Share VDE Virtual Discovery Environment

SVDE 2.0 Linked Data Management System and Entity Discovery Portal

Progress status of new developments

BIBFRAME Workshop in Europe

21st September 2021

Share-VDE Statement

In September 2021 the Share-VDE Advisory Council has approved and published an official statement that describes the role of the initiative in the broader context of Library Linked Open Data:

Share-VDE’s Role in Library Linked Open Data

Share-VDE 2.0 is launched

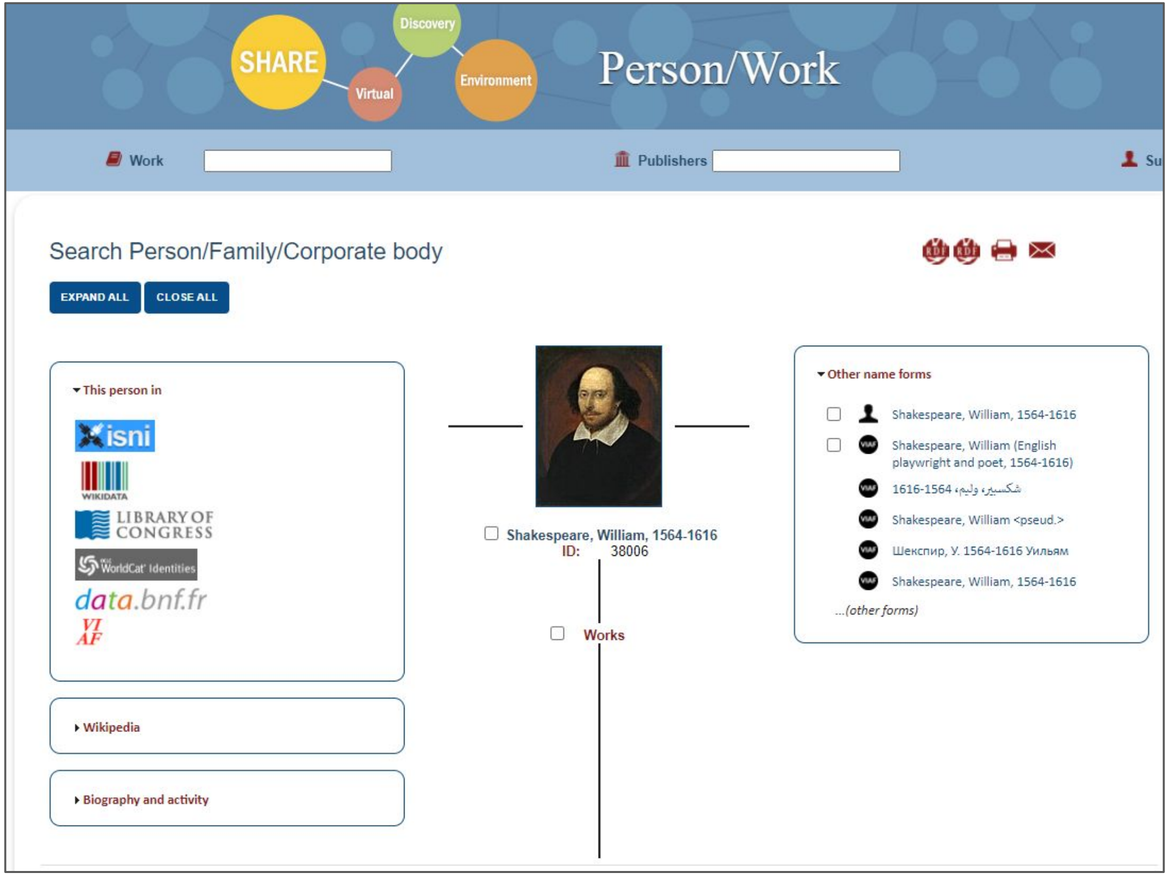
The new SVDE 2.0 is now live at https://svde.org

- new back-end infrastructure for the Linked Data Management and the Cluster Knowledge Base;
- new Entity Discovery Interface (web portal).

This results have been achieved

- with the guidance of the Advisory Council and the active involvement of the Working Groups and parallel project Kubikat-LOD;
- with the support of the SVDE founding members, the full members, and the LD4P project.

What was Share-VDE 1.0



Search Person/Family/Corporate body

EXPAND ALL | CLOSE ALL

▾ This person in

▸ Wikipedia

▸ Biography and activity

□ Shakespeare, William, 1564-1616
ID: 38006

□ Works

▾ Other name forms

- □ Shakespeare, William, 1564-1616
- □ Shakespeare, William (English playwright and poet, 1564-1616)
- 1616-1564 شکسپیر، ولیم،
- Shakespeare, William <pseud.>
- Шекспир, У. 1564-1616 Уильям
- Shakespeare, William, 1564-1616

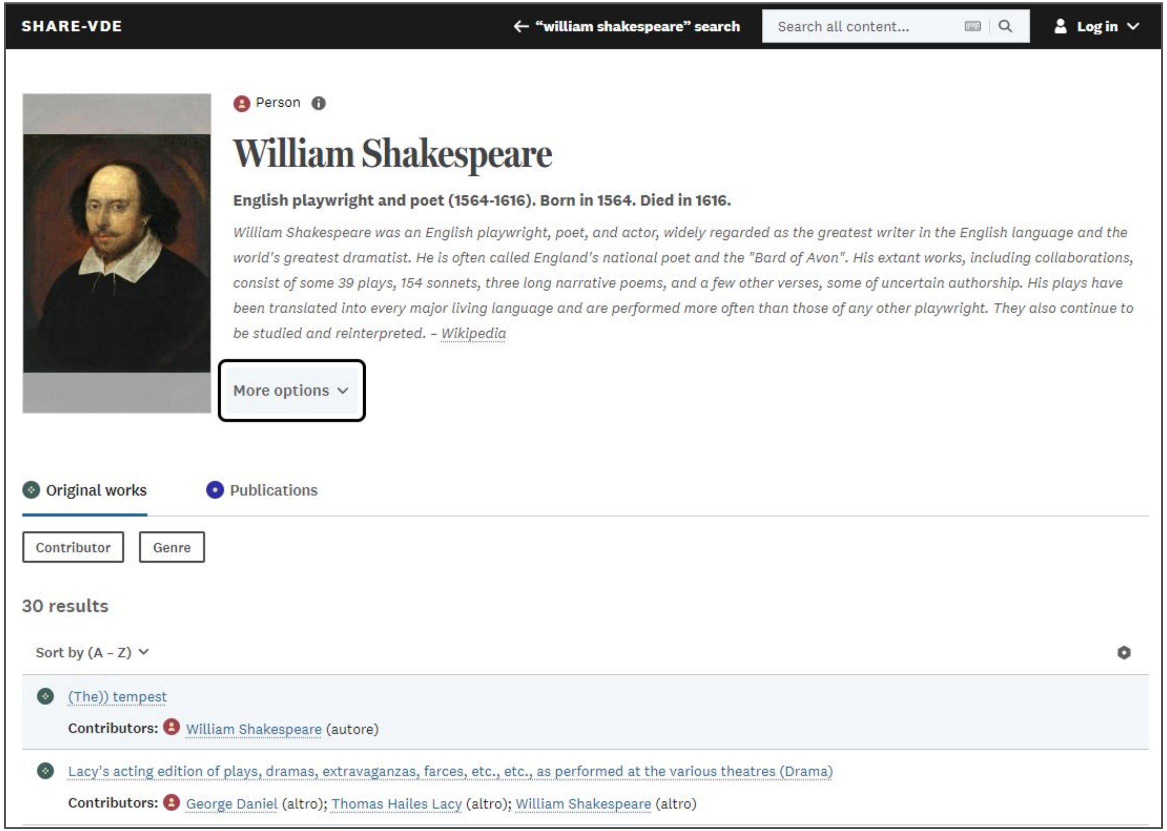
...(other forms)

A complex system designed from the perspective of expert users, where the user experience was quite complex

Back-end and front-end were not differentiated in the technological architecture

The system was not based on APIs

What is Share-VDE 2.0



Person

William Shakespeare

English playwright and poet (1564-1616). Born in 1564. Died in 1616.

