

The Share Family initiative to bring Linked Open Data into practice

November 17th 2023

https://www.share-family.org/ https://wiki.svde.org info@svde.org

Summary

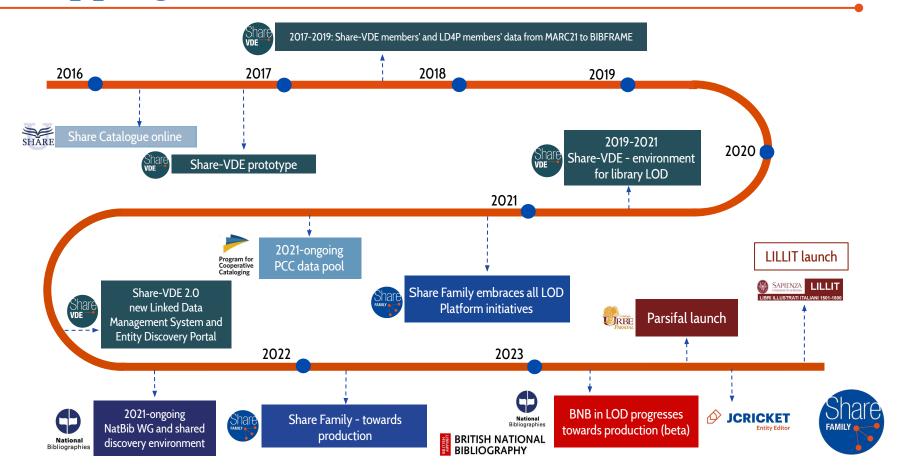
- ✓ Share-VDE background and the Share Family
- ✓ Working Groups and cooperation
- ✓ Towards an operational environment
- JCricket The Entity Management System



Share-VDE background and the Share Family



Stepping stones



The Share Family Linked Data Ecosystem



The Share Family is a global community built on collaboration that brings together libraries, archives, museums, consortia and Library Service Platforms (LSPs) to join their knowledge in an ever-widening network of interconnected bibliographic data.

Share-VDE - Virtual Discovery Environment



https://www.svde.org

Berkeley Law Library

Duke University

Library of Congress

National Library of Finland

National Library of Norway

New York University

Smithsonian Libraries and Archives

Stanford University

University of Alberta / NEOS Library Consortium

University of Chicago

University of Michigan Ann Arbor

University of Pennsylvania

Vanderbilt University

Yale University



Share Catalogue: Scholarly Heritage and Access to Research



Share Catalogue discovery portal

Università degli Studi di Napoli Federico II

Università degli Studi della Basilicata

Università degli Studi del Sannio

Università degli Studi di Salerno

Università degli Studi di Napoli Parthenope

Università degli Studi del Salento

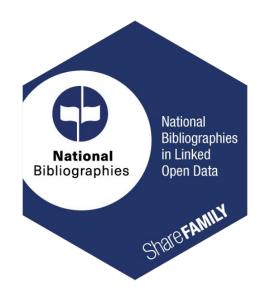
Università degli Studi di Napoli L'Orientale

Università degli studi della Campania Luigi Vanvitelli

Università degli Studi Suor Orsola Benincasa



National Bibliographies in Linked Open Data



https://natbib-lod.org/

The aggregation of data from National Bibliographies in a shared entity discovery environment; the first of these is the BNB - British National Bibliography, soon to go into production.

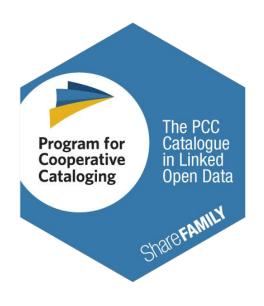


The linked open data BNB beta website is available at

https://bl.natbib-lod.org/



PCC Catalogue in Linked Open Data



The Share Family hosts a dedicated tenant for the data of the PCC - Program for Cooperative Cataloging, to provide PCC-quality BIBFRAME data housed in an ad hoc data pool

https://pcc-lod.org/



Parsifal - Integrated Catalogue in Linked Open Data



https://parsifal.urbe.it/parsifal/?l=en

Accademia Alfonsiana

Centro Pro Unione

Pontificia Facoltà di Scienze dell'Educazione "Auxilium"

Pontificia Facoltà Teologica "Marianum"

Pontificia Università Antonianum

Pontificia Università della Santa Croce

Pontificia Università di San Tommaso d'Aquino (Angelicum)

Pontificia Università Gregoriana

Pontificia Università Lateranense

Pontificia Università Urbaniana

Pontificio Ateneo Sant'Anselmo

Pontificio Istituto Biblico

Pontificio Istituto Orientale

Pontificio Istituto Teologico "Giovanni Paolo II" per le Scienze del Matrimonio e della Famiglia

Pontificium Institutum Patristicum Augustinianum

Università Pontificia Salesiana



Share Art, Share Music, Share MIA



Three pilot projects for shared Linked Open Data environments in the domains of Art, Music and Manuscripts, Incunabula and Ancient books

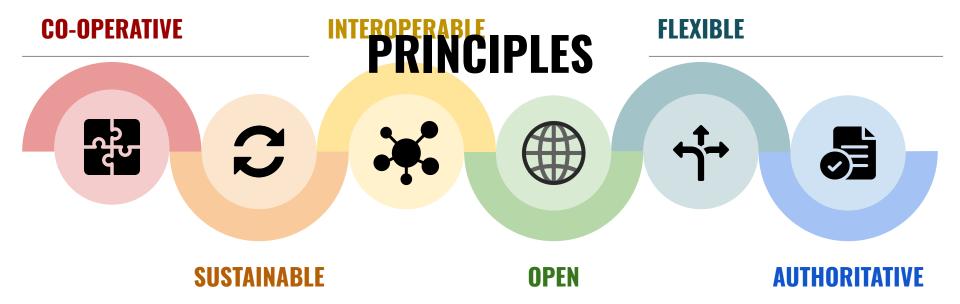


Share Family - Linked Data Ecosystem



The quality of data is guaranteed at left to use a live of dechnical processes and through collaborative data modeling, enrichment and sharing, handled collectively by member organizations.

----- j, --- ---- -----



Share Family - Linked Data Ecosystem



PROCESSES

DATA INPUT FROM INSTITUTIONS

MARC21 bib. and holding

MARC21 aut.

UNIMARC

RDF/BIBFRAME

Other formats (eg. FOLIO)

LOD PLATFORM SERVICES

Mapping

PROCESSES

Reconciliation

Creation of Linked Data Entities

Conversion to RDF/BIBFRAME

RESULTS

Data Publication

- End user discovery portals for each initiative of the Share Family
- Star Datas Pistribution ptron
- (GraphQL, REST, SPARQL)

 API / protocols for third parties integration
- Ska local LSPs and data editor such a Shared Data Management Wikidata, Sinopia BIBFISME editor
- Editing of Share Family entities with
- Ericket Downloadable datasets
- PIRFRAME ADE MARCIENTICH de la company de la company de la company systems
- Authority control in MARC and
- BIBISR OF THURSE and yell ARS representations in local library
- **SF\$te®**hare Family Index: registry of entity URIs

Outputs for consortia or single libraries

Linked Data Descriptions and Enriched MARC Records

- The Library catalogue is converted in linked data entities according to BIBFRAME 2.0. The entities are then enriched both with native and persistent SVDE URIs and URIs from external sources.
- MARC records from the original library catalogue are enriched both with native SVDE URIs and URIs from external sources.

JCricket Entity Editor

- It's a manual collaborative tool designed to manage properties (attributes, relations, and links) of entities in the Cluster Knowledge Base, improving data quality through tasks like creation, merging, and splitting.
- Data can always be traced back to each Institution through the Provenance.
- It can potentially support other workflows and connections with systems external to the Share Family

Discovery Portal

- Advanced entity discovery system based on BIBFRAME
- Customised UI (skin)
- Integration with local APIs
- Site mapping with additional meta-tagging
- Data conversion to Schema.org

Authoritative Services

- Innovative solutions that facilitate and improve authority control through automatic and manual procedures.
- Libraries to receive constantly updates on their bibliographic and authority records from authoritative sources.
- Authority Services currently available for MARC-based workflows offer automated URI enrichment, reconciliation and validation of library data.



