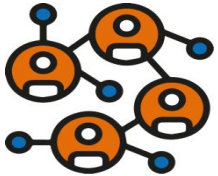
Share-VDE 2.0: a panel discussion among the Share-VDE working group chairs

Beth Picknally Camden, Filip Jakobsen, Ian Bigelow, Jim Hahn, Martin Knott, Erik Radio

Share-VDE Overview



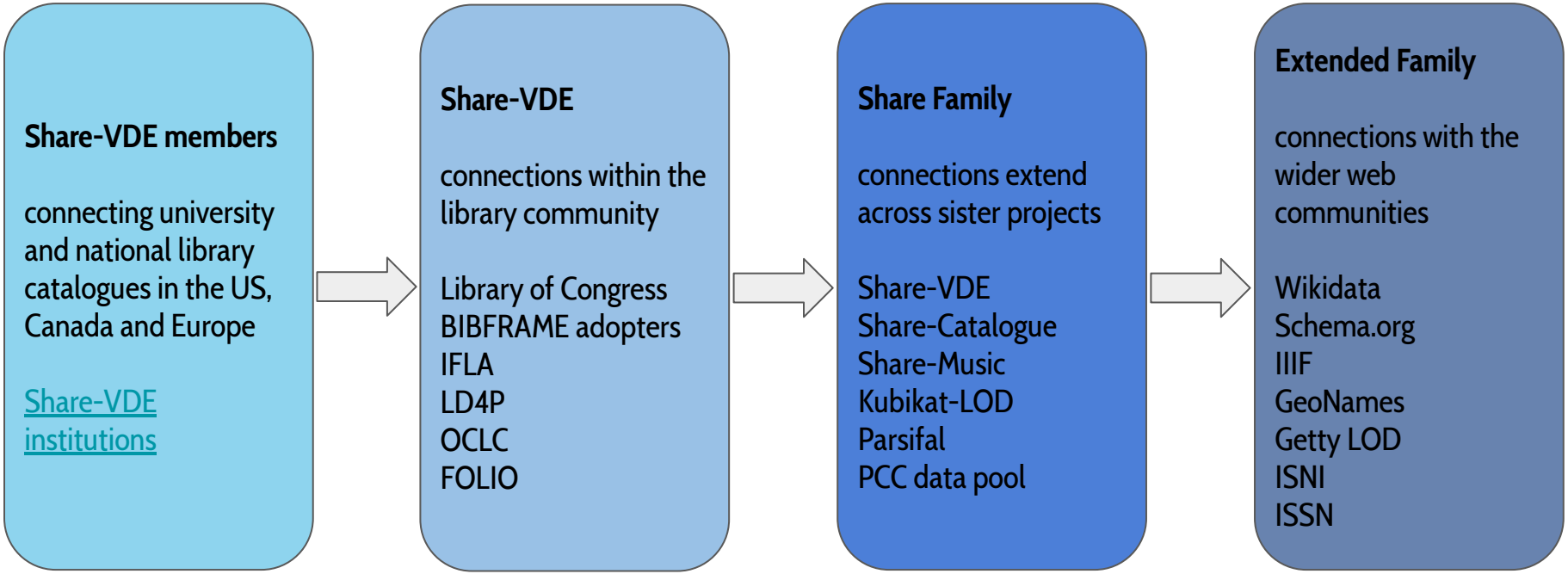
Share-VDE is a library-driven initiative to establish an effective working environment for the use of linked data by libraries within a global context.

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

A virtual discovery platform with the structure based on 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 entire Share-VDE community.

The Share Family



Share-VDE Advisory Council & Working Groups

The Share-VDE AC determines future uses and vision for the Share-VDE initiative; sets development priorities; monitors the work of working groups; and maintains communication among Share institutions and other groups.

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

Latest achievements

Design of the J.Cricket Cluster Knowledge Base Editor with the CKB Editor working group

Analysis of authority services with the Authority Identifier Management Services working group and initial release of the authority control features (delivered to Stanford)

Analysis for the integration with Wikidata and ISNI in AIMS and CKB working groups

Revision of the entity model with the Sapientia Entity Identification working group

Road map: 2021 Q3 & Q4

The SVDE 2.0 roadmap for the remainder of 2021 includes:

- Discovery:
  - Search Interface
  - Search APIs
- Cluster Knowledge Base 2.0 refactoring
- J.Cricket editor release
- Testing of authority services
- Ongoing updates of the PCC data pool
- Ongoing work on Sinopia/QA interaction

https://www.flickr.com/photos/revdrpepper/46815577174/

Sapientia Entity Identification Working Group (SEIWG)

Mandate

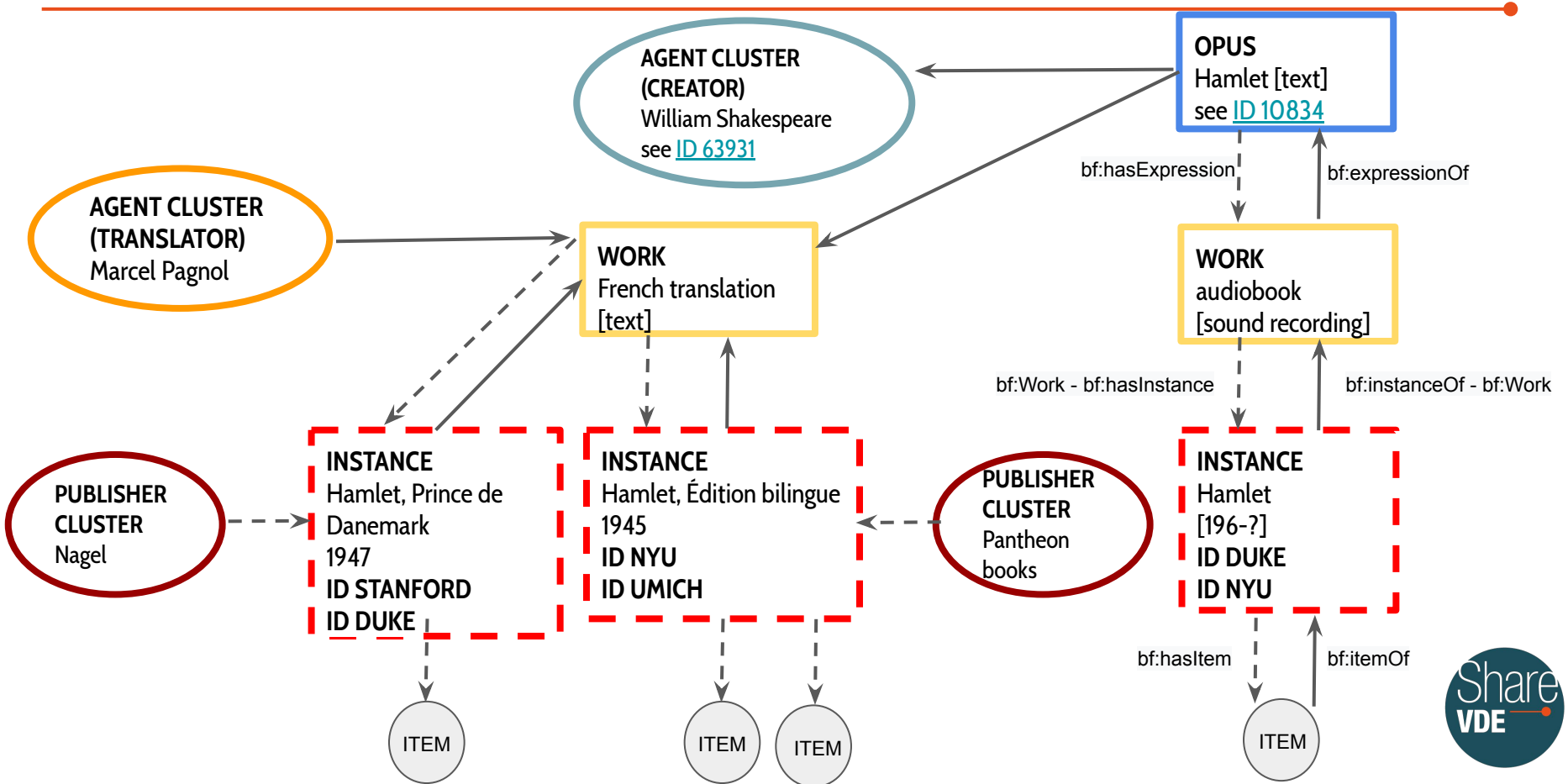
The SEIWG will:

1. Review use of entities, identifiers, and associated modelling in the Sapientia KB
2. Review and refine processes for Sapientia entity clustering in Share-VDE and the creation of associated open and stable URI for use in Share-VDE and the library community
3. Review MARC 21 to BIBFRAME and BIBFRAME to MARC conversion of elements related to entities in Sapientia
4. Engage with the library community to identify and/or develop best practices for use of Sapientia identifiers in BIBFRAME and MARC 21 data

In advance of Share-VDE 2.0 work has focused on:

- Further elaboration and refinement of Work and Opus entities
- Identification, conversion specifications and clustering for the Instance Entity
- In general, refinement and finalization of this iteration of the data model.
- Interoperability: Multiple communities, 1 model
    - BIBFRAME, LRM, and RDA working together

Share-VDE model (simplified version)



Opus

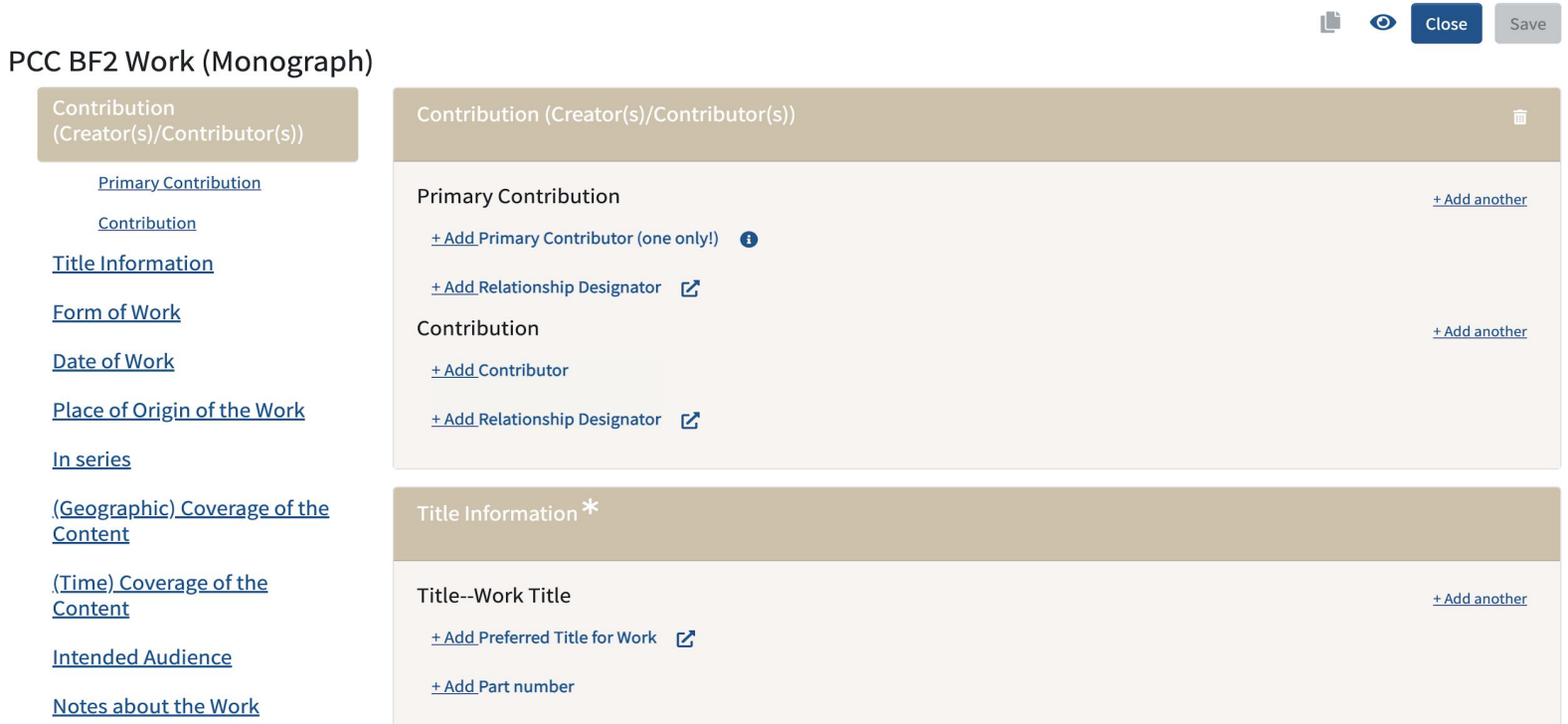
The Share-VDE Opus is a type of bf:Work.

This work type is similar to the Library of Congress Hub in the sense that it provides clustering features for bringing all the works under a canonical work type.

The SEI team have worked on defining properties that provide this clustering functionality for the related cluster of Works.

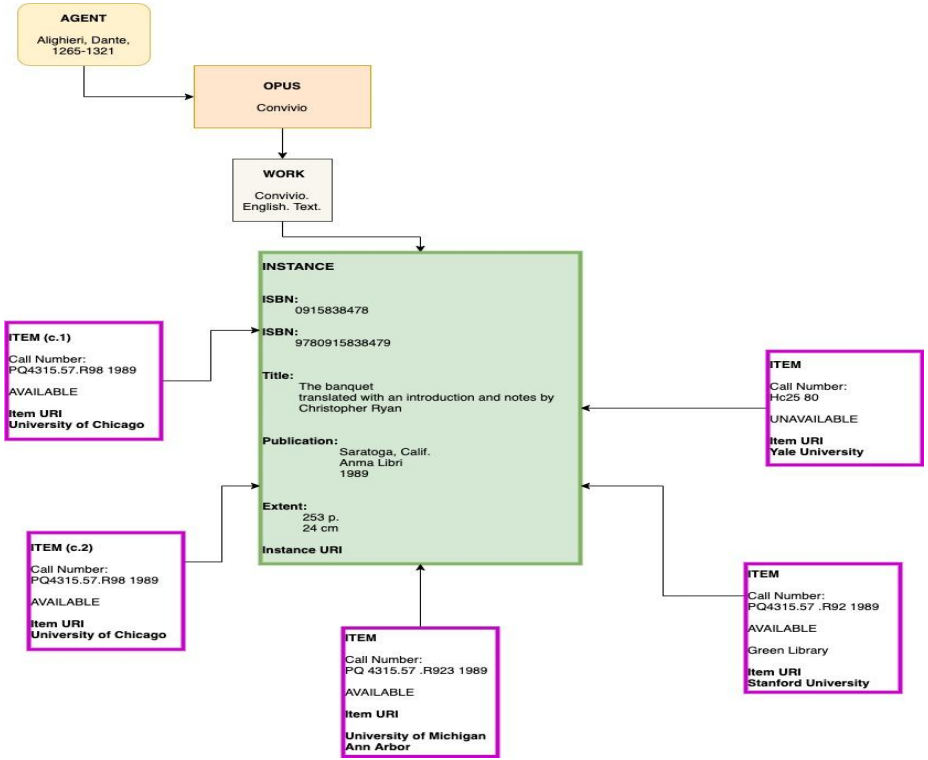
Work

Share-VDE Work maps to BIBFRAME Works and are compatible fully with BIBFRAME correspondances for Work properties.