William Shakespeare was an English playwright, poet, and actor, widely regarded as the greatest writer in the English language and the world's greatest dramatist. He is often called England's national poet and the "Bard of Avon". His extant works, including collaborations, consist of some 39 plays, 154 sonnets, three long narrative poems, and a few other verses, some of uncertain authorship. His plays have been translated into every major living language and are performed more often than those of any other playwright. They also continue to be studied and reinterpreted. – Wikipedia

30 results

Sort by (A – Z)

(The)) tempest

Contributors: William Shakespeare (autore)

Lacy's acting edition of plays, dramas, extravaganzas, farces, etc., etc., as performed at the various theatres (Drama)

Contributors: George Daniel (altro); Thomas Hailes Lacy (altro); William Shakespeare (altro)

A (much more) complex system with entity-based presentation layer, reflecting BIBFRAME and the ad hoc SVDE extensions

Improved user experience

Back-end infrastructure based on APIs and enhanced with a new version of the [LOD Platform framework]() and of the CKB

What data is available

Progressive load of SVDE libraries data into the new system:

- Share-VDE 2.0 is available at https://svde.org
  - progressive upload of Stanford’s bibliographic records + authority records from the Library of Congress → 14 millions of entity clusters
  - https://www.svde.org/about/about-share-vde
- Share-VDE 1.0 is available at https://share-vde.org

SVDE 1.0 and 2.0 will coexist until clustering iterations and data load on the new version will be completed.

The ultimate goal

The ultimate goal is to:

- create a linked data ecosystem where BIBFRAME entities benefit as much as possible from the wealth of data included in the original MARC catalogues
- act as a linked data node providing authoritative source of data through the CKB
- reconcile data from different libraries in a Union Catalogue and enrich with information from external sources (e.g. addition of URIs to entities from VIAF, ISNI, Wikidata etc.)
- provide a rich but simple user experience on the discovery portal
- expose the data on different layers that can serve many purposes (API layer, triple store, discovery portal)

How we got here

To meet this goal a much more powerful system is needed, and several steps to achieve it:

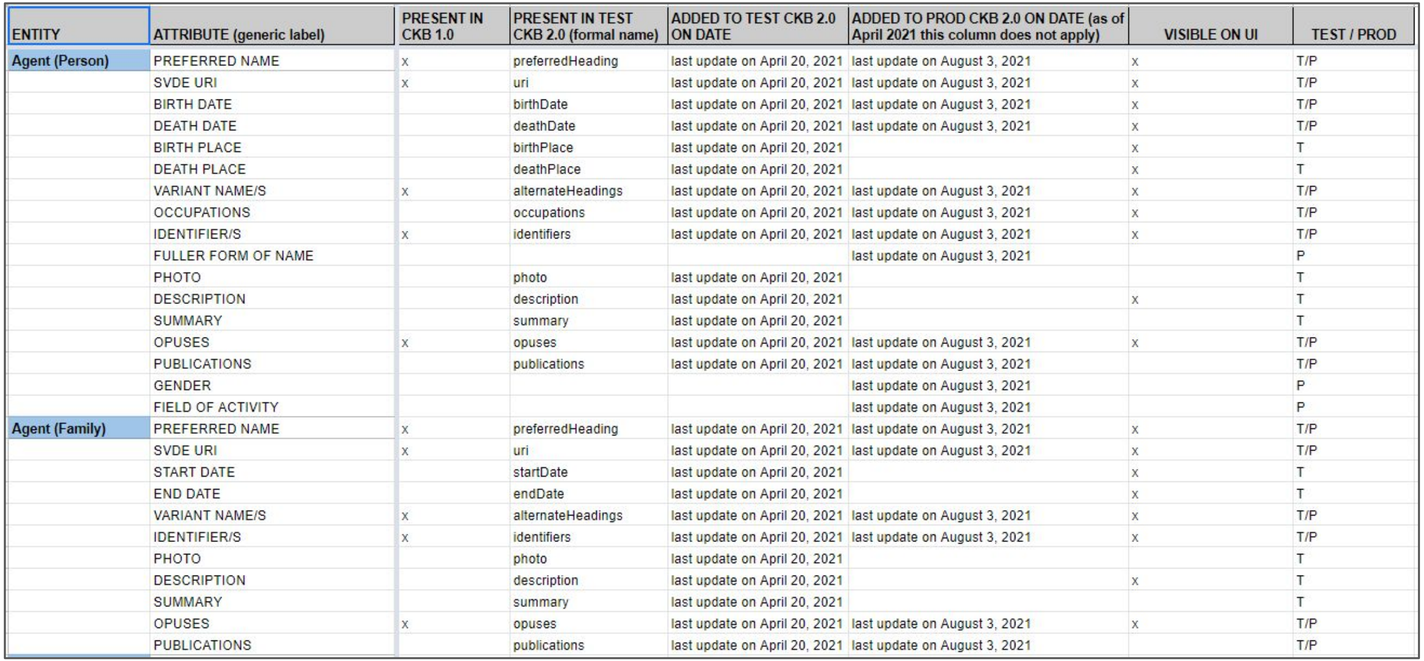
- complex search logic of the new discovery
- refactoring of the Cluster Knowledge Base including the data of many libraries (CKB 2.0)
- updates to the entity model and addition of many new attributes and properties

How we got here

- new API layer with different sets of APIs to support the search logics from the discovery portal and from external systems
- analysis from the SVDE team together with member libraries of the dedicated SEI - Sapientia Entity Identification WG for entity modeling and with members of Kubikat-LOD parallel project
- BIBFRAME extensions to support interoperability with other models, e.g. IFLA LRM (see Share-VDE entity model)
- updates to the entity model, updates to clustering specs, conversion specs in several iterations with the SEI WG
- new Cluster Knowledge Base, which means extended entity modeling and the whole refactoring of several storage mechanisms intertwined: RDBMS, Solr, triple store

Library-driven work

The work of the SVDE team is informed by member libraries. An example of some outcomes of the joint work around entity modeling with the Sapientia Entity Identification WG:



ENTITY	ATTRIBUTE (generic label)	PRESENT IN CKB 1.0	PRESENT IN TEST CKB 2.0 (formal name)	ADDED TO TEST CKB 2.0 ON DATE	ADDED TO PROD CKB 2.0 ON DATE (as of April 2021 this column does not apply)	VISIBLE ON UI	TEST / PROD
Agent (Person)	PREFERRED NAME	x	preferredHeading	last update on April 20, 2021	last update on August 3, 2021	x	T/P
	SVDE URI	x	uri	last update on April 20, 2021	last update on August 3, 2021	x	T/P
	BIRTH DATE		birthDate	last update on April 20, 2021	last update on August 3, 2021	x	T/P
	DEATH DATE		deathDate	last update on April 20, 2021	last update on August 3, 2021	x	T/P
	BIRTH PLACE		birthPlace	last update on April 20, 2021		x	T
	DEATH PLACE		deathPlace	last update on April 20, 2021		x	T
	VARIANT NAME/S	x	alternateHeadings	last update on April 20, 2021	last update on August 3, 2021	x	T/P
	OCCUPATIONS		occupations	last update on April 20, 2021	last update on August 3, 2021	x	T/P
	IDENTIFIER/S	x	identifiers	last update on April 20, 2021	last update on August 3, 2021	x	T/P
	FULLER FORM OF NAME				last update on August 3, 2021		P
	PHOTO		photo	last update on April 20, 2021			T
	DESCRIPTION		description	last update on April 20, 2021		x	T
	SUMMARY		summary	last update on April 20, 2021			T
	OPUSES	x	opuses	last update on April 20, 2021	last update on August 3, 2021	x	T/P
	PUBLICATIONS		publications	last update on April 20, 2021	last update on August 3, 2021		T/P
	GENDER				last update on August 3, 2021		P
	FIELD OF ACTIVITY				last update on August 3, 2021		P
Agent (Family)	PREFERRED NAME	x	preferredHeading	last update on April 20, 2021	last update on August 3, 2021	x	T/P
	SVDE URI	x	uri	last update on April 20, 2021	last update on August 3, 2021	x	T/P
	START DATE		startDate	last update on April 20, 2021		x	T
	END DATE		endDate	last update on April 20, 2021		x	T
	VARIANT NAME/S	x	alternateHeadings	last update on April 20, 2021	last update on August 3, 2021	x	T/P
	IDENTIFIER/S	x	identifiers	last update on April 20, 2021	last update on August 3, 2021	x	T/P
	PHOTO		photo	last update on April 20, 2021			T
	DESCRIPTION		description	last update on April 20, 2021		x	T
	SUMMARY		summary	last update on April 20, 2021			T
	OPUSES	x	opuses	last update on April 20, 2021	last update on August 3, 2021	x	T/P
	PUBLICATIONS		publications	last update on April 20, 2021	last update on August 3, 2021		T/P