Collaboration for enhancing library workflows and services

Member libraries send their records

both for their community of users and librarians

According to their <u>local</u> or <u>network</u> policies, libraries can implement their **services**

high quality data: the CKB includes the <u>clusters of entities</u> <u>created</u> in the reconciliation and conversion to linked data of the catalogues of all Share-VDE participating libraries

Sapientia Cluster Knowledge Base, a collaborative source of

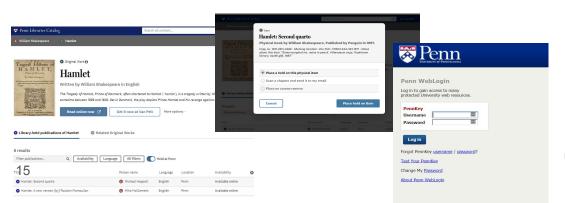
Integration with Other Systems

 Development of APIs for interoperability and cooperation with loca LSPs and third parties (including FOLIO, Wikidata, LD4P - Linked Data for Production)

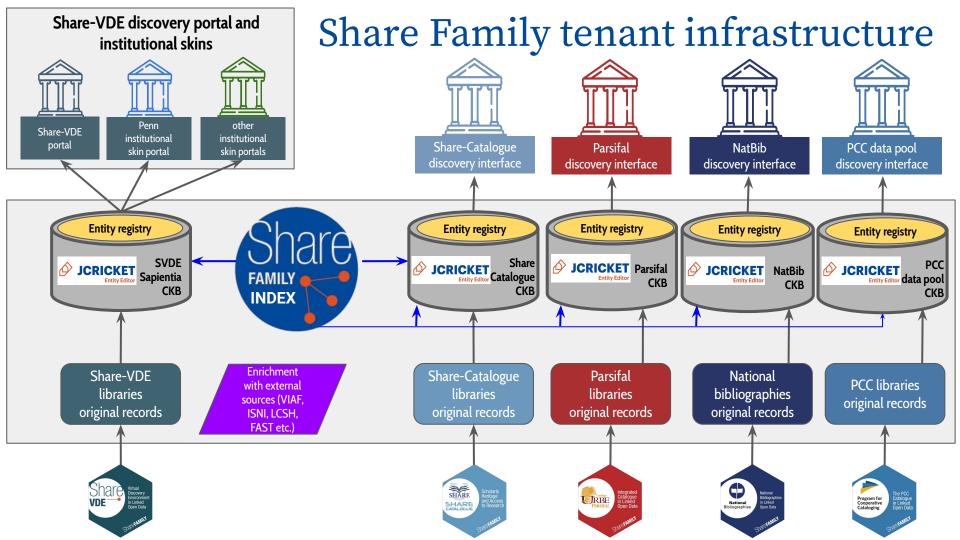
- discovery portal
- authoritative services
- shared cataloguing
- statistics and evaluation
- collection development
- interlibrary loan

and so on!

Data in the CKB is enhanced by JCricket Entity Editor







Default configuration: SVDE and PCC data pool

Simple search default configuration on SVDE and PCC data pool portals



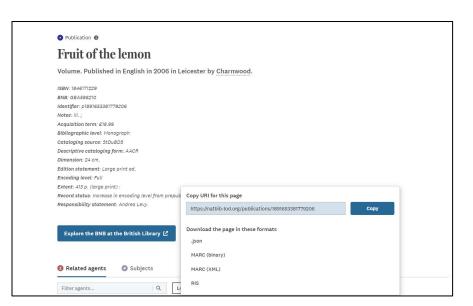




Default configuration: British National Bibliography

Simple search default configuration on Natbib tenant and the BNB - British National Bibliography skin*





* the British Library announced the launch of the beta version of the British National Bibliography Linked Open Data Portal



Default simple search configuration: the BNB



Simple search default configuration on <u>Natbib tenant</u> and the <u>BNB - British National Bibliography skin</u>* is set to Publications search, instead of the SVDE default.

This was done to comply with a different requirement whereby for the data stored in this tenant (ie. national bibliographies) it's meaningful to direct users to publications.

Different communities or types of institutions might need customised features



Share Family tenant infrastructure

- The Share Family of initiatives includes different branches and sister projects, supported by the same <u>LOD Platform technology</u>. Each branch or project is hosted in a specific tenant of the system architecture with a corresponding specific Cluster Knowledge Base and a dedicated web entity discovery portal.
 - For more details on the Share Family tenant infrastructure see the <u>Summary of Share Family</u> <u>tenants</u>.
- In some cases, within a single tenant a customised skin (ie. a sub-portal of the main entity discovery) can be created to address ad hoc needs of an institution, or group of institutions, willing to expose only their own data or to integrate local services in the Share environment.
 - For example, Share-VDE entity discovery portal at svde.org is one of such tenants, including a pool of data from a number of institutions, and the respective skin portals.

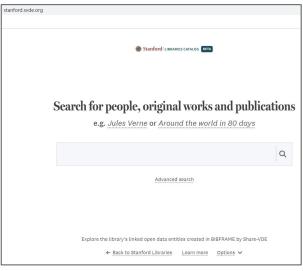
Institutional skin portals within a tenant

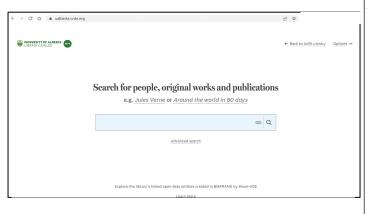
- While the main entity discovery portal of a tenant shows the data of all the institutions feeding the tenant's Cluster Knowledge Base, the skin portal gives the ability to filter only the data of the institution that the skin portal has been designed for.
- To this aim, the "held at" filter was added, allowing to filter publications by what is available at the current library. It is enabled on skin portals at Publication (= Instance) level in these cases:
 - in advanced search, see e.g. NYU data pre-filtered here <u>https://nyu.svde.org/advanced-search/publications?q=(title+does not contain+xyz)&heldAtLibrary=true</u>
 (see the toggle on the right of the screen, you can turn it on / off)
 - o in the Original work entity page that lists Publications, see e.g. https://nyu.svde.org/suite-de-la-mancha-flute-cello-piano-unknown-author-o781654264663247/library-publications (see the toggle on the right of the screen, you can turn it on / off)
 - o in simple search results in cases where the simple search default on the home page is the Publication simple search (e.g. <u>Natbib tenant</u>)

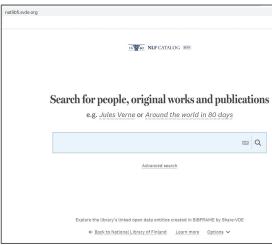


Example: institutional skin portals in SVDE

- SVDE tenant https://svde.org => with LC's authority data and the bibliographic data of member institutions
 - skin portals including: <u>Penn</u>, <u>Smithsonian</u>, <u>Stanford</u>, <u>University of Alberta</u>, <u>New York University</u>,
 <u>National Library of Norway</u>, <u>National library of Finland</u> (other skin portals will be set up following the load of libraries' catalogues to svde.org)









Working Groups and cooperation



Share-VDE and Share Family Working Groups

<u>Member institutions</u> of Share-VDE and Share Family <u>Working Groups</u> and parallel projects are constantly contributing with their Subject Matter Experts to requirements gathering, functional analysis and feedback to developments.

Share Family Working Groups:

- National Bibliographies Working Group involving SVDE members and external institutions
- Italian group for the conversion UNIMARC -BIBFRAME
- discussions in the field of photo libraries and audio-visual collections

Share-VDE Advisory Council and Working Groups:

- Share-VDE Advisory Council
- Sapientia Entity Identification WG
- Authority/Identifier Management
 Services WG (currently on hold)
- Cluster Knowledge Base Editor WG (currently on hold)
- User experience/User Interface WG



Share-VDE Advisory Council

The <u>Share-VDE AC</u> takes an active role in determining future uses and vision for the Share-VDE initiative; develop future use cases for Share-VDE, and set development priorities as needed; monitor and lead the work of the various Advisory Council Working Groups; maintain communication among the Share Family member institutions.

- Share-VDE Executive Summary, December 2022, summarising the scope of Share-VDE in the context of Linked Open Data for Libraries;
- Share-VDE Statement, September 2021, explaning the position of SVDE in the broader context of Library Linked Open Data;
- Library and community events sub-group, dedicated to monitoring conferences/events/initiatives of interest for the Share community, and to submitting proposals for presentations as appropriate;
 - see the SVDE wiki Resources page for details about SVDE presentations at conferences and events.



User Experience/User Interface WG

The <u>UX-UI WG</u> has redesigned the Share-VDE user interface to meet the requirements and expectations of both patrons and library staff. SVDE 2.0 entity discovery interface:

- reflects the components of the Share-VDE data model infrastructure;
- harnesses the potential of linked data and delivers wide-ranging and detailed search results;
- provides an intuitive user experience, hiding the complexity of the underlying data model;
- embeds partner APIs for interoperability with local library services (e.g. lending);
- allows dedicated skins, ie. customised sub-portals dedicated to individual institutions.

- Share-VDE 2.0 Entity Discovery https://svde.org
- new Entity Discovery Portal and new back-end infrastructure for the Linked Data Management;
- other Share Family discovery portals supported by the same technology;
- review and enhancement of portal features, in conjunction with the National Bibliographies
 Working Group.



