hidden, promote a **culture of open access to knowledge**, with advantages for each link in the information chain.

Libraries, archives and museums all benefit from the possibility of more well-structured and sharable data which provide users with a **vast wealth of information**, and **create new cooperative scenarios**.

Some examples on SHARE-VDE platform

Emily Bronte:

http://share-vde.org/sharevde/searchNames?n_cluster_id=318705

And the Work Cime tempestose:

<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

and his Works:

http://share-vde.org/sharevde/searchTitles?t_cluster_id=7961&l=en

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

[Reconciliation of the same instances present in different catalogues \(Attention: is in the test db\):](#)

http://dev-vde.atcult.it/sharevde/search?t_cluster_id=7961;Bufera%20e%20altro&v=I&dls=true&l=en



EURIG Annual meeting Fiesole 2017

local host



Thanks

Tiziana Possemato
Casalini Libri - @Cult