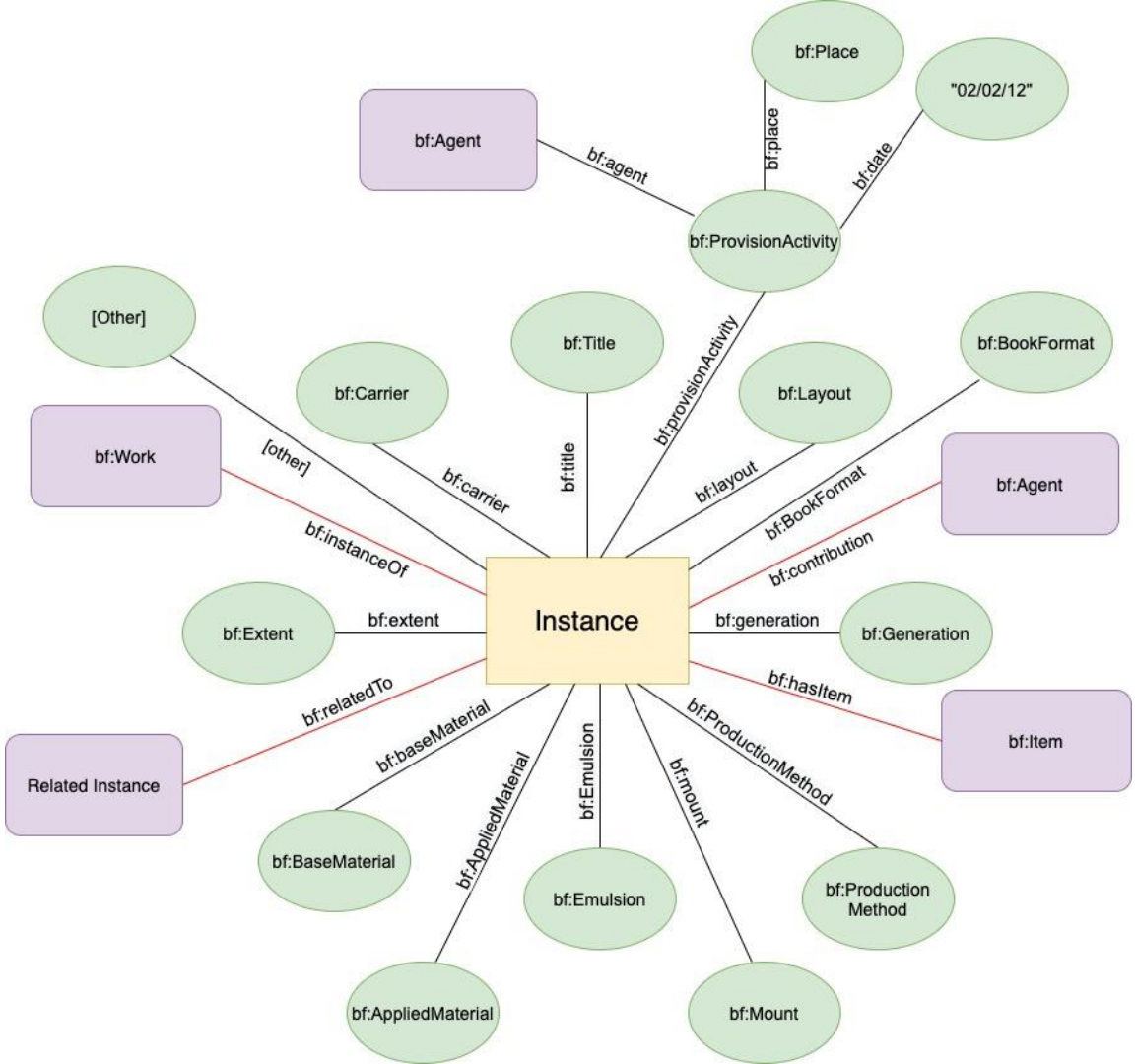


Instance

Development of Instance Clustering Process, Introducing the canonical Instance. The SEI team has worked on defining properties that provide this clustering functionality for the related Instances.



Instance modeling



Stardog Example

DATA EXPLORATION

Search Databases ...