Sapientia Entity Identification WG

The <u>SEI WG</u> reviews the use of entities, identifiers, and associated modelling in the Sapientia CKB; evaluates and refines processes for Sapientia entity clustering in Share-VDE and the creation of associated open and stable URI for use in Share-VDE and in the library community; reviews MARC to BIBFRAME and BIBFRAME to MARC conversion; engages with the library community to outline and/or develop best practices for use of Sapientia identifiers in BIBFRAME and MARC data.

- definition of the <u>SVDE Ontology</u>; see also Jim Hahn's presentation at the <u>BFWE 2023</u>;
- svde:Work is subclass of bf:Work \rightarrow this ensures interoperability;
- review of clustering and conversion rules;
- cooperation in the IFLA context: the mapping UNIMARC-BIBFRAME is being prepared and a formal liaison with SVDE has been approved by the IFLA Bibliography Section Standing Committee.



National Bibliographies WG

The <u>National Bibliographies Working Group</u> is dedicated to facilitating practical cooperation among National Bibliographies, and addressing the needs of National Libraries and institutions that hold National Bibliographies within the framework of a shared entity discovery environment such as the Share Family of initiatives.

- overview document <u>National Bibliographies Share Family initiative 2022-June.pdf</u>
- involvement of SVDE / Share Family members and external institutions;
- IFLA Bibliographic Section liaison (Maud Henry from KBR Royal Library of Belgium);
- discussion on topics of interest related to hosting national bibliographies as an ad hoc tenant;
- main tenant of the shared discovery environment for national bibliographies: https://natbib-lod.org/;
- implementation of the skin portal for the BNB British National Bibliography https://bl.natbib-lod.org (this is a preview of a beta site);
- joint work with the SVDE UX-UI working group to design end user services and user interface/discovery features.

Authority/Identifier Management Services WG

The <u>AIMS WG</u> defines guidelines and best practices for Authority/Identifier management; describes 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.

- definition of use cases;
- functional analysis;
- study of the interaction with Wikidata and ISNI (joint work with CKBE WG to design JCricket functionalities);
- pilot of MARC-based authority services with Stanford University Libraries;
- assessment of services for authority control in linked data workflows.



Cluster Knowledge Base Editor WG

The <u>CKBE WG</u> examines how libraries interact with the *Sapientia* Cluster Knowledge Base (CKB) and their use of the JCricket Editor for modifying (correcting / enriching), deleting, merging and separating clusters.

- back-end developments for JCricket entity editor are completed, front-end features in progress;
- definition of use cases;
- design of manual editing features;
- analysis of interaction with Wikidata and ISNI to be incorporated into JCricket and authority dataflows that feed the Cluster Knowledge Base (joint work with AIMS WG to design JCricket functionalities).



Towards an operational environment



An integrated and hybrid environment

The mutual exchanges in the BIBFRAME / linked data community are bringing the Share Family towards:

- an integrated, "hybrid" operational environment...
- ...based on a variety of tools and diverse data sources...
- ...including traditional workflows (eg. new authority services for MARC workflows) as well as advanced models for data exchange



An integrated and hybrid environment

DONE ONGOING EXPLORING

New authority services for MARC-based workflows - designed with SVDE AIMS working group and Stanford's input

Third parties integration with ILSs/LSPs - local library services (eg. Alma circulation APIs)

Finer granularity level of the CKB - Cluster Knowledge Base to make it format-agnostic and extend input data capabilities (MARC21, UNIMARC, BF/RDF eg. Sinopia profile etc.)

Third parties integration with ILS/LSP - Sinopia

Third parties integration with ILS/LSP - FOLIO

Third parties integration with authority sources:

- LD4P Questioning Authority
- Wikidata (initial specs by SVDE working groups)
- ISNI (initial specs by SVDE working groups)

Application of Share Family tech to other domains (Art, Music)



Focus on Authority services

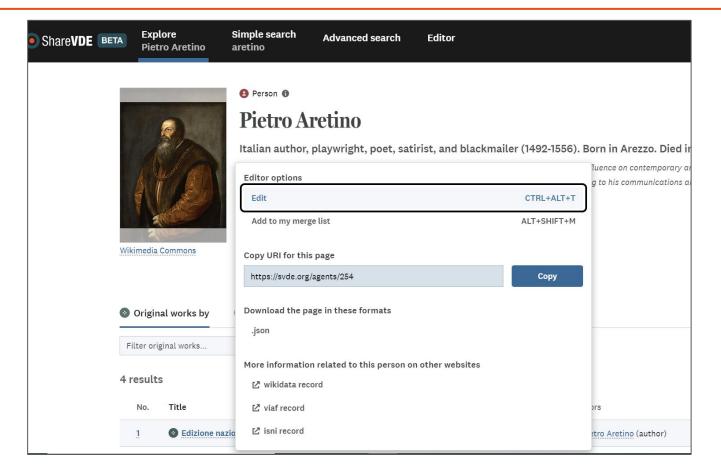
Automatic services for Share Family libraries, piloted by Stanford University:

- validation of MARC 21 bibliographic records (correction of MARC 21 fields and obsolete forms, update of tags and subfields etc.);
- enrichment of MARC 21 fields with SVDE original URIs and URIs from external sources according to ad hoc profiling, including LCNAF, VIAF, ISNI;
- matching processes on external authority files (LCNAF, LCSH, LCGFT, FAST);
- import of authority records from external authority files (LCNAF, LCSH, LCGFT, FAST);
- reporting features providing complete details of the validation and corrections done to the records.

Next step developments: Authority Services fully integrated in the Linked Open Data environments.

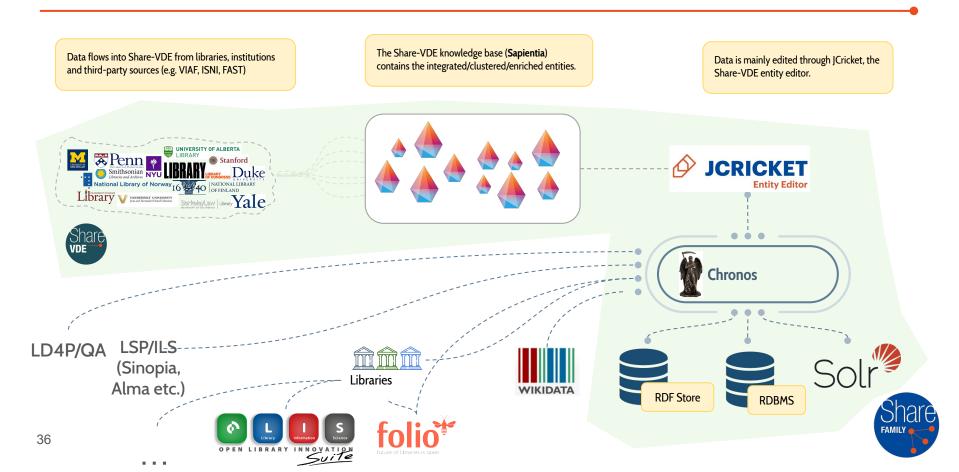


Enhance data quality and authority control

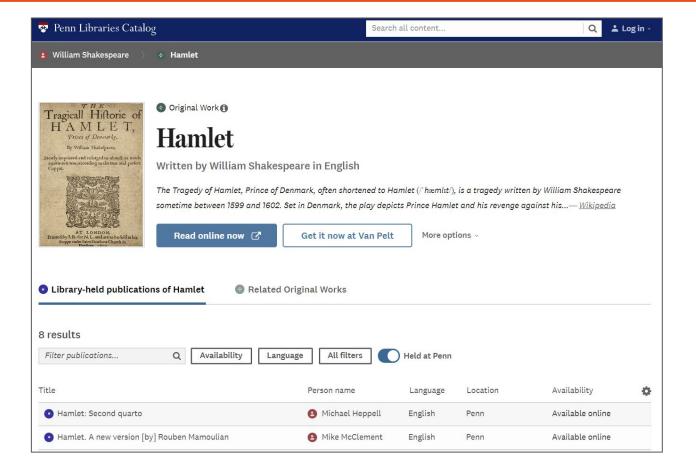




Third party integration - Outbound Connectors Architecture



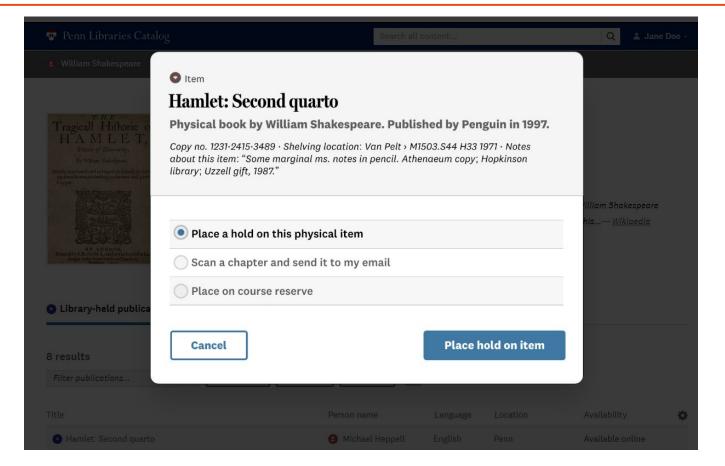
Integration with local services - circulation info



Integration with local services, e.g. connection to Alma APIs for <u>University of Pennsylvania</u> circulation services



Integration with local services - lending





SINOPIA integration: high-level milestones

- set-up the connector to fetch data from Sinopia
- ingested subset of Sinopia data from Stanford
- now creating the parser so that RDF data coming from Sinopia can be clustered by Share-VDE processes
- ot the end of this process, Sinopia data will be included in the Share-VDE CKB Cluster Knowledge Base



folio integration: high-level milestones

See a possible model for ILS/LSP interaction through FOLIO

Level 1: Instance correlation

- Folio inventory instances are retained in dedicated faces of Share-VDE prisms
- The inbound connector receives FOLIO data (instances) and feeds the Cluster Knowledge Base (CKB)
- The outbound connector communicates back data changes to FOLIO

Level 2a: Agents (and works) correlation

Same interaction as above, but using authority records (agents, works), instead.

