The future of bibliographic services in light of new concepts of authority control

12th February 2021
Michele Casalini, Casalini Libri
New era, new needs

- Technological innovations + cataloguing changes: machines take on proactive roles in the bibliographic world.
- The cultural heritage sector is called to react to the growing need for supervisory authorities (higher risk of unreliable and perishable data).
- Authority control guarantees: homogeneity, stability and formal quality of access entries.
- The world-wide system of bibliographic control must respond to this need, more than ever, adapting its tools to the demands of the globalized world.
Where we are and what we need

Record → Entity: New authority control

- More flexible notion of entity identification and relationships between entities.
- Transition from FRBR to IFLA LRM.

Interoperability of data models

- Global cooperation (libraries, archives, museums, publishers, providers).
- Data structure: RDF quads and URIs.
The Provenance and its supporting role

The role of the Provenance to manage a shared environment (in building or in retrieving the authoritative data):

- in local authority: it’s important for data retrieval, not for clusters building;

- in external authoritative sources: it’s important for clusters building, not considered for data retrieval.
Putting new models into practice

Share-VDE is a library-driven initiative initiated in 2016 to establish an effective working environment for the use of linked data by libraries within a global context.

Library data are enriched with additional information and relationships, and bibliographic and authority data are converted into linked data.

A virtual discovery platform with a structure based on the BIBFRAME data model is created to simplify the way in which that data is consumed.

The network of resources created is the basis for the Share-VDE Sapientia Cluster Knowledge Base, the common authoritative source of clusters accessible in RDF, open to the Share-VDE community and beyond.
... in a cooperative environment

Share-VDE is a collaborative endeavour based on the needs of libraries, developed by:

- the joint efforts of the Share-VDE Advisory Council and of the Working Groups;
- Casalini Libri, provider of bibliographic and authority data as member of the Program for Cooperative Cataloguing;
- @Cult, provider of ILS, Discovery tools and Semantic web solutions for the cultural heritage sector;
- influenced by the vision of the Linked Data for Production initiative;
- with input and active participation from an international group of national and research libraries.
The Share family map around the world

Details can be found on wiki.svde.org
Share-VDE overall goals

- **Enrichment** of MARC records with identifiers from external sources (e.g. ISNI, VIAF)
- **Reconciliation and clusterization** of entities and publication of the Cluster Knowledge Base
- **Conversion from MARC to RDF** using the BIBFRAME vocabulary together with other ontologies
- **Batch/automated authority services**, data updating and data dissemination services
- **Manual entity management tool** (J.Cricket)
- **Data publication on an entity oriented user interface** (www.share-vde.org)
The Share-VDE AC takes an active role in determining future uses and vision for the Share-VDE initiative with the following aims: to develop future use cases for Share-VDE, and set development priorities as needed; to monitor and lead the work of the various Advisory Council Working Groups; to maintain communication with authorities and among the Share member institutions.

There are 4 sub-committees focusing on specific areas:

- Sapientia Entity Identification Working Group
- Authority-Identifier Management Services Working Group
- Cluster Knowledge Base Editor Working Group
- User experience/User Interface Working Group
Authority/Identifier Management Services WG

- defines guidelines and best practices for Authority/Identifier management in the linked data environment;

- defines scope and data-flow for the creation and implementation of automated services based on preliminary documentation;

- proposes additional use cases identified as essential for effective knowledge base management.
Some of the external sources

- RDA Registry
- OCLC Worldcat Identities
- ID.LOC.GOV
- Other sources
- FAST Linked Data
- WIKIDATA
- VIAF
- ISNI
- DATA.BNF.FR
- DNB Integrated Authority File
- Other sources
- OCLC
Integration - Wikidata

Wikidata is increasingly authoritative and is used in the library community as a source for entity identification (SVDE property on Wikidata Share-VDE author ID)

Query the source and enrich SVDE data with Wikidata entities information and vice versa → connection with Casalini participation in the PCC Wikidata pilot

Ad hoc SVDE working group is studying the use cases for interaction (e.g. starting points for the analysis are API:Main page + Wikibase/API, and other documentation)

Major challenge: alignment between Wikidata and SVDE entities
Integration - the technological perspective
Comparison IFLA-LRM/BIBFRAME/Share-VDE

IFLA-LRM

- Work
  - Expression
    - Manifestation
      - Item
      - Item

BIBFRAME

- Hub
  - Work
    - Instance
      - Item
      - Item

Share-VDE

- Opus
  - Work
    - Instance
      - Item
      - Item
The challenge of data models interoperability