ilibriup_prod
Triple Count: ~423M

ilibriup_test
Triple Count: ~48M

Share-VDE_CKB_prod
Triple Count: ~401M

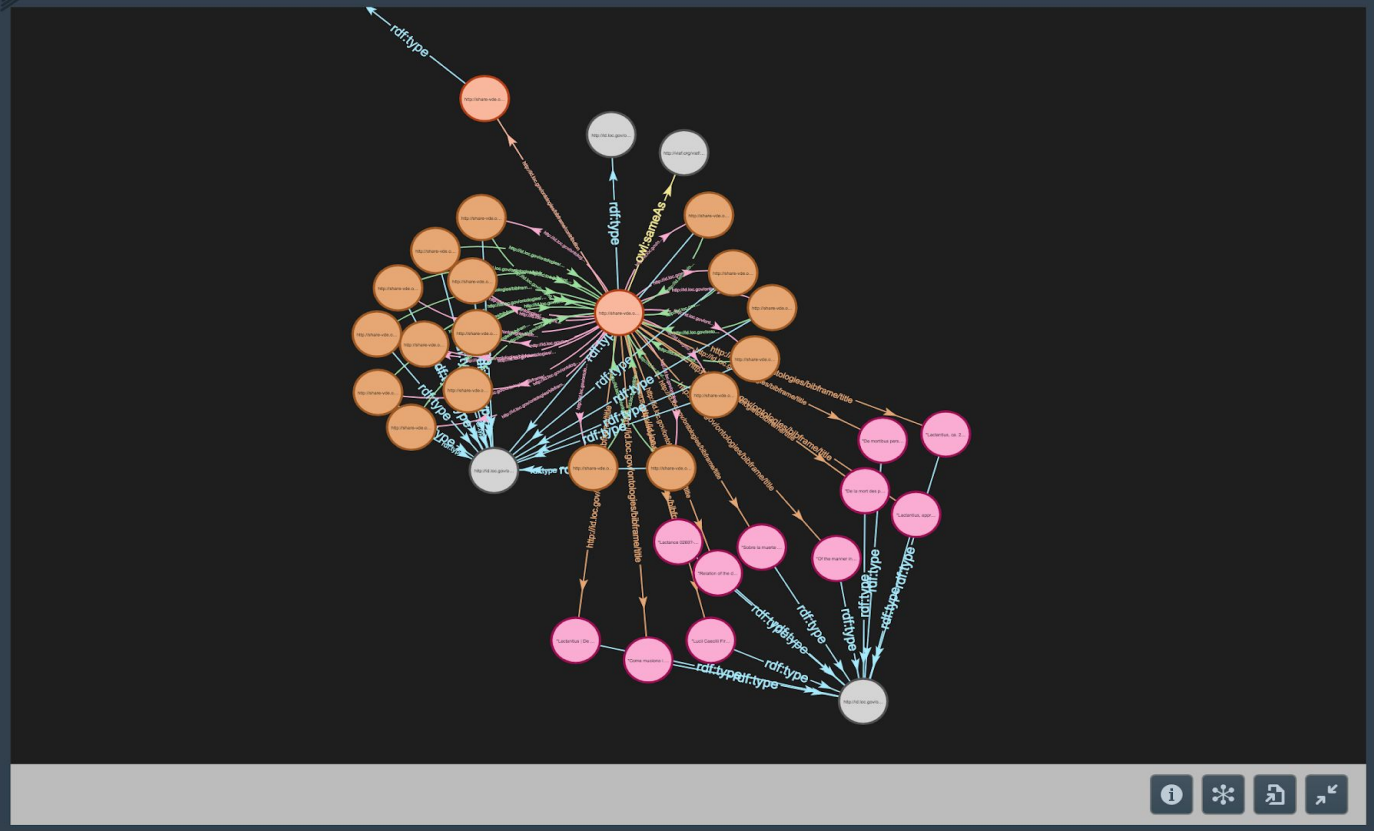
Share-VDE_prod
Triple Count: ~24B

Share-VDE_CKB_prod

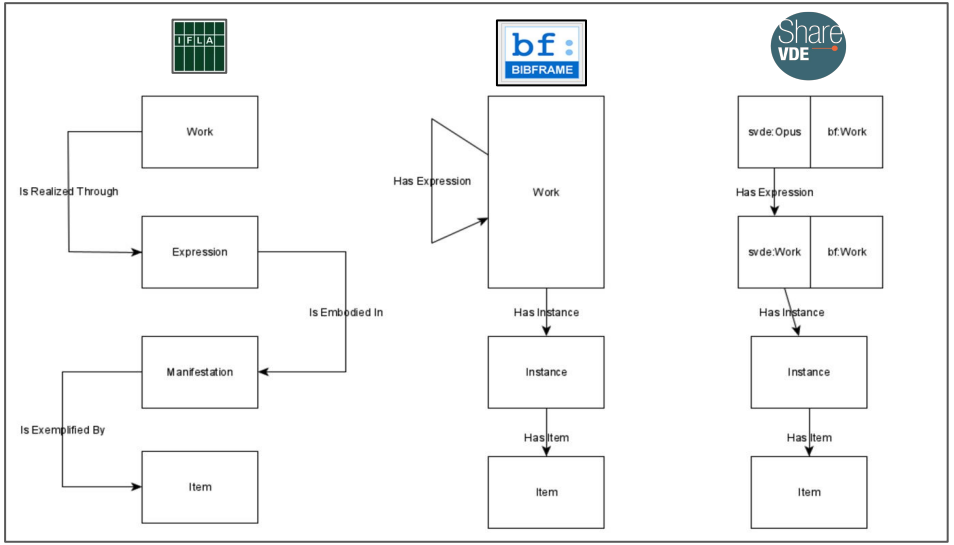
Search Share-VDE_CKB_prod

Classes — Properties

99504 (http://share-vde.org/sharevde/rdfBibframe/Work/99504)

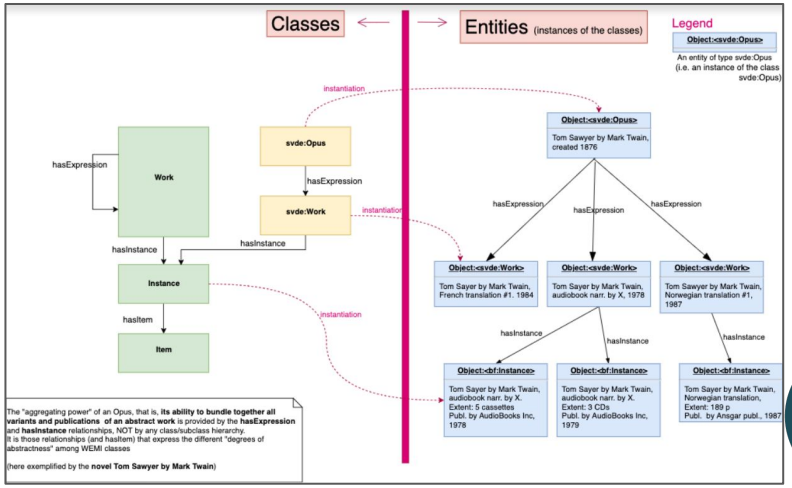

rdf:type
owl:sameAs
rdf:type
rdf:type
rdf:type

Facilitate interoperability between entity models…

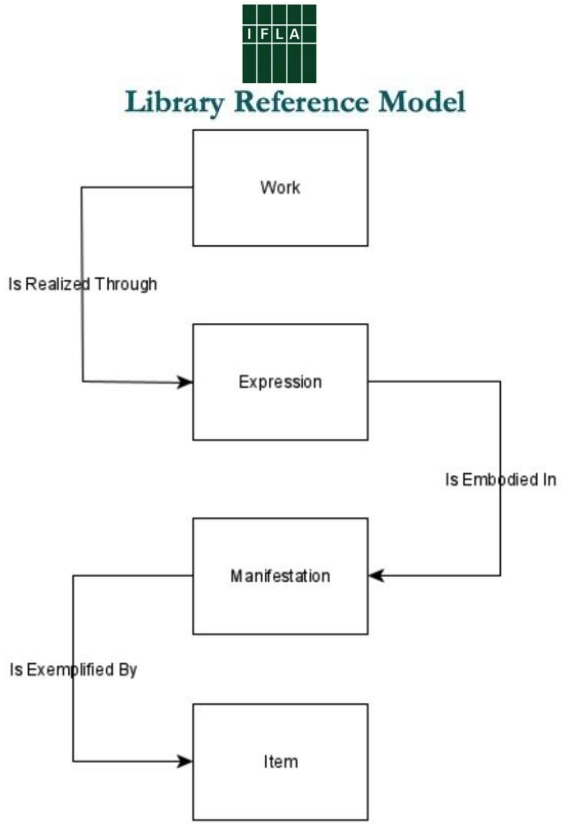


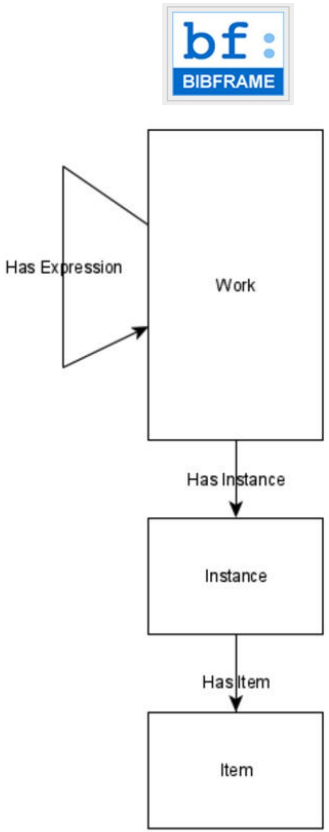
Participation in LD4P3: the challenge of data models interoperability

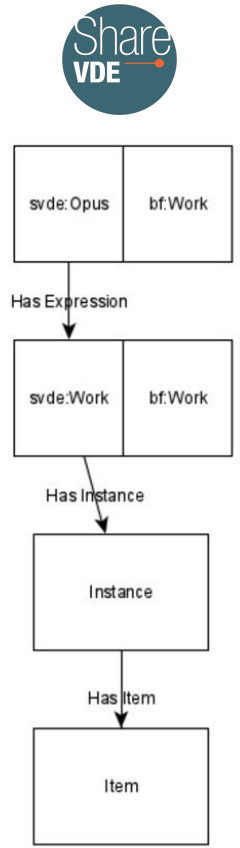
See the SVDE entity model compared to BF and LRM and an example of application of the model