Level 2b: JCricket UI App in FOLIO

Using the FOLIO built-in "pluggable" nature, the FOLIO UI SDK and the Share-VDE (GraphQL) API





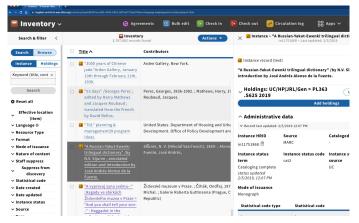
The Entity Management System

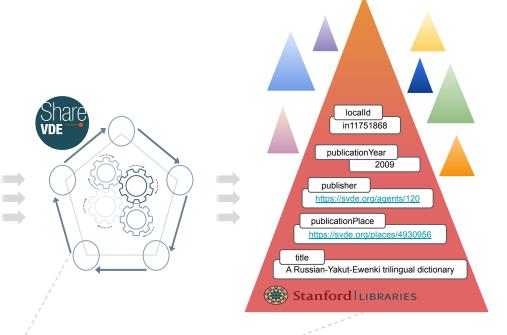


From Library Data to Share-VDE

A Share-VDE member (Stanford, in the example) uses FOLIO for managing its data.





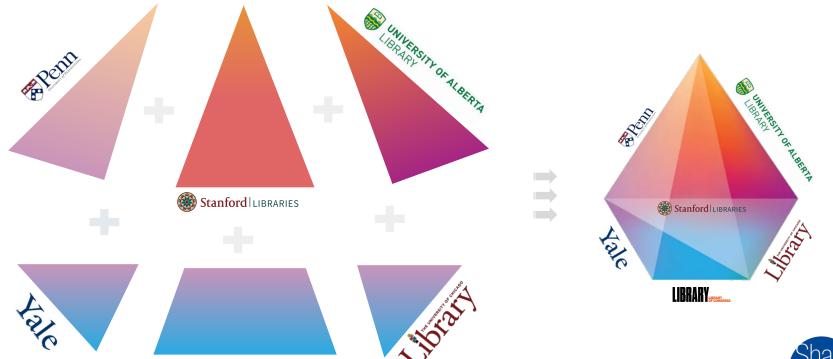


FOLIO instance (or instances in case of massive export) is sent to Share-VDE

FOLIO instance data is split across the entities that form the Share-VDE domain model. In this example we focus on the properties that are assigned to a Share-VDE instance (red triangle above)

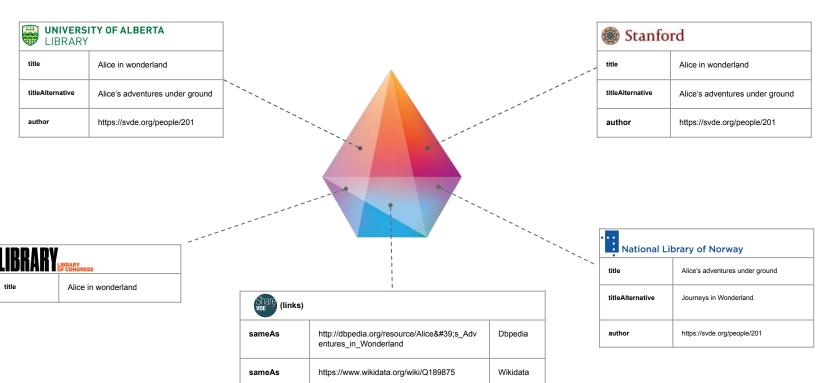


Prism, faces: the Share-VDE Entity





Faces (aka Contributions & Provenances)



https://data.bnf.fr/ark:/12148/cb358500385#a

bnf

sameAs

bout



Properties: Attributes, Relationships, Links



Name	Value	Provenance
title	Alice in wonderland	LIBRARY WRANT Stanford
titleAlternative	Alice's adventures under ground	LIBRARY OF CONGRESS
titleAlternative	Journeys in Wonderland	National Library of Norway

An attribute is a data property, having a literal as value

Share			
sameAs	http://dbpedia.org/resource/Alice's_Adventures_in_Wonderland	Dbpedia	
sameAs	https://www.wikidata.org/wiki/Q189875	Wikidata	•
sameAs	https://data.bnf.fr/ark:/12148/cb358500385#about	bnf	

A link is a connection between a Share-VDE Prism and an external reference

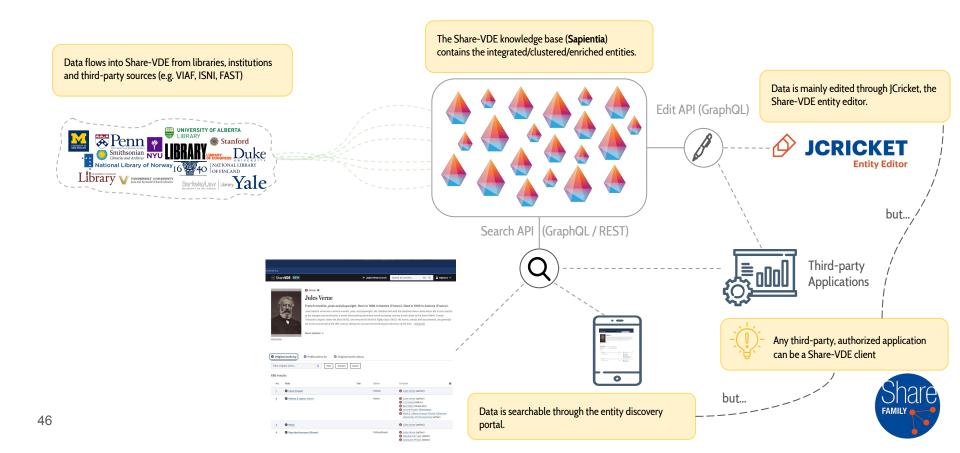
Name	Provenance
author	LIBRARY SPRANKARES (Stanford







The Big Picture: Genesis, Search, Edit



JCricket: Available Operations

Edit: a property is added/updated/deleted



Lewisss Carroll

Lewiss Carroll

is author of

ttps://svde.org/opuses/1827349

https://svde.org/opuses/920302

Invalidate







Merge: multiple prisms are merged into one







For example, two prisms, "Mark Twain" and "Samuel Clemens", should be actually part of the same entity.

Split: a prism is split into multiple prisms



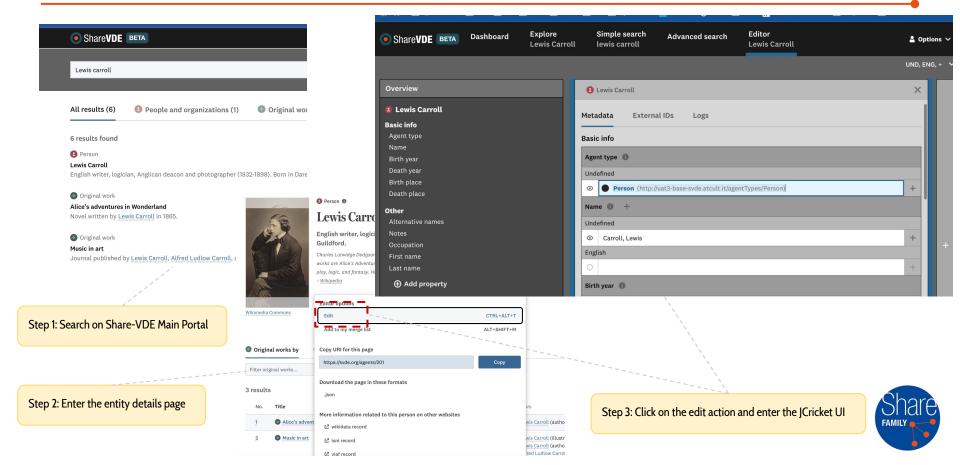




A prism (wrongly) contains information belonging to multiple entities (e.g., "Wallace David" and "David Wallace")