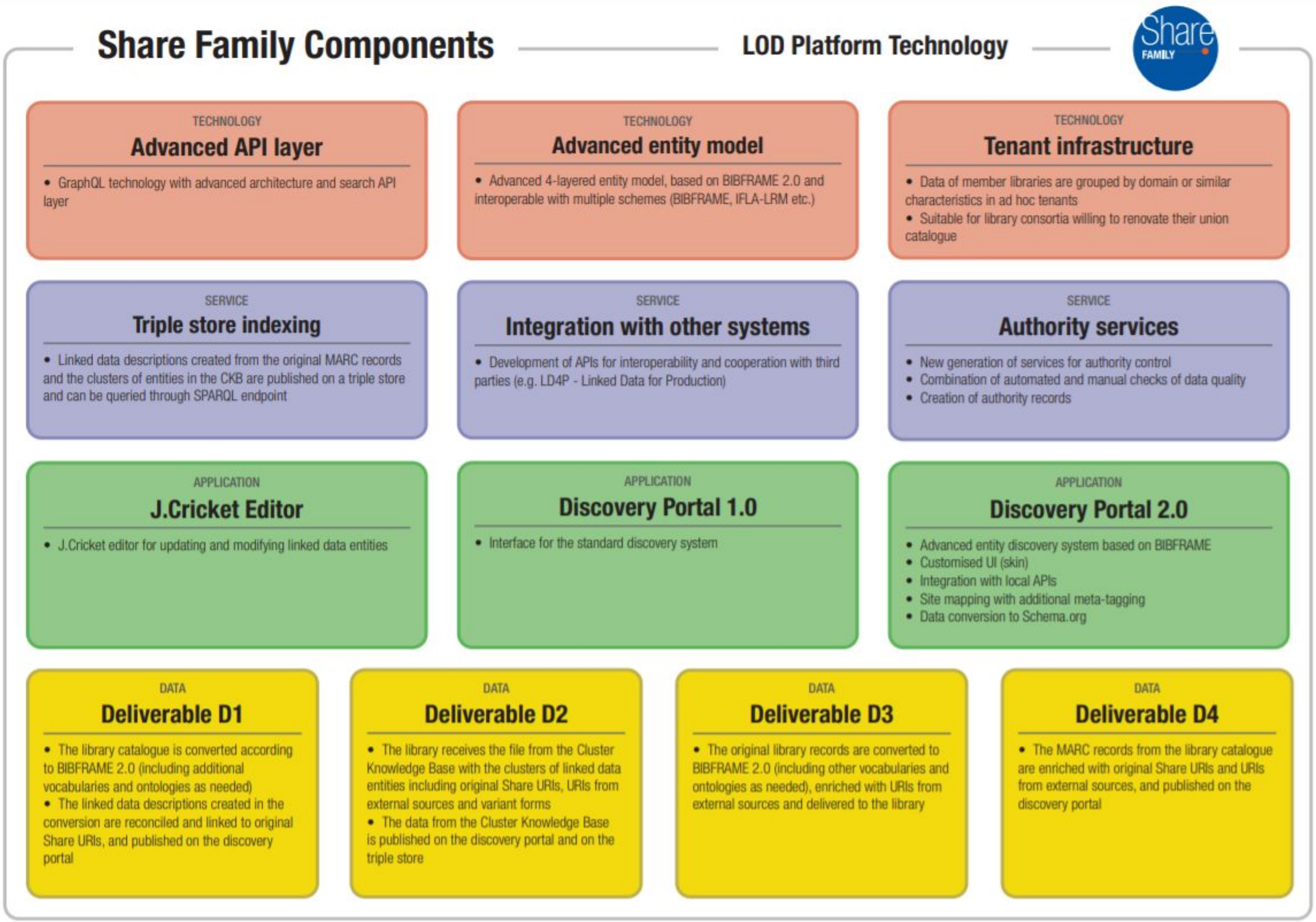
Overview of Share-VDE 2.0: front-end

- Simple search, including: exact match suggestions, explanations of search results; Wikidata descriptions
- Advanced search, including: search for any Agent type, for Original work, for Publication
- Entity pages (Agent, Original work, Publication, first version of Item), including wiki content (images from Wikimedia, summaries from Wikidata, descriptions from Wikipedia)
- Configuration of the system for the connection with local library services via API, for ad hoc customised skin portals
- Optimisation of the system for the J.Cricket editing features that will be developed over the next period
- Optimisation of accessibility features

Overview of Share-VDE 2.0: back-end

- Two API protocols: GraphQL API and REST API
- All Share-VDE entities are exposed through (read-only) API
- Search API provide several shapes / context behaviour (e.g. simple, advanced search, partial or full match, exact matches suggestions, terms modifiers, results explanation)
- Three query languages: TermsQL, SVDEQL, StructQL
- Search Quality Evaluation Tools
- Analysis and design of URI resolution and content negotiation mechanisms: dereference URIs and access to different formats of the entities
- Controlled vocabularies represented as entities (e.g. Roles, Places, Languages, Agent types, Forms, Genre etc.): this allows to dereference such vocabularies using URIs
- Authorization/Authentication infrastructure
- Continuous Integration

LOD Platform components



Share Family Components

LOD Platform Technology

TECHNOLOGY

Advanced API layer

- GraphQL technology with advanced architecture and search API layer

TECHNOLOGY

Advanced entity model

- Advanced 4-layered entity model, based on BIBFRAME 2.0 and interoperable with multiple schemes (BIBFRAME, IFLA-LRM etc.)

TECHNOLOGY

Tenant infrastructure

- Data of member libraries are grouped by domain or similar characteristics in ad hoc tenants
- Suitable for library consortia willing to renovate their union catalogue

SERVICE

Triple store indexing

- Linked data descriptions created from the original MARC records and the clusters of entities in the CKB are published on a triple store and can be queried through SPARQL endpoint

SERVICE

Integration with other systems

- Development of APIs for interoperability and cooperation with third parties (e.g. LD4P - Linked Data for Production)

SERVICE

Authority services

- New generation of services for authority control
- Combination of automated and manual checks of data quality
- Creation of authority records

APPLICATION

J.Cricket Editor

- J.Cricket editor for updating and modifying linked data entities

APPLICATION

Discovery Portal 1.0

- Interface for the standard discovery system

APPLICATION

Discovery Portal 2.0

- 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

DATA

Deliverable D1

- The library catalogue is converted according to BIBFRAME 2.0 (including additional vocabularies and ontologies as needed)
- The linked data descriptions created in the conversion are reconciled and linked to original Share URIs, and published on the discovery portal

DATA

Deliverable D2

- The library receives the file from the Cluster Knowledge Base with the clusters of linked data entities including original Share URIs, URIs from external sources and variant forms
- The data from the Cluster Knowledge Base is published on the discovery portal and on the triple store

DATA

Deliverable D3

- The original library records are converted to BIBFRAME 2.0 (including other vocabularies and ontologies as needed), enriched with URIs from external sources and delivered to the library

DATA

Deliverable D4

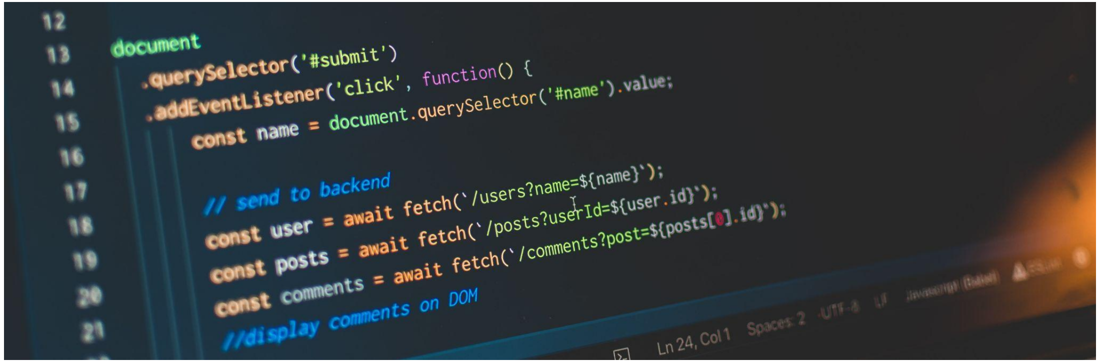
- The MARC records from the library catalogue are enriched with original Share URIs and URIs from external sources, and published on the discovery portal

How to query Share-VDE and provide feedback

Share-VDE data can be queried through several methods:

- entity discovery portal (web user interface available at https://svde.org)
  - https://www.svde.org/about/about-share-vde
- via API through GraphQL and RESTful API endpoints
- via Stardog triple store (the Stardog db including the new CKB 2.0 will soon be available)

Report bugs and suggestions on the forum https://forum.svde.org/



```
document
  .querySelector('#submit')
  .addEventListener('click', function() {
    const name = document.querySelector('#name').value;

    // send to backend
    const user = await fetch(`/users?name=${name}`);
    const posts = await fetch(`/posts?userId=${user.id}`);
    const comments = await fetch(`/comments?post=${posts[0].id}`);
    //display comments on DOM
```