Share-VDE entity model



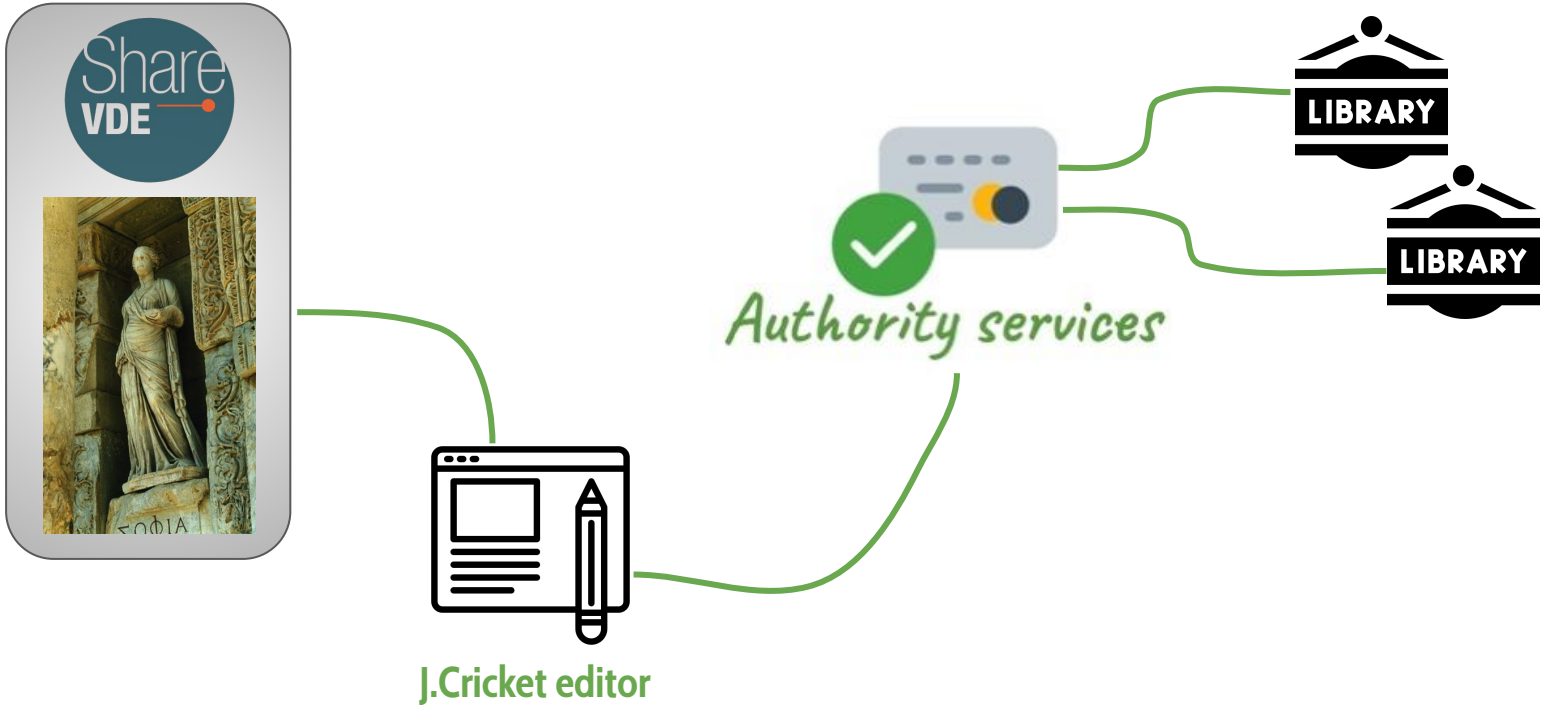




Authority/Identifier Management Services working group - Background

- Defines guidelines and best practices for Authority/Identifier management services in linked data environments
- Defines scope and data-flow for the creation and implementation of automated services based on use cases defined by the group
- Proposes additional use cases identified as essential for effective knowledge base management.
- Works closely with Cluster Knowledge Base Group
- Contains members from several different types of libraries across Europe and North America
- Serves as knowledge sharing space for innovative approaches to authority workflows at each institution

Integration - new Authority services generation



AIMS - Integrations

AIMS - Current and Future Work

- MARC Authority Workflow established and in testing phase at Stanford
- AIMS members currently undertaking review of that service to provide feedback for enhancements
- Additional integrations with other sources of authority data
- Working with LD4 Change Management for Authoritative Data WG
- Development of BIBFRAME authority workflow

The Cluster Knowledge Base and the CKB editor

Sapientia Cluster Knowledge Base:

- Sapientia includes the clusters of entities created in the reconciliation and conversion to linked data of the catalogues of all Share-VDE participating libraries;
- the first release of Sapientia is online and the database is constantly enriched with the new data created by libraries and converted by Share-VDE;
- more than 100 millions of bibliographic records and 24 millions of authority records have been processed;
- Sapientia contains 400 millions triples in its triplestore, and 24 billions quads of converted bibliographic records.

J.Cricket Cluster Knowledge Base editor:

- J.Cricket is the module dedicated to the editing of the SVDE entities, essential for the management of Share-VDE database;
- the initial model of the J.Cricket Cluster Knowledge Base Editor has been analysed from a functional perspective, along with the design of its user interface;
- the result will be a collaborative environment enabling the editing of the entities managed in Share-VDE (works, authors etc.).