JCricket user interface



What JCricket is



- it's a linked data entity / authority editor
- it applies to linked data entities created within all tenants of the Share Family (svde.org, pcc-lod.org, natbib-lod.org)
- it's a manual application that manages properties (attributes, relations and links) of entities in the CKB Cluster Knowledge Base
- it's a collaborative tool shared across member institutions
- it can be a new tool for entity sharing in LOD



What JCricket is not



- not a traditional bibliographic data editor
- not an original cataloguing tool
- not in contrast with Sinopia or Marva
- not impacting original data that reside in member libraries' systems (unless libraries want to use ad hoc APIs for entity updates both in SVDE and in their systems)



Next generation cataloguing



The JCricket editor is an example of how the LOD Platform technology, within the Share Family Linked Data Ecosystem, is pursuing a new way of managing library cataloguing in a cooperative way:

- ★ aggregation of data from multiple sources
- ★ managed through standard protocols (linked data)
- ★ in a collaborative and integrated environment
- ★ that makes available open data and resources
- ★ to end users and professionals (researchers, scholars etc.)
- ★ for reuse in the library community and beyond



Where we are now



- ★ The back-end APIs that manage JCricket behind the scenes are ready
- ★ The respective front-end functions for the end users to actually use JCricket are under development ∠



JCricket references

Useful references:

- JCricket overview https://wiki.share-vde.org/w/images/7/74/JCricket_overview 2023-Jan-26.pdf
- for more technical details on JCricket https://wiki.share-vde.org/w/images/e/e8/|Cricket_entity_editor_presentation.pdf
- on how JCricket has been conceived
 https://wiki.share-vde.org/w/images/b/b3/Share-VDE perspective on Cluster Knowledge Base and Provenance.pdf



See the live demos!

JCricket premiered its demo at the ALA Conference 2023, during the Share-VDE Workshop:

https://bit.ly/SVDE_Workshop2023_recording

|Cricket demo at the LD4 Conference 2023:

https://www.youtube.com/watch?v=wbrqvWGnvfl

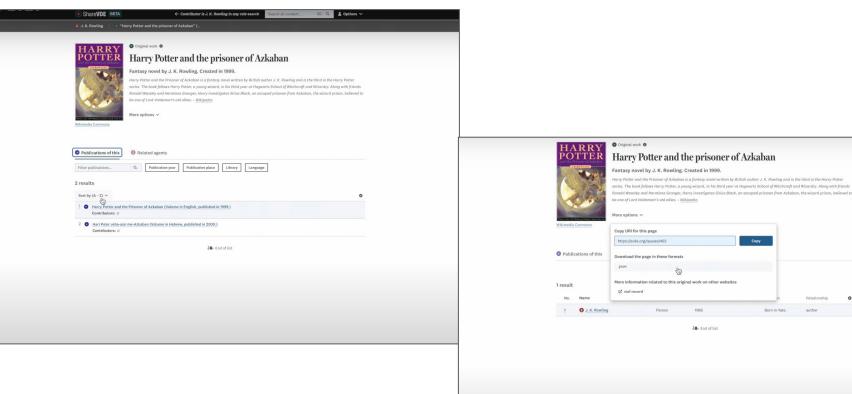
Share-VDE discovery portal general demo: https://bit.ly/SVDE-discovery-live-demo



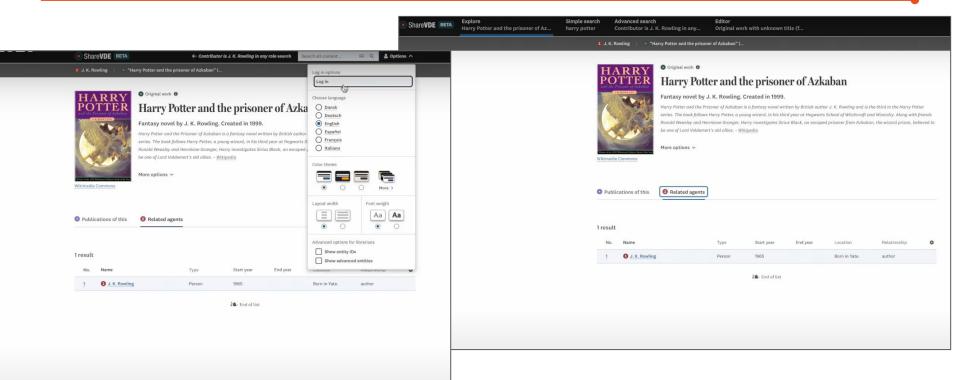
JCricket Editor - examples from the user interface