Share-VDE Sapientia Entity Identification (SEI) Working Group decision of June 10th 2020

Resource that is a svde:Opus is also a bf:Work

Likewise resource that is a svde:Work is also a bf:Work
Towards the Share-VDE Sapientia CKB ecosystem
Automatic authority processes

For record environments:

→ MARC record validator
→ MARC corrections for errors and obsolete forms
→ MARC matching/enrichment with SVDE and external profiled sources

For RDF environments:

→ Access point enrichment (including Series and Subjects)
→ Matching, import and interaction with the Sapientia Cluster Knowledge Base

(Enabled through the LOD Platform)
Manual authority processes

→ Manual control for similar matches and for non matches
→ Entity Work and Agents (including Publisher) access point management
→ Integration with the ISNI registration processes
  (Enabled through the URI Registration Platform)

→ Cluster Knowledge Base entity management for Works and Agents
  (Enabled through the J. Cricket CKB Editor)
Integration of J.Cricket: the professional perspective

The editing tool J.Cricket will allow for editing of the SVDE Cluster Knowledge Base, Sapientia, enabling several actions on the clusters of entities, including creation, modification, merge of clusters of works, of agents etc.

J.Cricket is going to extend authority capabilities through integration with external data sources such as Wikidata and ISNI.
How J.Cricket interacts with Wikidata

Person
Isaiah Thomas

Cluster metadata

<table>
<thead>
<tr>
<th>Source</th>
<th>ID</th>
<th>Name</th>
<th>Born</th>
<th>Died</th>
<th>Confidence</th>
</tr>
</thead>
<tbody>
<tr>
<td>OCLC</td>
<td>21X1241</td>
<td>Isaiah Thomas</td>
<td>1989</td>
<td></td>
<td>Manualy added by Anna Doe 4 months ago (June 2nd, 2021 at 04:32PM) View log</td>
</tr>
<tr>
<td>VIAF</td>
<td>1568479</td>
<td>Isaiah Thomas</td>
<td>1989</td>
<td></td>
<td>Automatically added on June 2nd, 2021 at 04:32PM based on identical LC authority ID in VIAF and SVDE Validate</td>
</tr>
<tr>
<td>LC</td>
<td>8925663</td>
<td>Isaiah Thomas</td>
<td>1989</td>
<td></td>
<td>Automatically added on June 2nd, 2021 at 04:32PM based on identical LC authority ID in VIAF and SVDE Validate</td>
</tr>
<tr>
<td>Wikidata</td>
<td>Q05322233</td>
<td>Isaiah Thomas</td>
<td>1773</td>
<td>1819</td>
<td>Automatically added on June 2nd, 2021 at 04:32PM based on identical LC authority ID in VIAF and SVDE Validate</td>
</tr>
</tbody>
</table>

Connected external IDs

Actions

Add another external URI
### Connect to external IDs

<table>
<thead>
<tr>
<th>Name</th>
<th>OCLC</th>
<th>VIAF</th>
<th>Wikidata</th>
<th>LC</th>
<th>GND</th>
<th>BNF</th>
<th>Other</th>
</tr>
</thead>
</table>

#### Isaiah Thomas
- [Wikidata Item](Q356030)
- **Personal Name**
  - Birth name: Isaiah Jamar Thomas
  - Born: 1989

#### Isaiah Thomas, Sr.
- [Wikidata Item](Q1550343)
- **Personal Name**
  - Birth name: I. Thomas. J. T. J.
  - Born: 1749–1831

#### Isaiah Thomas, Jr.
- [Wikidata Item](Q85322233)
- **Personal Name**
  - Birth name: I. Thomas, Jun.
  - Born: 1773–1819

- Create new wikidata item for “John Doe”
J.Cricket’s purpose is to manage authority services for libraries, by providing automatic and manual data quality control procedures (APIs for automatic flows and an ad hoc web user interface for manual interaction are being designed). This means that J.Cricket purpose is twofold: the creation and handling of linked data entities within SVDE, and the direct interaction with library systems both in MARC and in RDF formats. The authority capabilities will be extended through integration with external data sources with the ability to mutually enrich both SVDE data and, for example, the Wikidata and ISNI platforms.
Linked data for the future

● Create connections across the LAM domain in order to preserve and promote existing knowledge while forwarding progress.

● Collaborative programs and initiatives through the players of the information chain are crucial to implementation for cultural heritage collections.

● Bibliographic agencies must strive to provide constantly updated and authoritative information as well as promoting an open and cooperative approach to authority control steered by the library community.
Thank you

Michele Casalini
CEO, Casalini Libri

michele@casalini.it