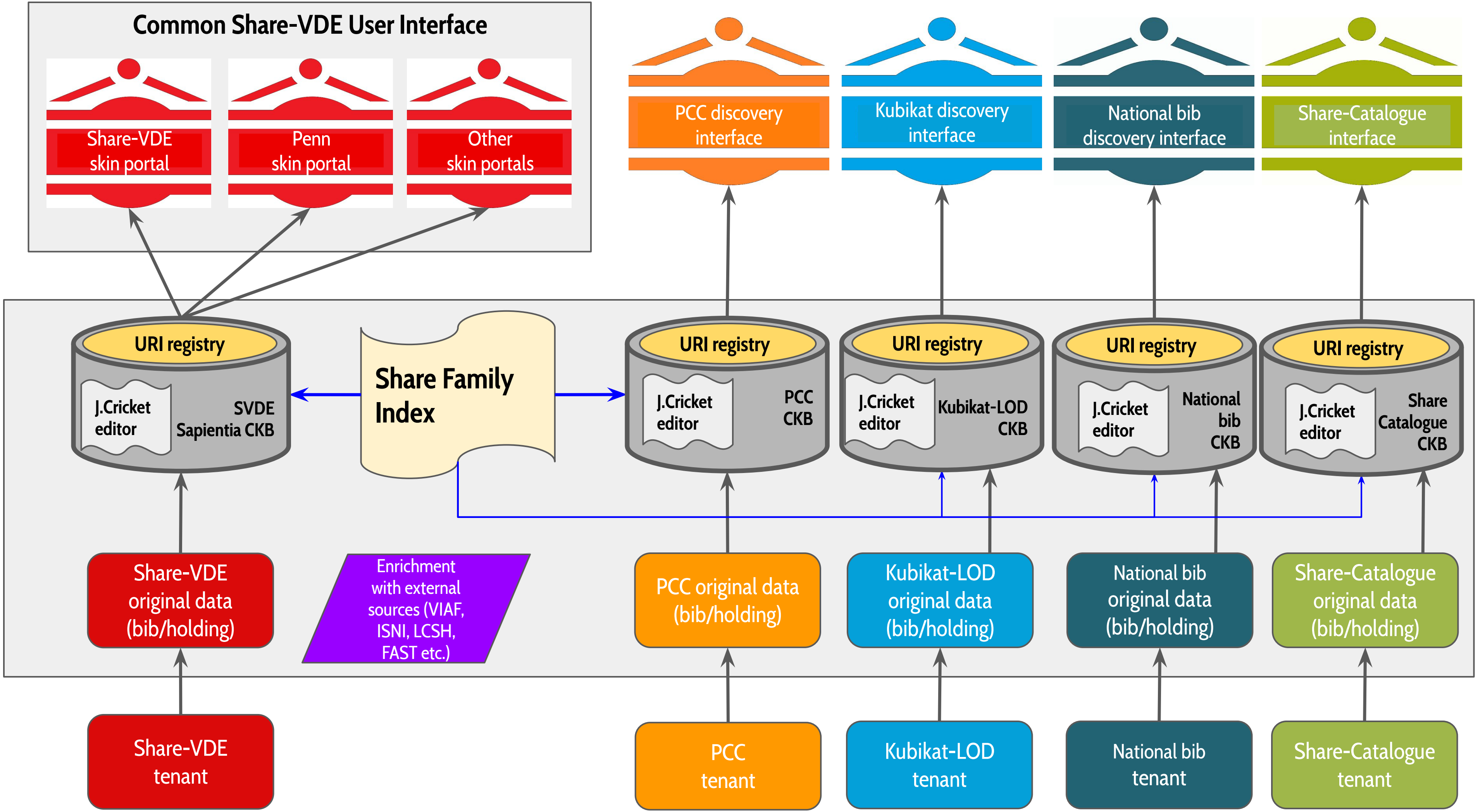
Next steps

What comes next:

- progressive load of SVDE members’ catalogues to populate the front-end portal
- complete connection with ad hoc skin portals
- continue the work on ad hoc features for Kubikat-LOD parallel project
- a huge work will be done to manage the Instances and related descriptions from the different libraries
- further enrichment of the CKB with new properties and refinements according to the joint work with the SEI WG
- developments of J.Cricket CKB editor
- Tenant architecture with Share Family Index (SFI) implementation

Further activities of the Share family

Share VDE


Common Share-VDE User Interface
Share-VDE skin portal
Penn skin portal
Other skin portals
PCC discovery interface
Kubikat discovery interface
National bib discovery interface
Share-Catalogue interface
URI registry
J.Cricket editor
SVDE Sapientia CKB
Share Family Index
URI registry
J.Cricket editor
PCC CKB
URI registry
J.Cricket editor
Kubikat-LOD CKB
URI registry
J.Cricket editor
National bib CKB
URI registry
J.Cricket editor
Share Catalogue CKB
Share-VDE original data (bib/holding)
Enrichment with external sources (VIAF, ISNI, LCSH, FAST etc.)
PCC original data (bib/holding)
Kubikat-LOD original data (bib/holding)
National bib original data (bib/holding)
Share-Catalogue original data (bib/holding)
Share-VDE tenant
PCC tenant
Kubikat-LOD tenant
National bib tenant
Share-Catalogue tenant

Further activities

- Kubikat-LOD: work is going on in parallel with the new SVDE infrastructure; important components emerged in the Kubikat group that serve the whole infrastructure are being tackled (e.g. serials) and work is progressing for the specific Kubikat tenant
- Parsifal, union catalogue of ecclesiastic libraries in Rome: work is going on in parallel to go live with the first version of the system
- Authority services: towards completion of the MARC-based services (tests ongoing at Stanford and among SVDE libraries); next steps: authority control based on linked data
- Continuous dialogue within the community and with other initiatives such as LD4P3, the PCC and the institutions involved

Further activities

- New working group dedicated to the practical cooperation among the National Bibliographies, to address the needs of National Libraries and institutions that hold National Bibliographies in the framework of a shared entity discovery environment such as the Share Family of initiatives
  - this could be a new tenant of the Share family

- Resources about the Share family
  - Share family presentation https://www.casalini.it/linked-data-for-libraries/
  - Share family resources https://wiki.share-vde.org/wiki/ShareFamily:Main_Page

Share VDE Virtual Discovery Environment

Thank you!

tiziana.possemato@atcult.it
tiziana.possemato@casalini.it
anna.lionetti@casalini.it