JCricket features - The user interface

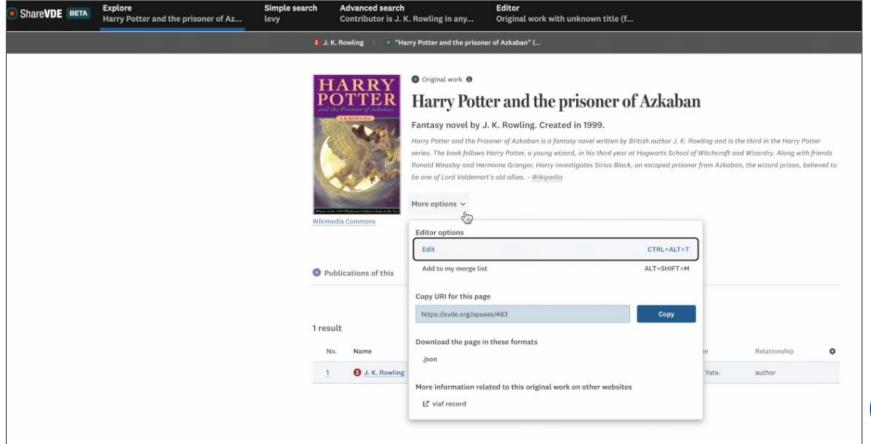


JCricket log-in



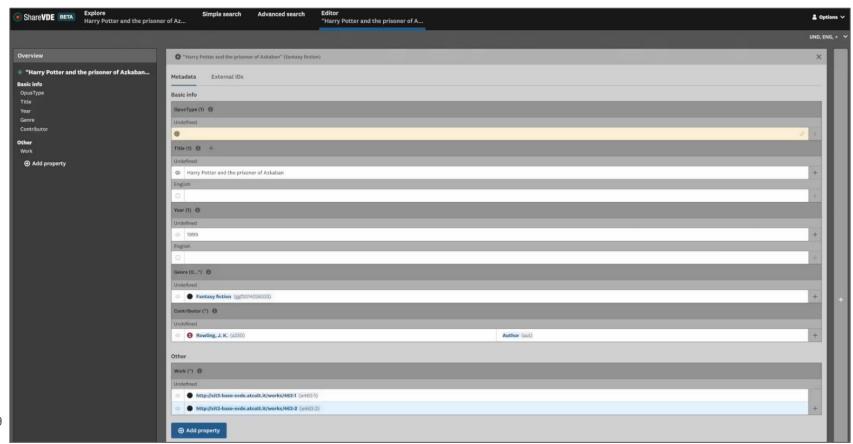


JCricket feature selection

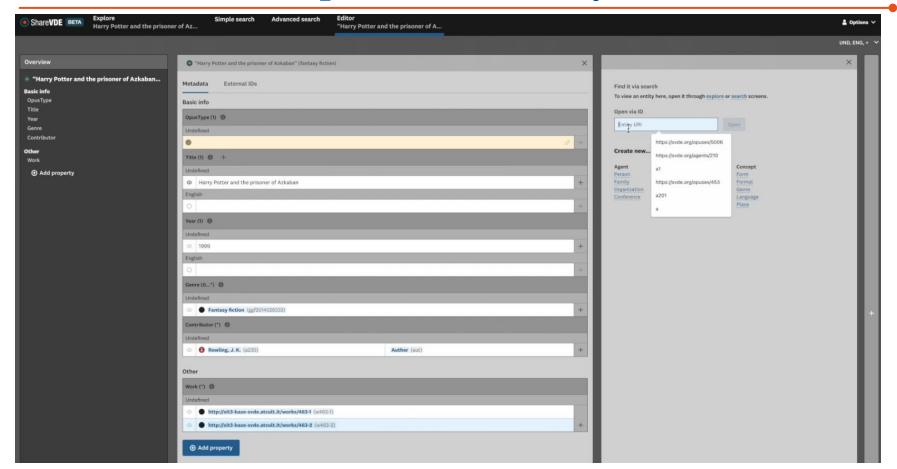




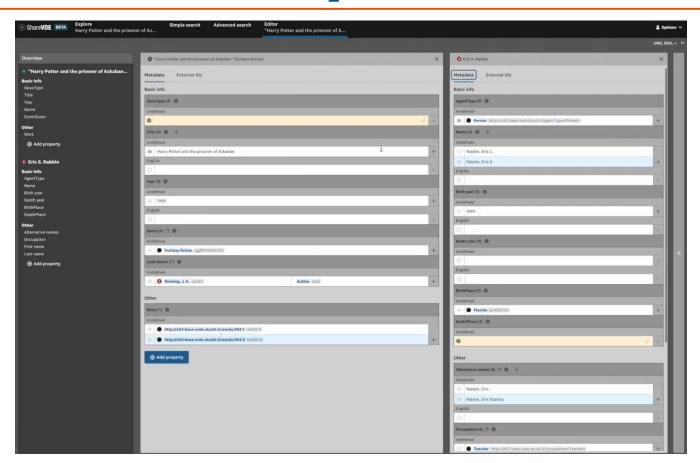
JCricket edit



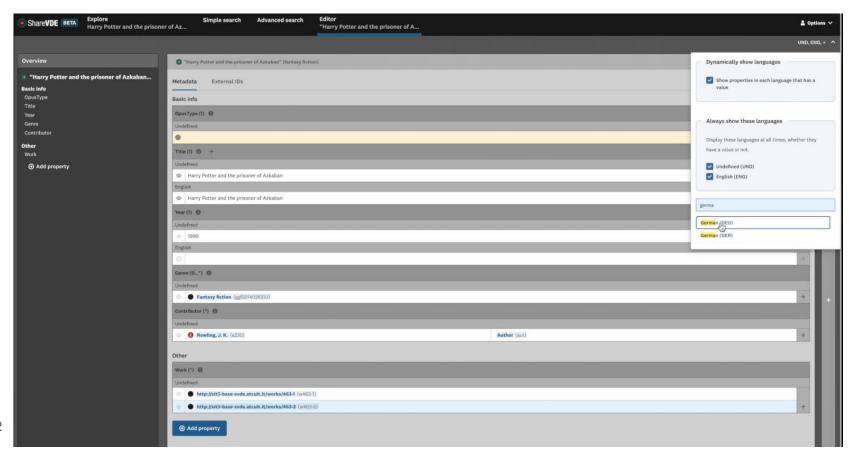
JCricket edit - open a new entity via ID

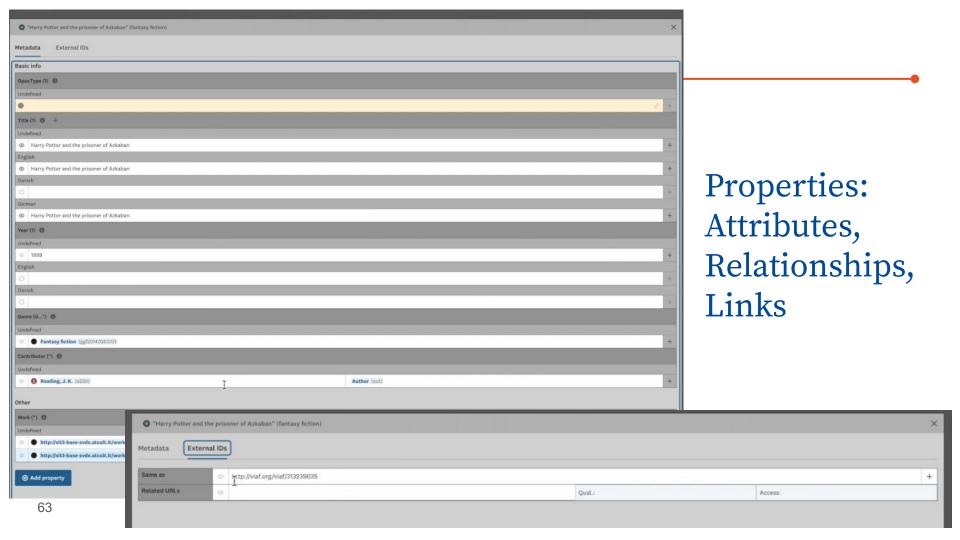


JCricket edit - Multiple entities in one screen

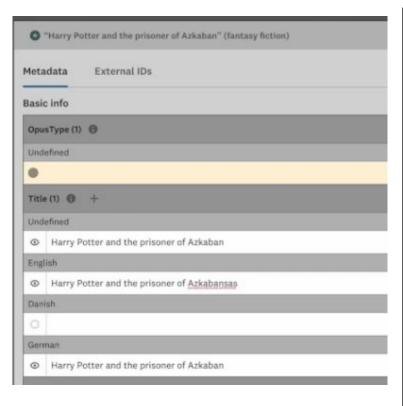


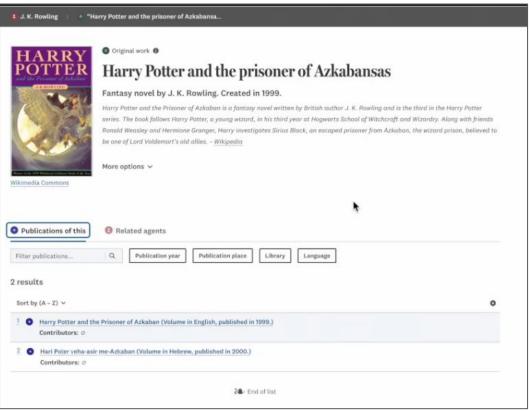
JCricket edit - languages



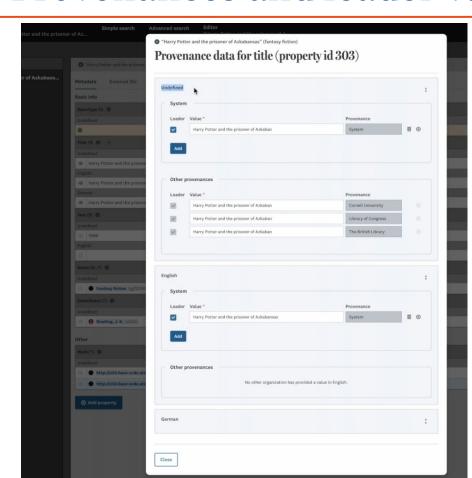


JCricket edit - Add variant forms

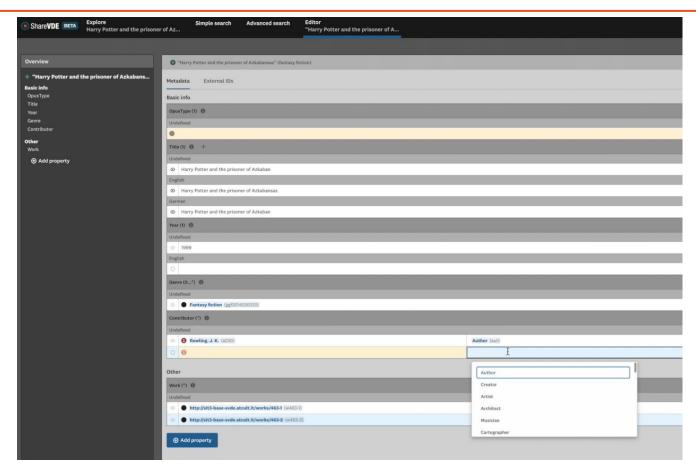




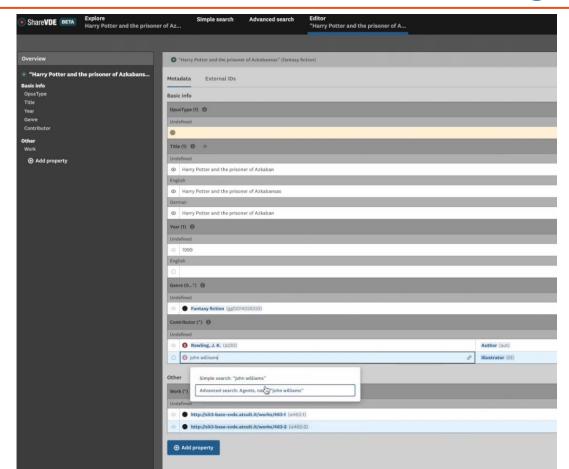
Provenances and leader value for end users