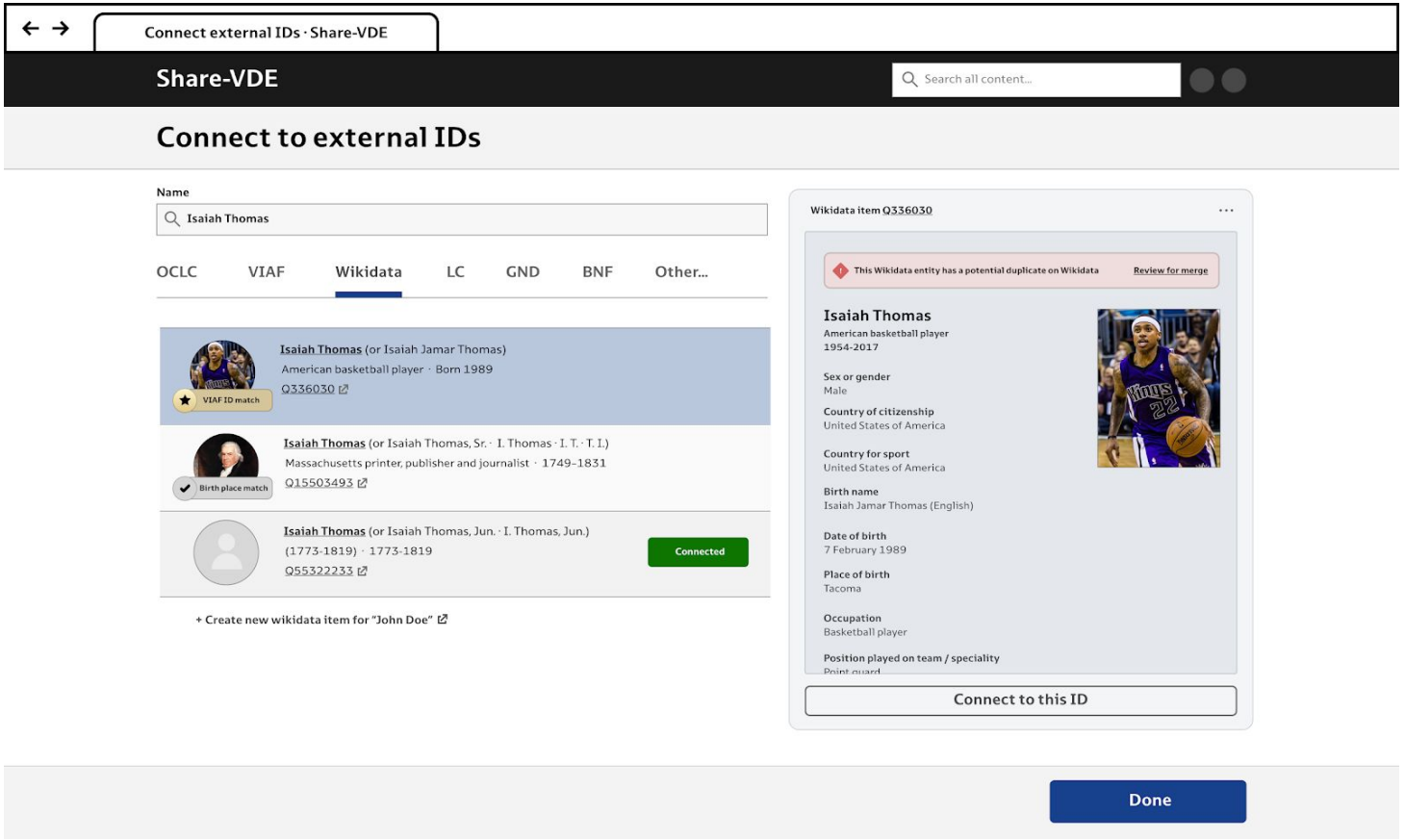
CKB WG Initial Goals

- Review and define use workflow cases for updating and creating entities and relationships
- Propose an editor interface (J.Cricket) to enable management of entity clusters
- Contribute metadata knowledge to the design of the interface

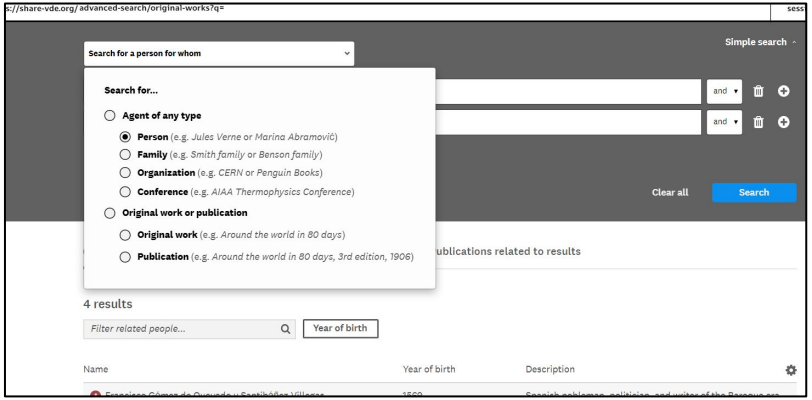
CKB Collaborations

- Worked with Filip Jakobsen (Samhæng) to sketch interface designs for J.Cricket
- Rejoined up with the AIMS group to address crossover functions for J.Cricket
- Reviewed use cases for integration of Wikidata into J.Cricket interfaces sketches

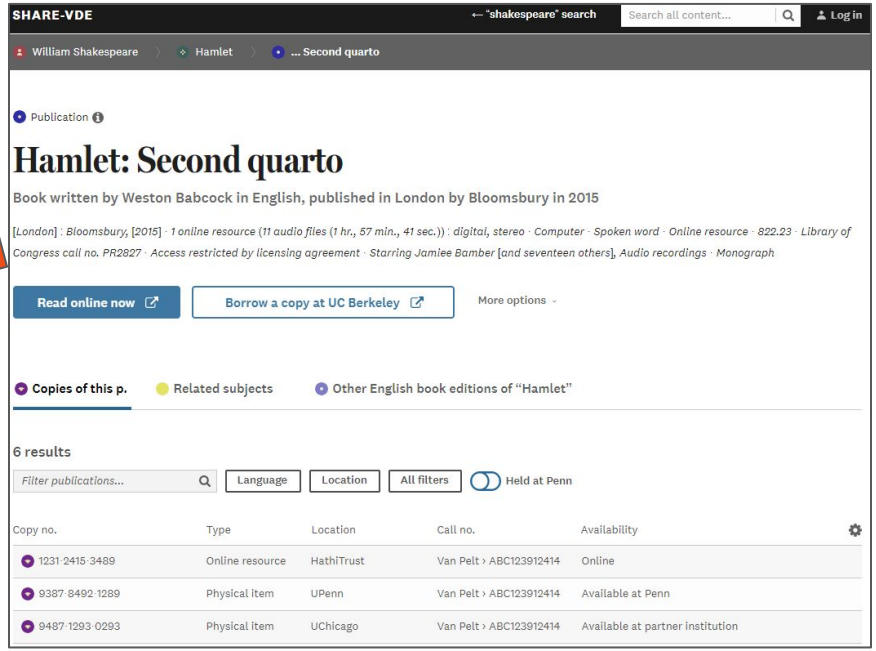
CKB + AIMS


Connect external IDs · Share-VDE
Share-VDE
Search all content...
Connect to external IDs
Name
Isaiah Thomas
OCLC
VIAF
Wikidata
LC
GND
BNF
Other...
Isaiah Thomas (or Isaiah Jamar Thomas)
American basketball player · Born 1989
Q336030
VIAF ID match
Isaiah Thomas (or Isaiah Thomas, Sr. · I. Thomas · I. T. · T. I.)
Massachusetts printer, publisher and journalist · 1749–1831
Q15503493
Birth place match
Isaiah Thomas (or Isaiah Thomas, Jun. · I. Thomas, Jun.)
(1773-1819) · 1773-1819
Q55322233
Connected
+ Create new wikidata item for "John Doe"
Wikidata item Q336030
This Wikidata entity has a potential duplicate on Wikidata
Review for merge
Isaiah Thomas
American basketball player
1954-2017
Sex or gender
Male
Country of citizenship
United States of America
Country for sport
United States of America
Birth name
Isaiah Jamar Thomas (English)
Date of birth
7 February 1989
Place of birth
Tacoma
Occupation
Basketball player
Position played on team / speciality
Point guard
Connect to this ID
Done

Integration - the end user perspective

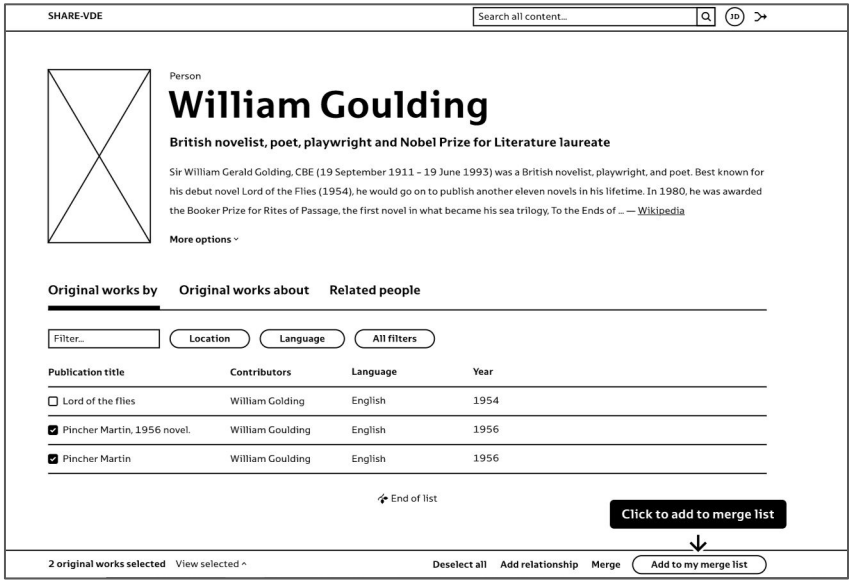


Integration is the core of the SVDE platform that aggregates data from multiple libraries to form clusters of entities.