https://wiki.svde.org/
https://svde.org
info@share-vde.org

Share VDE

Virtual Discovery Environment

Share VDE - Backend

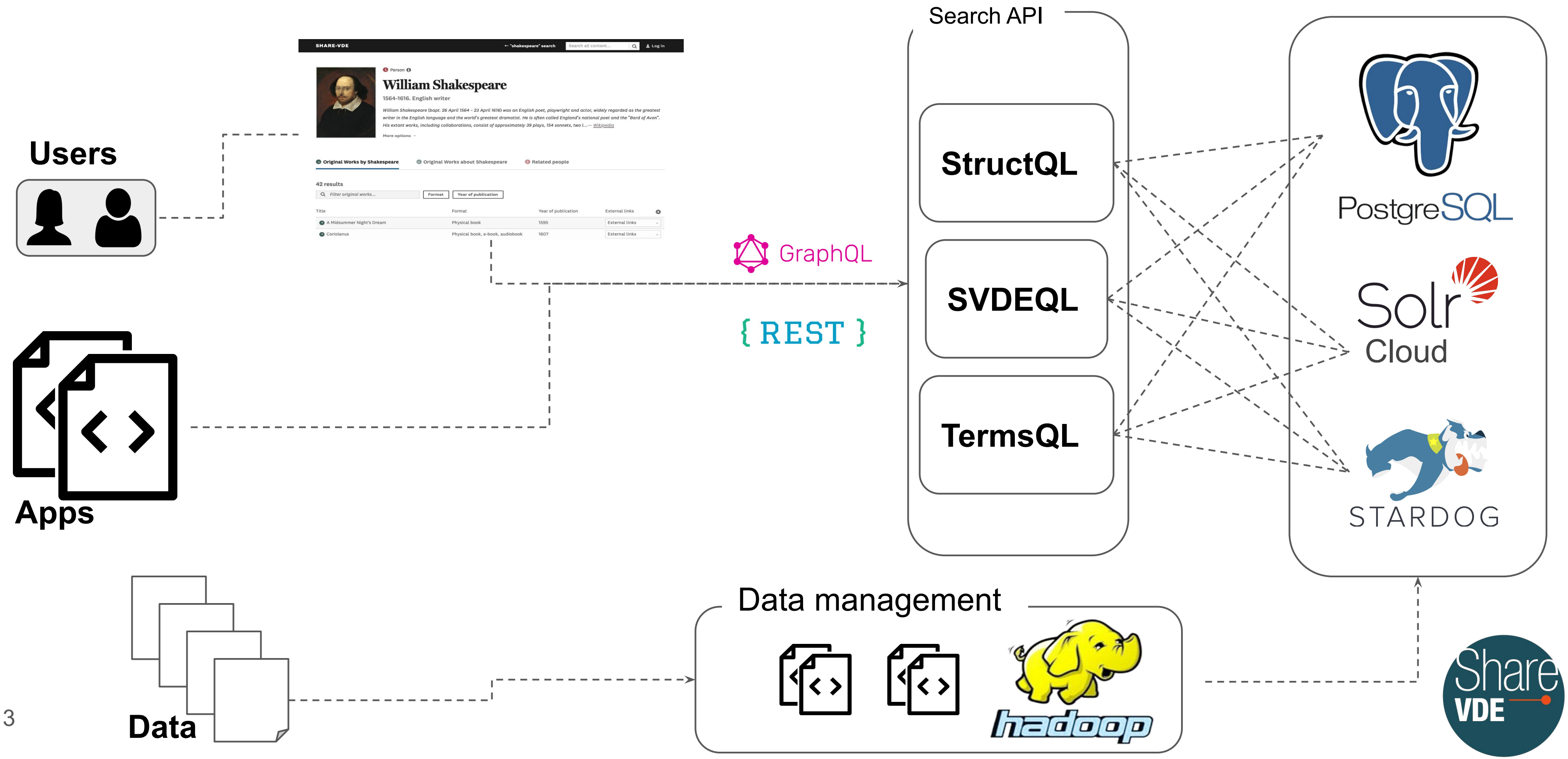
Andrea Gazzarini

www.share-vde.org
info@share-vde.org

Overview

Share VDE

Search API: Overview



SHARE-VDE ← "shakespeare" search Search all content... Log in

Person

William Shakespeare

1564-1616. English writer

William Shakespeare (bapt. 26 April 1564 – 23 April 1616) was an English poet, playwright and actor, widely regarded as the greatest writer in the English language and the world's greatest dramatist. He is often called England's national poet and the "Bard of Avon". His extant works, including collaborations, consist of approximately 39 plays, 154 sonnets, two l...— Wikipedia

More options

Original Works by Shakespeare | Original Works about Shakespeare | Related people

42 results

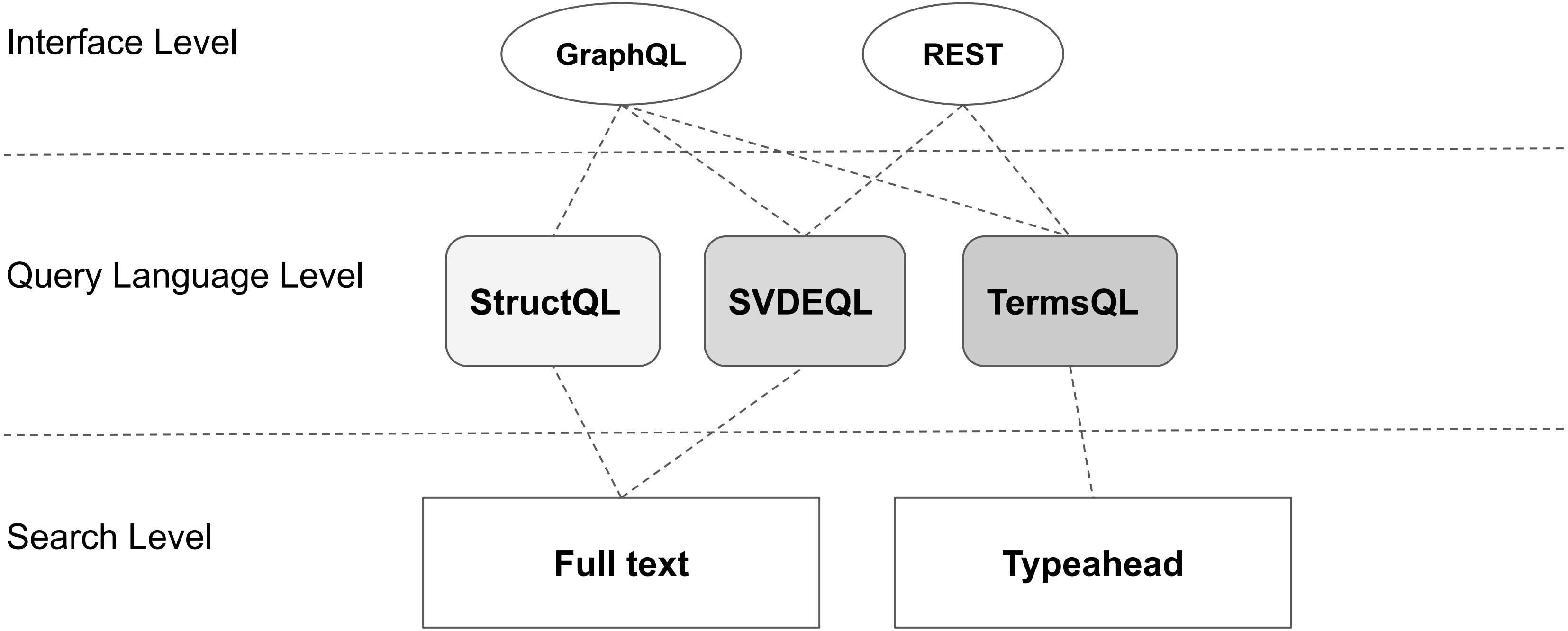
Filter original works... | Format | Year of publication

Title	Format	Year of publication	External links
A Midsummer Night's Dream	Physical book	1595	External links
Coriolanus	Physical book, e-book, audiobook	1607	External links

How can we interact with API?

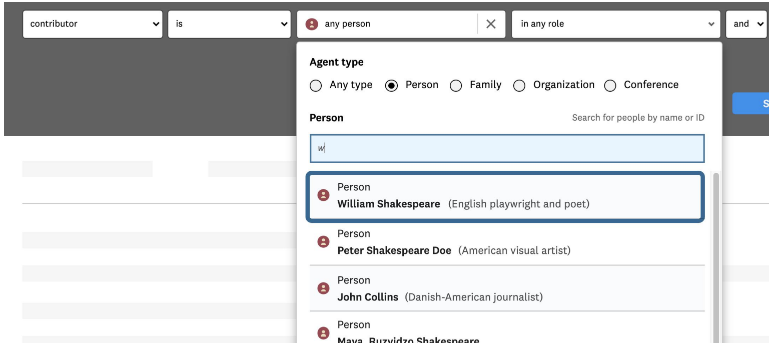
Share VDE

Query Languages



TermsQL

- Available in **GraphQL** and **REST** endpoints
- It **isn't a query language** itself
- The **query string** is composed only by **terms**
- When used, it triggers a **typeahead** search (see here for a detailed functional description)
- The response contains
  - matching entities with highlighting snippets
  - (optionally*) a "correctlySpelled" flag which indicates if some correction has been applied to the original terms
  - (optionally*) matching entities in other languages



contributor | is | any person | in any role | and

Agent type

○ Any type ● Person ○ Family ○ Organization ○ Conference

Person — Search for people by name or ID

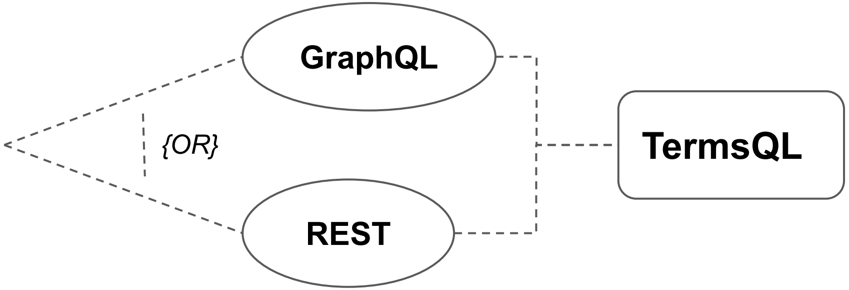
w

Person
William Shakespeare (English playwright and poet)

Person
Peter Shakespeare Doe (American visual artist)

Person
John Collins (Danish-American journalist)