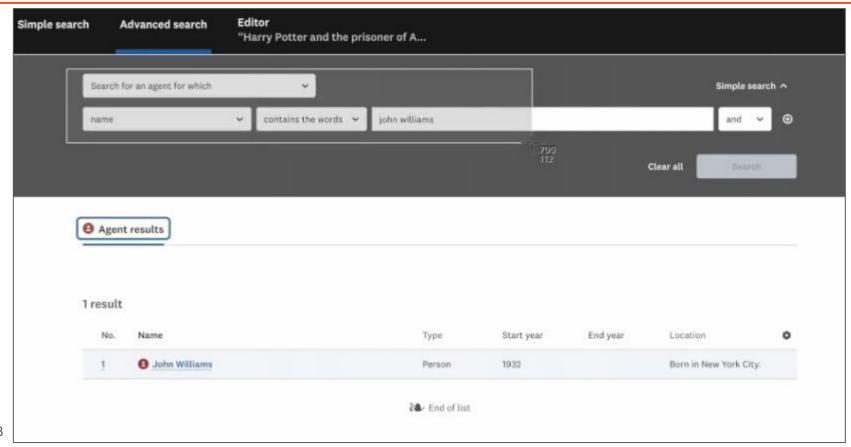
JCricket edit - controlled value list



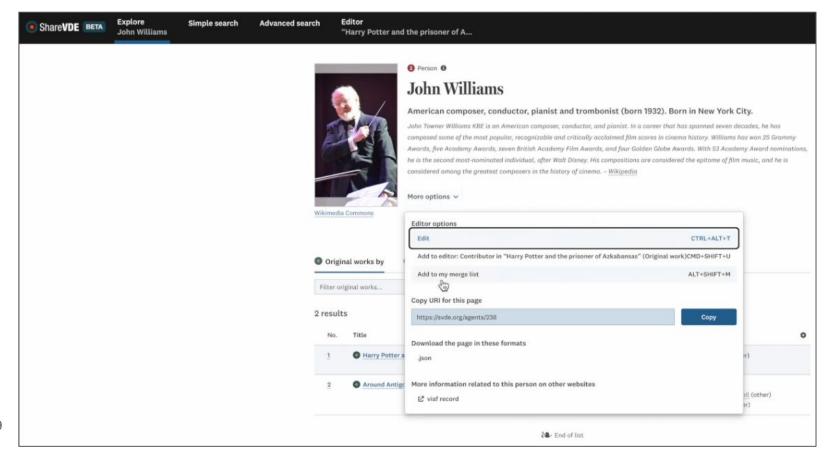
JCricket edit - search while editing



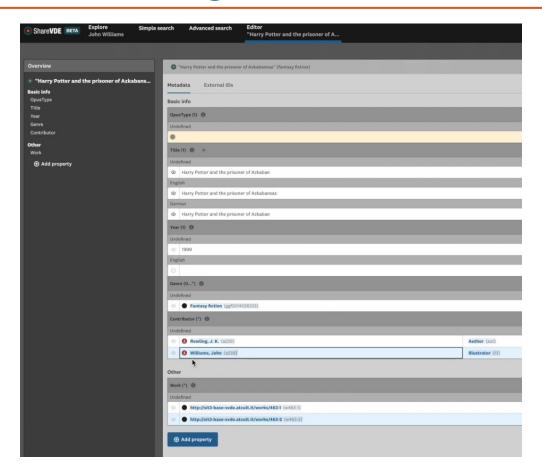
JCricket edit - search while editing



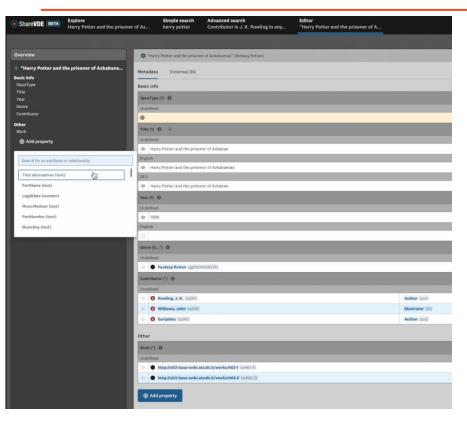
JCricket edit - search while editing

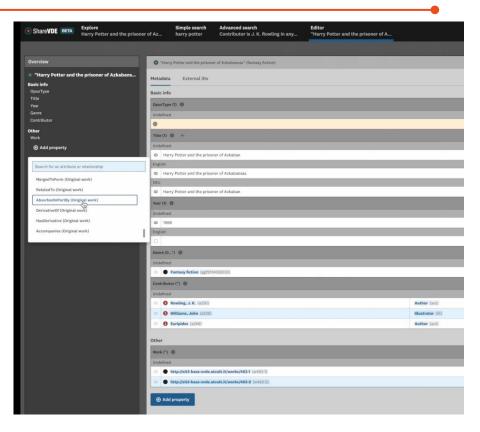


JCricket edit - add agent

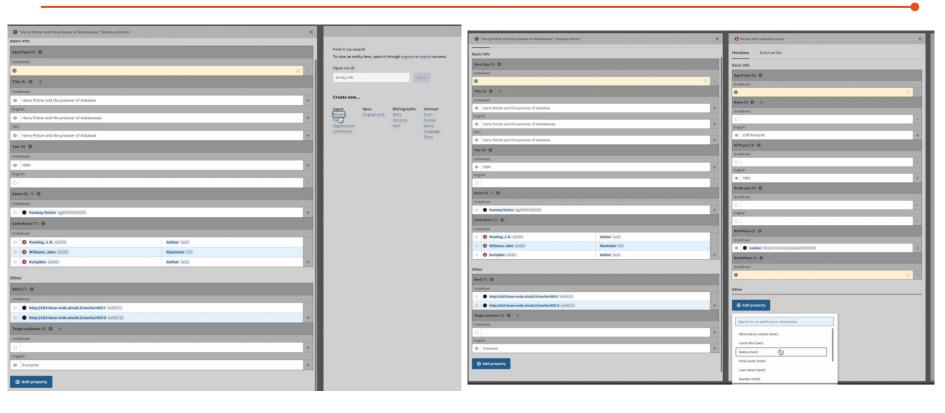


JCricket edit - add attribute



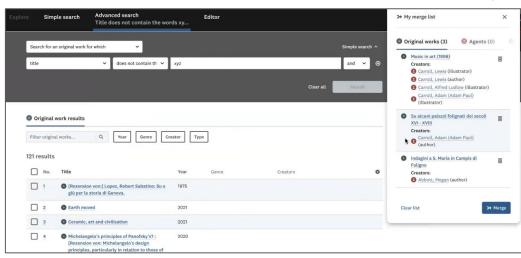


JCricket - create an entity from scratch

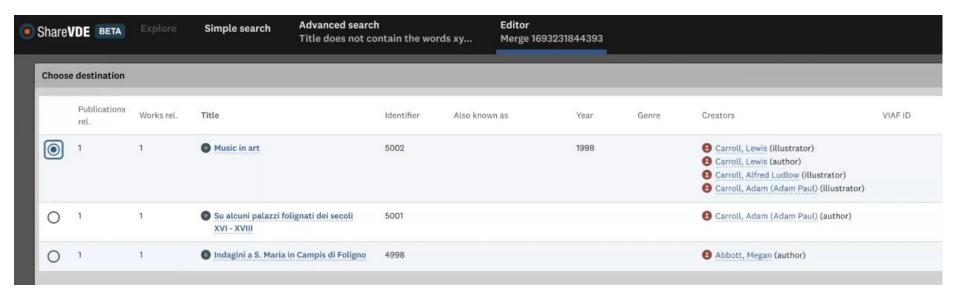


JCricket merge - select merged entities

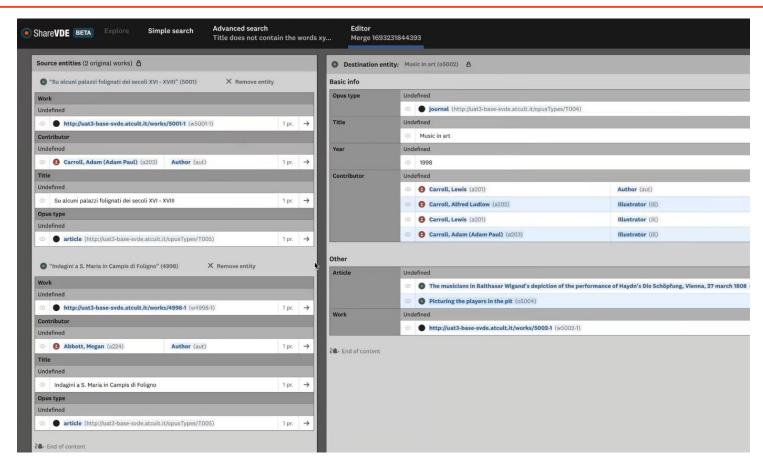




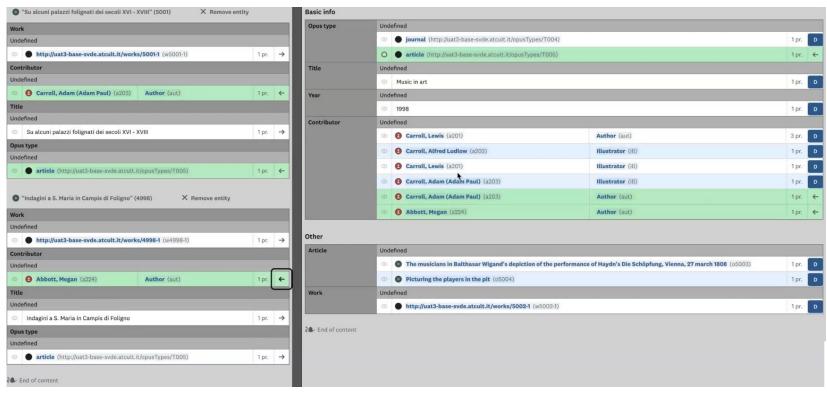
JCricket merge - select destination entity