From the end user perspective: the discovery platform has to accurately represent the entity model, but also provide an intuitive experience, seamless navigation and rich resources to the end users. A new, advanced discovery interface is being developed to harness the potential of linked data.

Integration - J.Cricket: the professional perspective

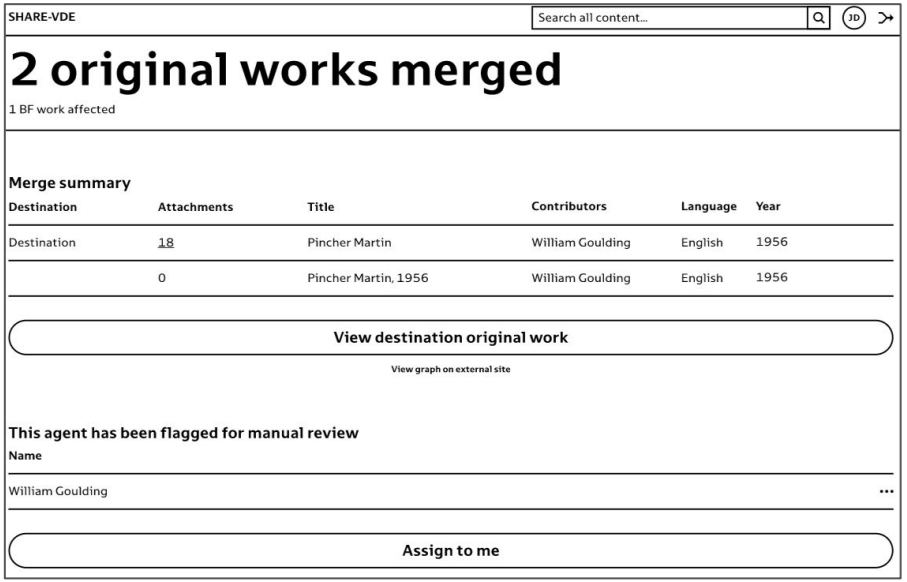


SVDE is evolving from a discovery platform that converts MARC data of libraries in Linked Open Data to an interactive authoritative source providing real services for libraries. This transition is happening through the editor named J. Cricket, that is the new application dedicated to the editing of SVDE data in a collaborative environment.

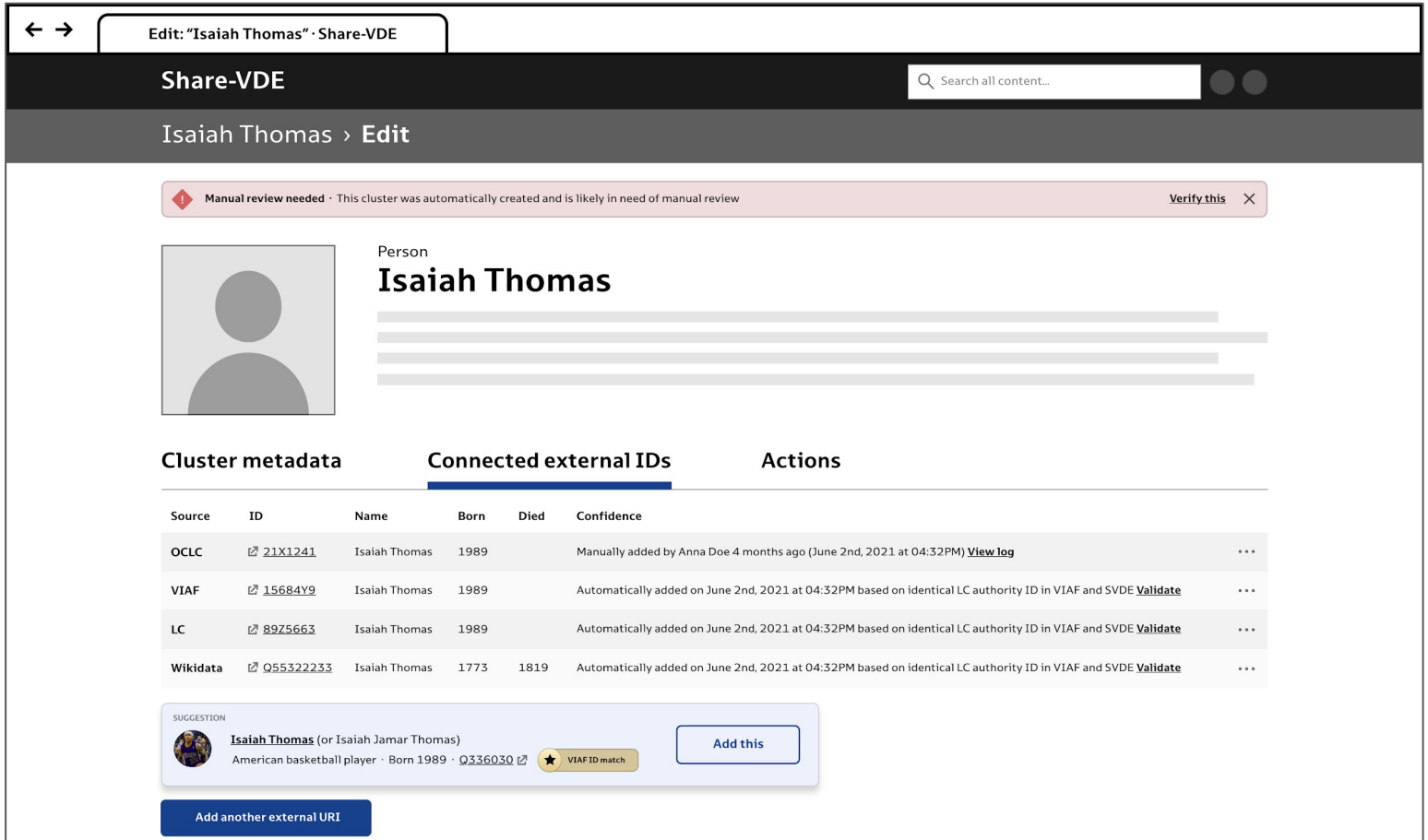
Integration - J.Cricket: the professional perspective

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

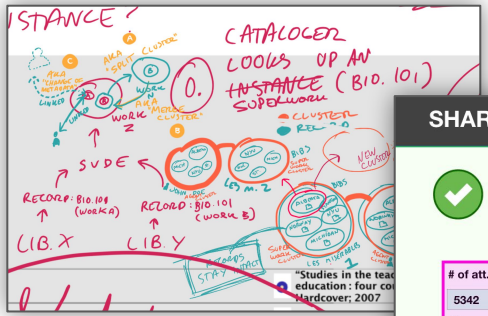
J.Cricket will extend authority capabilities through the integration with external data sources such as Wikidata and ISNI.

