JCricket entity editor

Overview of the Share Family tool for a collaborative entity curation

info@svde.org
https://share-family.org
https://wiki.svde.org
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 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)
What it does

- Integrated in the discovery portal web interface, for authenticated users
- User types: basic and advanced
- Entities aggregate data from different contributing libraries (aka Provenances):
  - ★ an entity is also called Prism, where each face represents data coming from a given Provenance
The Big Picture: from Genesis to Edit

Data flows into Share-VDE from libraries, institutions and third-party sources (e.g. VIAF, ISNI, FAST).

The Share-VDE knowledge base (Sapientia) contains the integrated/clustered/enriched entities.

Data is mainly searchable through the entity discovery portal.

Data is mainly edited through JCricket, the Share-VDE entity editor.

Any third-party, authorized application can be a Share-VDE editor or reader.

Third-party Applications

but...

Edit API

Search API
## Properties: Attributes, Relationships, Links

<table>
<thead>
<tr>
<th>Name</th>
<th>Value</th>
<th>Provenance</th>
</tr>
</thead>
<tbody>
<tr>
<td>title</td>
<td>Alice in wonderland</td>
<td>Stanford</td>
</tr>
<tr>
<td>titleAlternative</td>
<td>Alice’s adventures under ground</td>
<td>Stanford</td>
</tr>
<tr>
<td>titleAlternative</td>
<td>Journeys in Wonderland</td>
<td>National Library of Norway</td>
</tr>
</tbody>
</table>

### Attributes

- **sameAs**: http://dbpedia.org/resource/Alice%26%238217%3B_Adventures_in_Wonderland - Dbpedia
- **sameAs**: https://www.wikidata.org/wiki/Q189875 - Wikidata
- **sameAs**: https://data.bnf.fr/ark:/12148/cb356500385#about - bnf

### Relationships

<table>
<thead>
<tr>
<th>Name</th>
<th>Provenance</th>
</tr>
</thead>
<tbody>
<tr>
<td>author</td>
<td>Stanford</td>
</tr>
</tbody>
</table>

- **A relationship is a connection between two Share-VDE Prisms**
- **An attribute is a data property, having a literal as value**
- **A link is a connection between a Share-VDE Prism and an external reference**
Prism: Record-Level Provenance

Each record coming from a provenance contributes in building/enriching one or more Share-VDE prisms.

A Share-VDE entity can be seen as a prism where each face represents data coming from a given provenance.

Each Share-VDE cluster maintains a link to the records it originated from.
JCricket: Edit Scenarios

**Edit:** a property* of a prism is added/updated/deleted

- Lewis Carroll
- is author of
- https://svde.org/opuses/1827349
- https://svde.org/opuses/920302

**Merge:** multiple prisms are merged into one

- For example, the editor detects multiple prisms that belong to the same entity.

**Split:** a prism is split into multiple prisms

- Several entities have been (wrongly) clustered into one.
What it does

**Edit function** to change entities’ properties
★ add, remove and amend attributes, relationships and links belonging to a single entity
★ real time notifications about cluster property changes

**Merge function**: reconcile multiple entities into one (e.g. two authors who are in fact the same person)
★ choose the properties to copy to the merged entity
★ multiple phases: create the merge list, edit the merge list, edit clusters, request for review, approve (or deny) the merge
What it does

**Split and Create functions**: move one or more properties between two entities; optionally create a new one

★ choose the properties to move from entities

★ properties can be moved from an entity to create a brand new one

★ multiple phases: create the split-set, edit the split-set, edit clusters, request for review, approve (or deny) the split
What it does

Review workflow: edits are reviewed by advanced editors
★ notifications to manage the review workflow

Dictionary API: what are the available cluster types? Which attributes belong to a cluster type? What cardinality? Which relationships?

Entity Event Log: tracks the history of changes

Data changes synchronization across SVDE storages (e.g. RDF Store, Search Engine, RDBMS)
JCricket live demo
Why JCricket is valuable

★ A collaborative tool, shared across the Share Family community and improving the data created within it
★ Linked data conversion in a shared discovery environment and direct entity management capabilities in one place
★ Increases data quality where massive automated processes are necessarily lacking
★ Collaboration ensures higher quality and authoritativenss of data
★ Conceived by SVDE and Share Family community, ie. real users
★ Potentially complementary to other tools, e.g. Sinopia - as they cover different aspects of the entity management flow
★ Could potentially support other workflows and connections with systems external to the Share Family
★ You can always track back to your data through the Provenance
★ JCricket will extend authority capabilities through the integration with external data sources such as Wikidata
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 ✍️
  - progressive releases through 2023
  - the first live demo of *Edit* function was presented at ALA - Chicago 2023
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
References

Useful references:

- [https://wiki.share-vde.org/w/images/c/c7/Share_Family_data_flow.pdf](https://wiki.share-vde.org/w/images/c/c7/Share_Family_data_flow.pdf)
- for more technical details on JCricket [https://wiki.share-vde.org/w/images/e/e8/JCricket_entity_editor_presentation.pdf](https://wiki.share-vde.org/w/images/e/e8/JCricket_entity_editor_presentation.pdf)
info@svde.org
https://share-family.org
https://wiki.svde.org