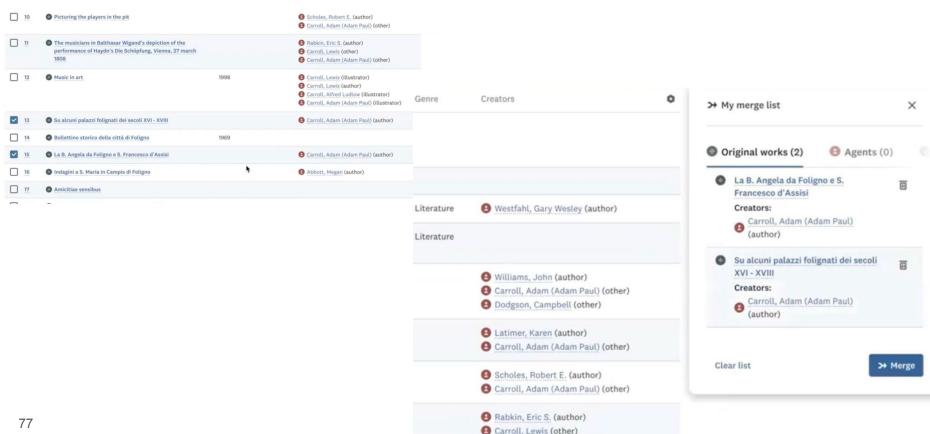
JCricket merge - select properties



JCricket merge - select properties

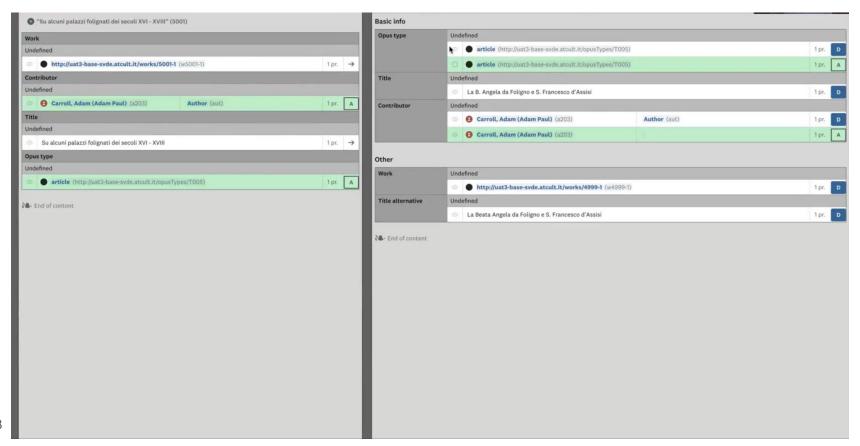


Merge entities where the role is the same

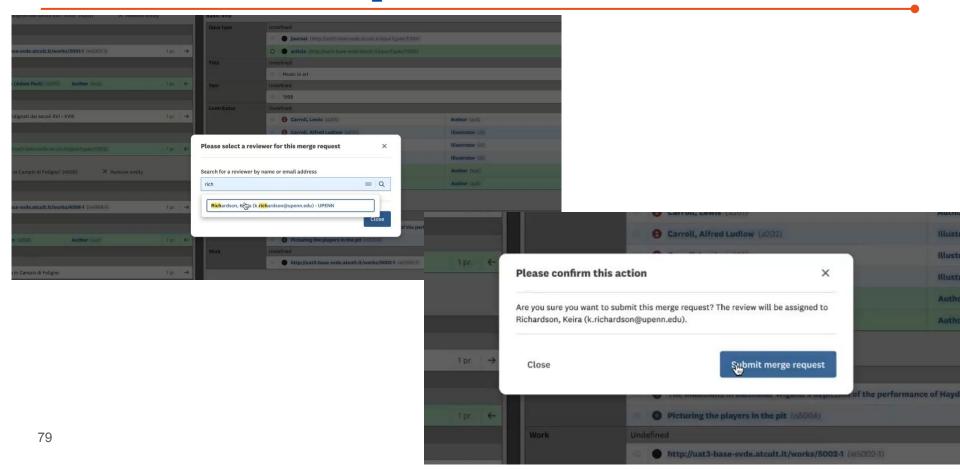


Carroll Adam (Adam Davil) (ather)

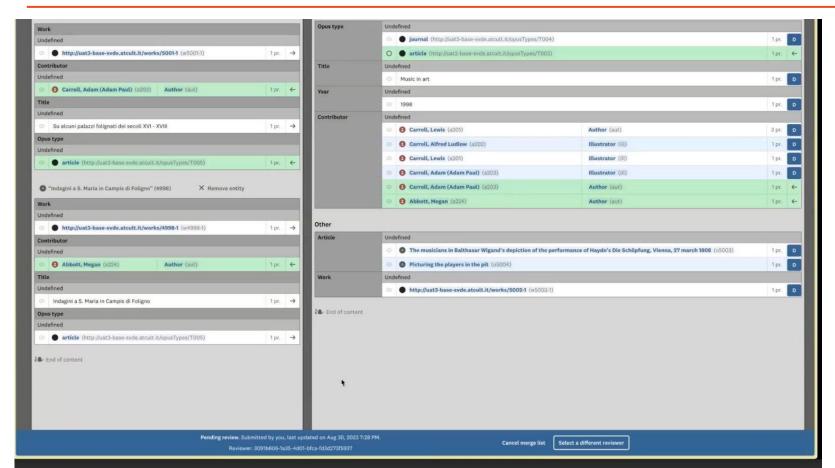
Properties <u>Automatically</u> added (full matching)



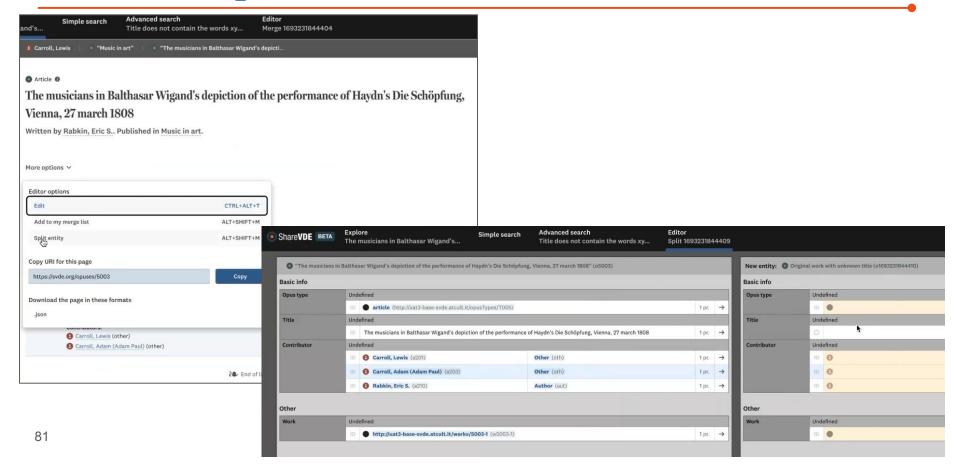
JCricket - review process



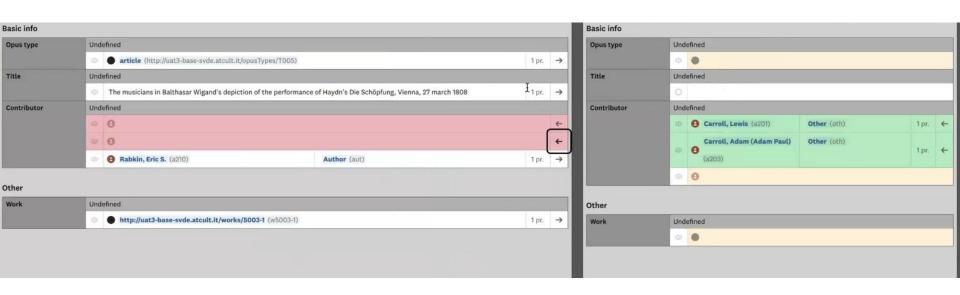
Merge - pending review



JCricket split



JCricket split - select properties





Thank you!

info@svde.org https://svde.org https://wiki.svde.org/ https://www.share-family.org/