Person
Maya Ruzvidzo Shakespeare



GraphQL

{OR}

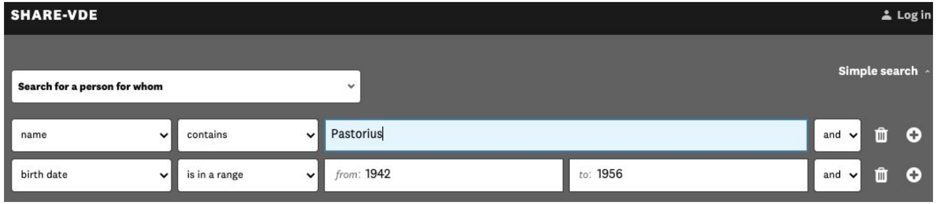
REST

TermsQL

** Only if the feature has been enabled and in case of zero results (using original terms + requestor language)*

SVDEQL / StructQL

- **Simple**: we don't want to build another query language
- **Dedicated grammar** for expressing and validating the "**advanced search**" syntax
- **Easy** and **human-readable** way to express an **information need**
- **Communication protocol** between **frontend** and **backend**
- Could be potentially **exposed** to **advanced users** in the future
- Only a "**format**" **difference** between the two query languages



SHARE-VDE

Log in

Simple search

Search for a person for whom

name | contains | Pastorius | and

birth date | is in a range | from: 1942 | to: 1956 | and

=

people whose name contains Pastorius
AND birth date is in range from 1942 to 1956

StructQL vs SVDEQL

StructQL

- Only available in the GraphQL endpoint
- JSON-like structure
- Grammar enforced by GraphQL schema

Example

```
{
   q: [
         { name : {p: CONTAINS, o: "Carroll"}, op:AND },
         { name : {p: CONTAINS, o: "Lewis"}}
   ]
}
```

SVDEQL

- Available in the GraphQL and REST endpoints
- Pseudo-natural syntax
- Grammar enforced by JAVACC

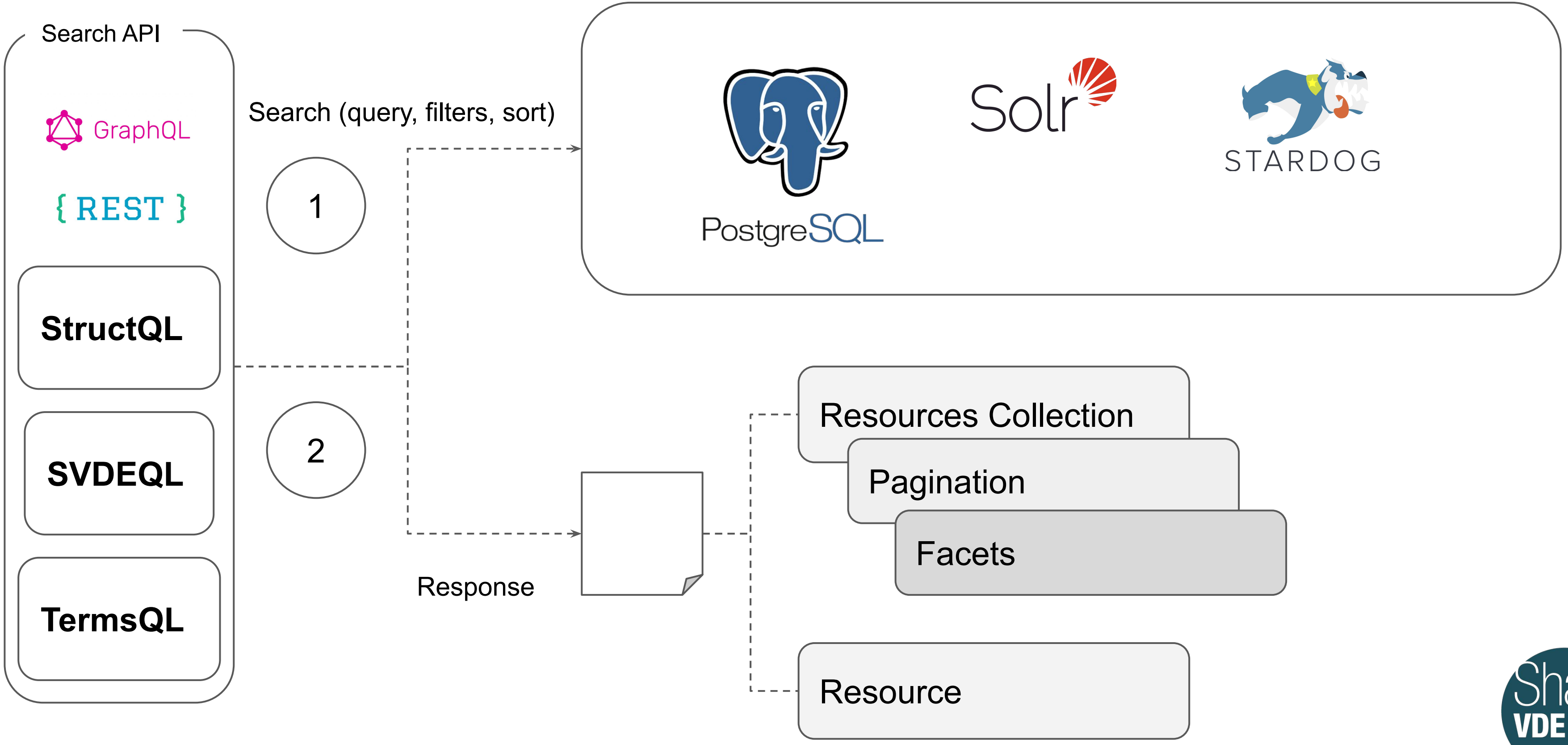
Example

agents whose
        name contains *Carroll* **AND**
        name contains *Lewis*

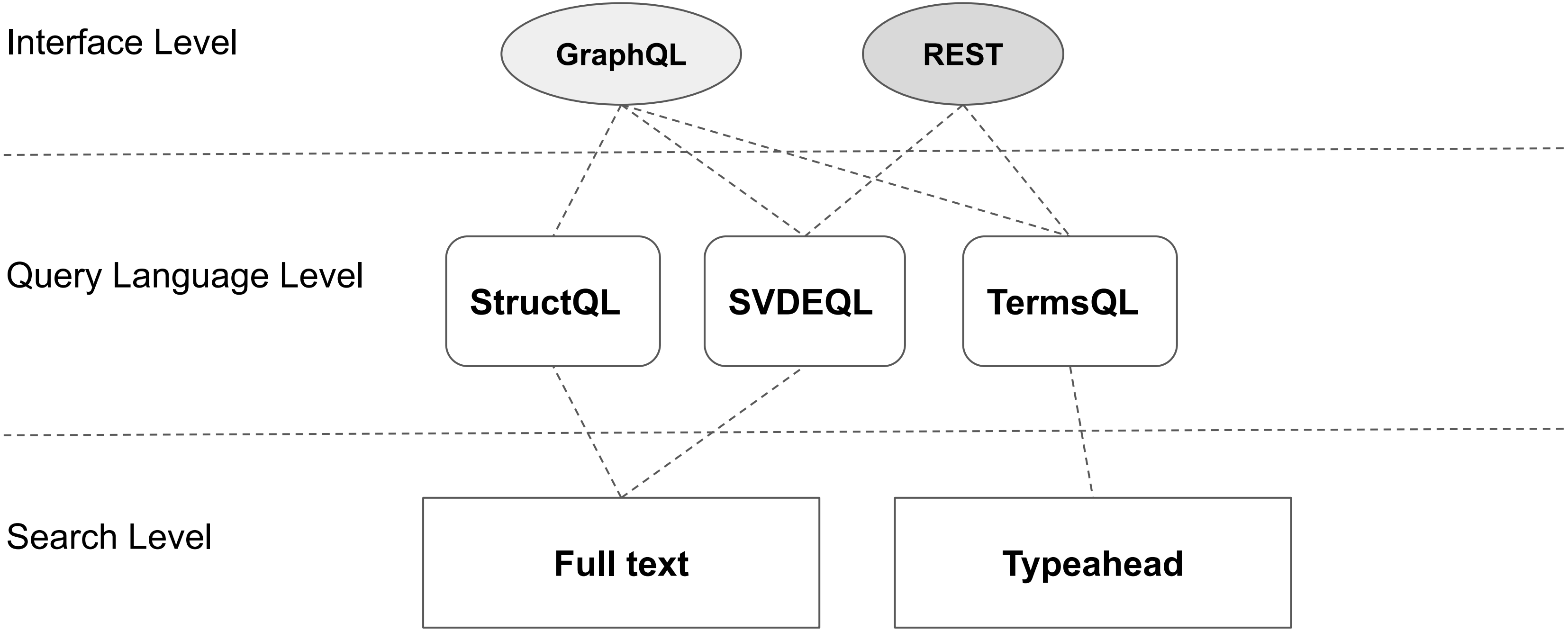
What do we get in output?