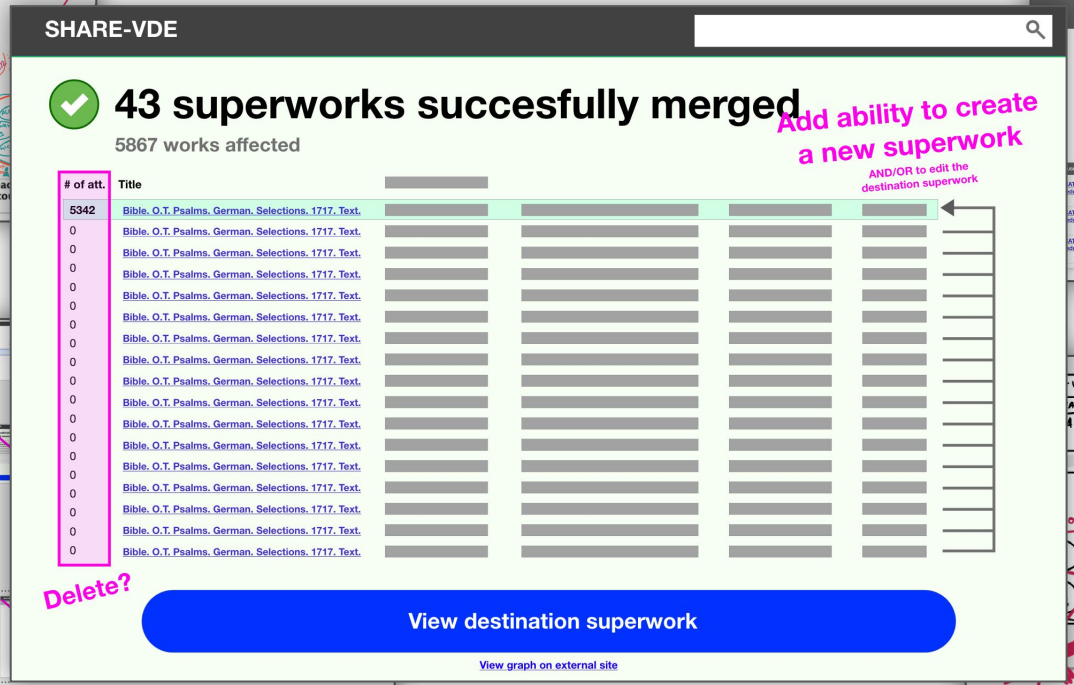


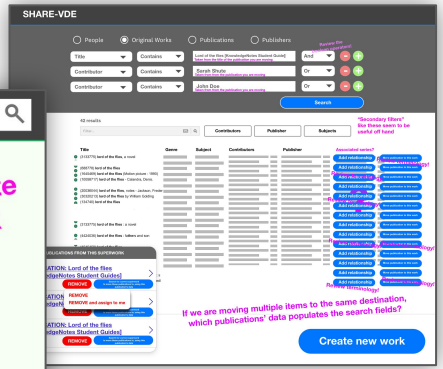
How J.Cricket interacts with Wikidata

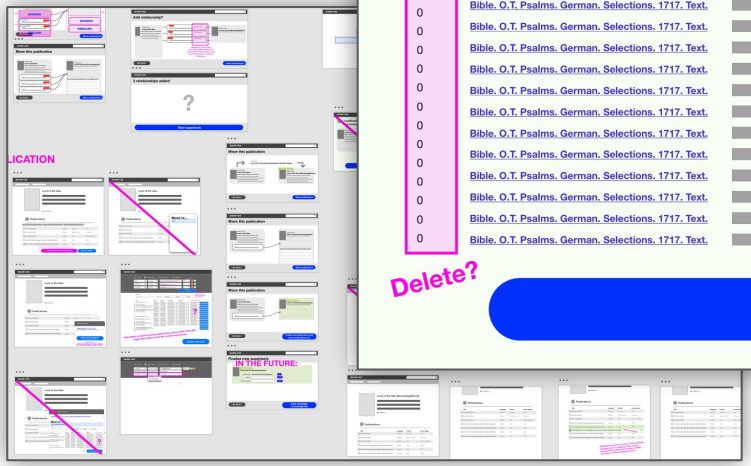


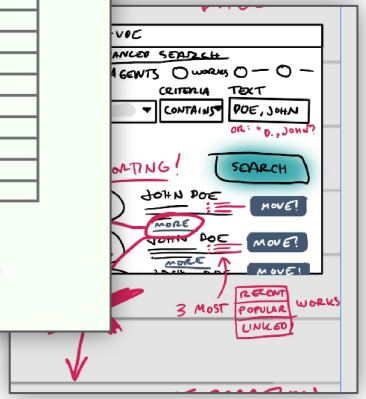
User experience/User Interface Working Group





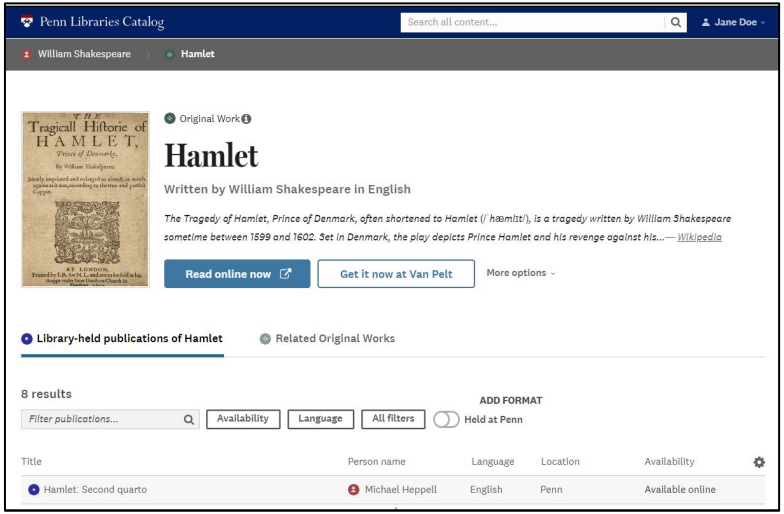




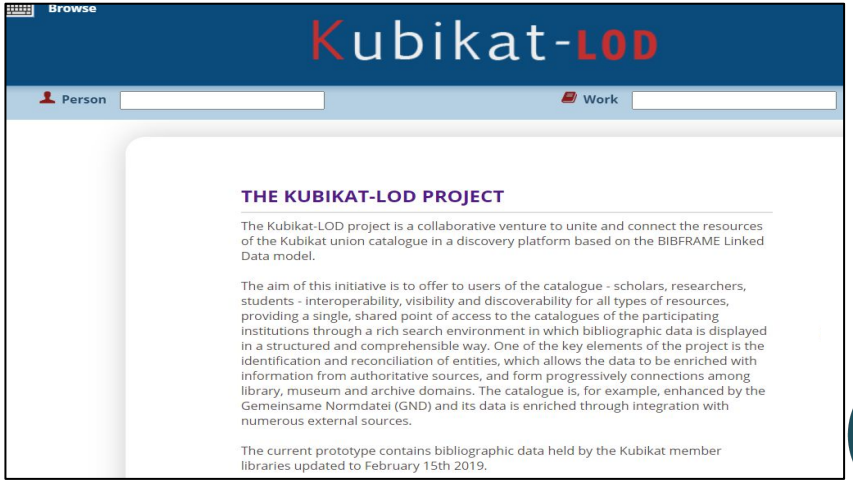


Share VDE

Autonomy - the end user perspective



SVDE localisation for the University of Pennsylvania



Kubikat-LOD project platform

Resources

Resources about Share-VDE and the Share family:

- Share-VDE wiki
- Share-VDE brochure
- Video introducing Share-VDE
- LOD Platform (the technology underlying the Share system)

Heading to New Heights Together - Thanks!