Share VDE

Hybrid Response: Realtime and Near Realtime



Query Languages



RESTFul API

- **H**ypermedia **A**s **T**he **E**ngine **O**f **A**pplication **S**tate (**HATEOAS**)
- Share VDE resources provide information dynamically through **hypermedia**
- client actions are **discovered** within **resource representations** returned from the server
- **linked resources are expressed through URIs**, in dedicated sections of the resource representation
- https://en.wikipedia.org/wiki/HATEOAS

RESTFul API: Resource (Person)

```
{
 "heading": "Carroll, Lewis",
 ...
 "_links": {
   "self": [
     { "href": "https://share-vde.org/people/201" },
     { "href": "https://share-vde.org/agents/201" },
     { "href": "http://isni.org/isni/000000012137136X", "type": "ISNI" },
     { "href": "https://viaf.org/viaf/66462036", "type": "VIAF" }
   ],
   "photo": { "href":  "https://commons.wikimedia.org/wiki/lc_1863.jpg" },
   "birth_place": { "href": "https://share-vde.org/places/7295222" }
 },
 "alternate_headings": [ "Dodgson, Charles Lutwidge", "Karol, Luis" ],
 "birth_date": 1832,
 "death_date": 1898
}
```

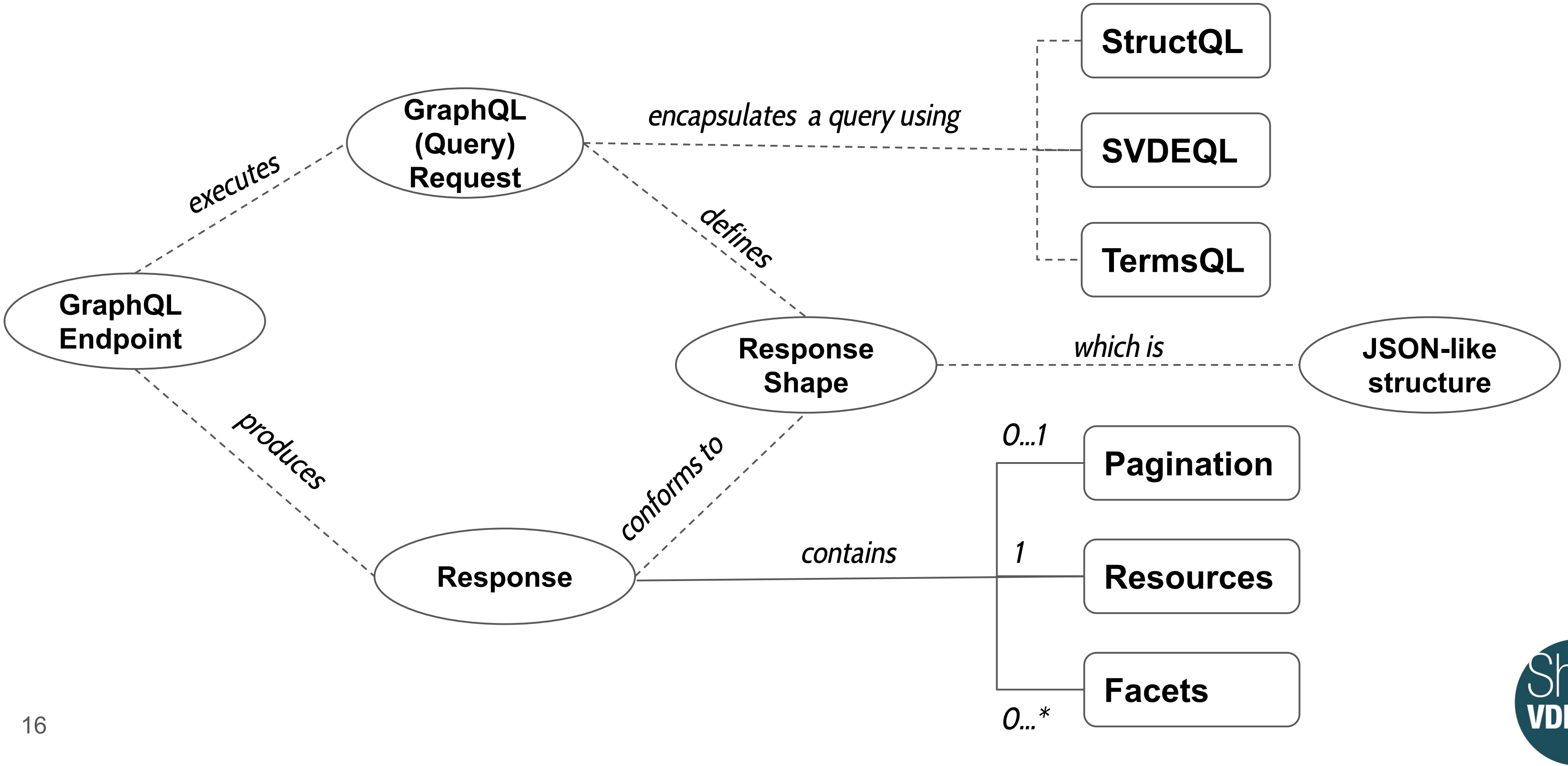
RESTFul API: Collection (People)

```
{
 "_embedded": {
   "agentList": [
     {
       "heading": "Carroll, Lewis",
       "_links": {
            ...
     },
     {
       "heading": "Dodgson, Campbell",
       "alternate_headings": [ "Dodgson, C." ],
       "birth_date": 1867,
       "death_date": 1948
     },
     ...
```

HATEOAS: Faceted Collection (Agents)

```
...
"facets": {
  "facet_fields": {
    "location": {
      { "https://share-vde.org/places//5387877": 3,
        "https://share-vde.org/places/2650225": 2,
      ...
    },
    "type": {
      "https://share-vde.org/agentTypes/Meeting": 6,
      "https://share-vde.org/agentTypes/Person": 4,
      "https://share-vde.org/agentTypes/Family": 3,
      "https://share-vde.org/agentTypes/Organisation": 1
    }
```

GraphQL


GraphQL Endpoint
executes
GraphQL (Query) Request
encapsulates a query using
StructQL
SVDEQL
TermsQL
defines
Response Shape
which is
JSON-like structure
produces
conforms to
Response
contains
0...1
Pagination
1
Resources
0...*
Facets

GraphQL: Example Request

```
{
 people(svdeql: {q: "people whose name contains meyer AND birthDate is in range from 1950 to 1988 "}) {
  resources {
   uri
   preferredHeading
   alternateHeadings
   birthDate
   opuses {
    resources {
     uri
     preferredHeading
     works {
      resources {
       uri
       language {
        uri
        preferredHeading
       }
      }
     }
    }
   }
  }
 }
}
```

Main query

For each person return these literal attributes

For each person return the opuses he contributed

For each opus return these literal attributes

For each opus return the child works

For each work return the language, specifically the URI and the heading

GraphQL: Example Response (extract)

```
{
 people(svdeql: {q: "people whose name contains meyer AND birthDate is in range from 1950 to 1988 "}) {
  resources {
   uri
   preferredHeading
   alternateHeadings
   birthDate
   opuses {
    resources {
     uri
     preferredHeading
     works {
      resources {
       uri
       language {
        uri
        preferredHeading
       }
      }
     }
    }
   }
  }
 }
}
```

```
{
 "uri": "https://share-vde.org/agents/211631289194108",
 "preferredHeading": "Meyer, Han",
 "alternateHeadings": [
  "Meyer, Valentin Johannes",
  "Meyer, Han"
 ],
 "opuses": {
  "resources": [
   {
    "uri": "https://share-vde.org/opuses/181631290147810",
    "preferredHeading": "Stad en de haven",
    "works": {
     "resources": [
      {
       "uri": "https://share-vde.org/works/181631290147812",
       "language": {
        "uri": "https://share-vde.org/languages/21631288253353",
        "preferredHeading": "Inglese"
       }
      }
     ]
    }
   },
   {
    "uri": "https://share-vde.org/opuses/961631362475595",
    "preferredHeading": "Delta urbanism",
    "works": {
   …
```

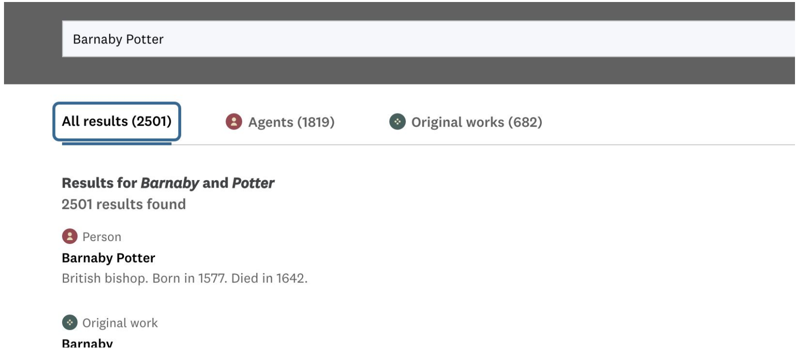
(other people follow)

Search Features

Share VDE

FullText Search

Simple (terms/phrases) Search



Barnaby Potter

All results (2501) | Agents (1819) | Original works (682)

Results for *Barnaby* and *Potter*

2501 results found

Person

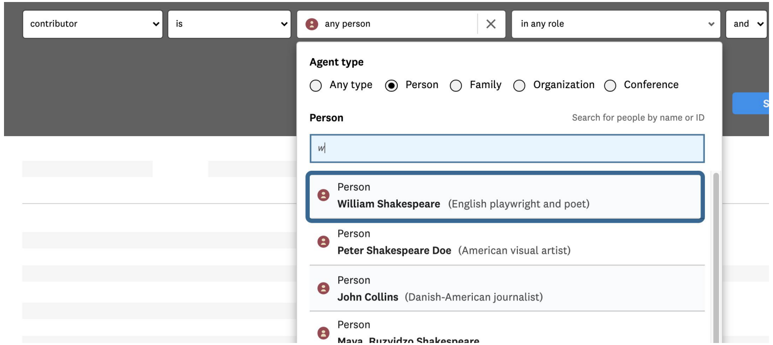
Barnaby Potter

British bishop. Born in 1577. Died in 1642.

Original work

Barnaby

Typeahead Search



contributor | is | any person | in any role | and

Agent type

○ Any type ● Person ○ Family ○ Organization ○ Conference

Person Search for people by name or ID

w

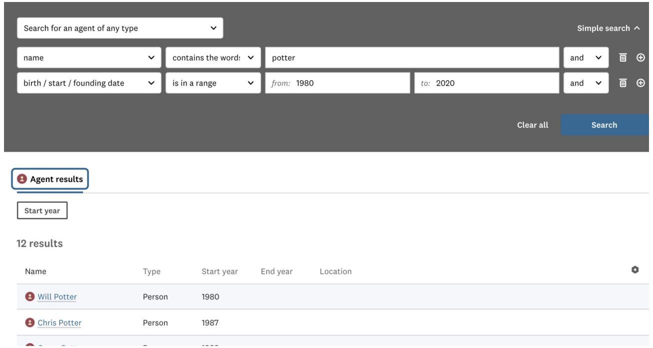
Person
William Shakespeare (English playwright and poet)

Person
Peter Shakespeare Doe (American visual artist)

Person
John Collins (Danish-American journalist)

Person
Maya Ruzvidzo Shakespeare

Advanced Search



Search for an agent of any type

name | contains the word: | potter | and

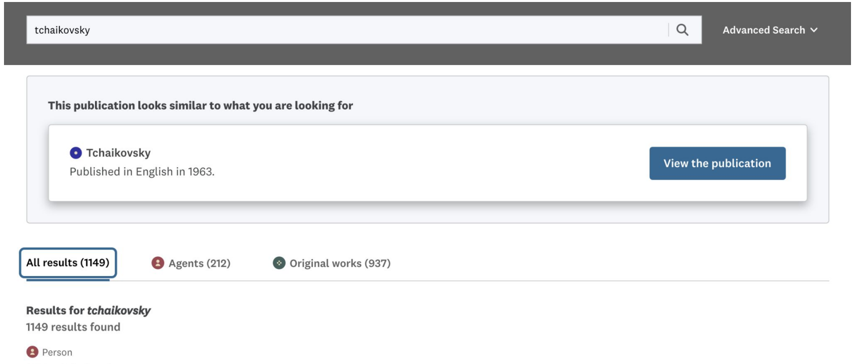
birth / start / founding date | is in a range | *from:* 1980 | *to:* 2020 | and

12 results

Name	Type	Start year	End year	Location
Will Potter	Person	1980		
Chris Potter	Person	1987		

Exact Match Suggestions

A search which targets only entities whose headings or identifiers exactly match the query, partially or totally.



tchaikovsky

Advanced Search

This publication looks similar to what you are looking for

Tchaikovsky
Published in English in 1963.

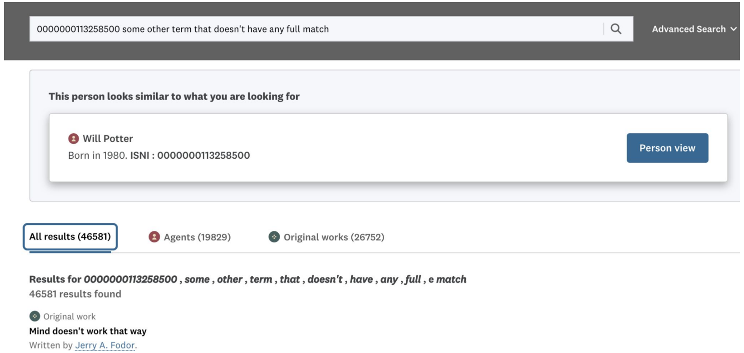
View the publication

All results (1149) | Agents (212) | Original works (937)

Results for *tchaikovsky*
1149 results found

Person

```
{
  exactMatches(q: "tchaikovsky") {
    resources {
      ... (other )
    }
  }
}
```



0000000113258500 some other term that doesn't have any full match

Advanced Search

This person looks similar to what you are looking for

Will Potter
Born in 1980. **ISNI : 0000000113258500**

Person view

All results (46581) | Agents (19829) | Original works (26752)

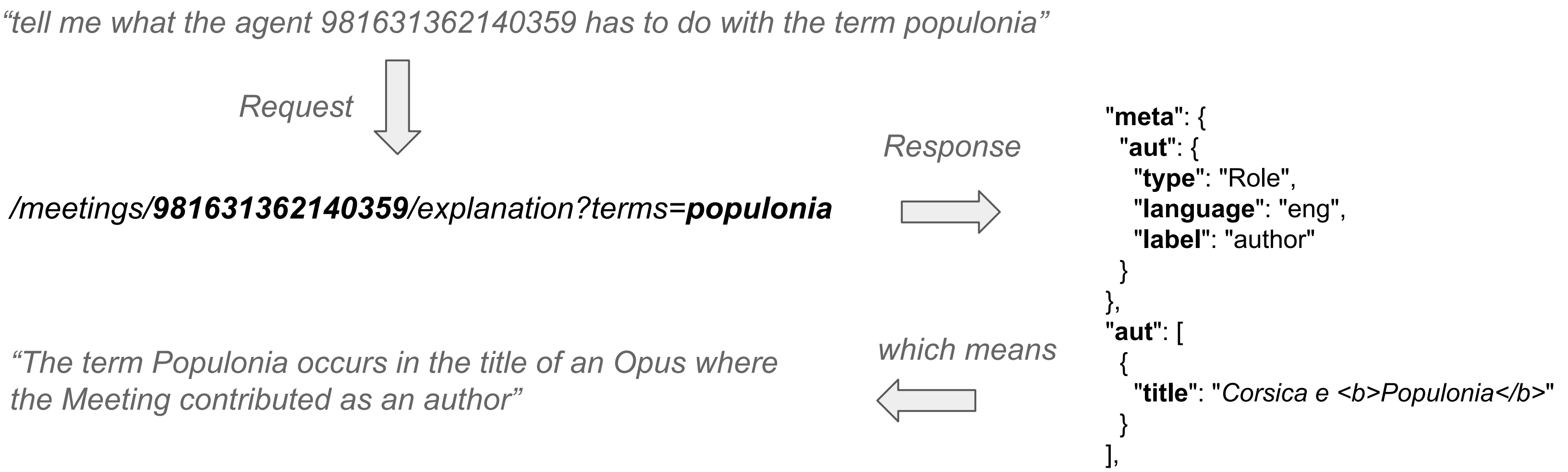
Results for *0000000113258500 , some , other , term , that , doesn't , have , any , full , e match*
46581 results found

Original work
Mind doesn't work that way
Written by Jerry A. Fodor.

```
{
  exactMatches(q: "0000000113258500 ...") {
    resources {
      ... (other )
    }
  }
}
```

Query Explanation

A virtual explanation entity, associated to a core entity (e.g. Agent, Opus) which provides insights about the reason why a given resource has been included in search results.



“tell me what the agent 981631362140359 has to do with the term populonia”

Request

*/meetings/**981631362140359**/explanation?terms=**populonia***

Response

```
"meta": {
  "aut": {
    "type": "Role",
    "language": "eng",
    "label": "author"
  }
},
"aut": [
  {
    "title": "Corsica e <b>Populonia</b>"
  }
],
```

which means

“The term Populonia occurs in the title of an Opus where the Meeting contributed as an author”

Share VDE Virtual Discovery Environment

Share VDE - Backend

Thank you!

Andrea Gazzarini

www.share-vde.org
info@share-vde.org