

# Penn Libraries Hybrid Linked Data Updates

---

Jim Hahn

April 9, 2024

LD4 Discovery Affinity  
Group





## Penn Libraries Linked Data Framework: Interoperability Focus

[Introduction](#)[Near Term Framework for Linked Data](#)[Supporting Penn Libraries Strategic Objectives](#)[Interoperability Focus](#)[How cross-organizational collaborations can move linked data forward](#)[Conclusion](#)[Appendix: Linked Data and Other Formats](#)[PDF Download](#)

### 4. Interoperability Focus

The implementations enumerated in the previous section are not meant to replace existing systems, but rather extend functionality by building on MARC data with linked data to enhance discovery. We can implement these services in a modular fashion, supported by independent and de-coupled APIs and standard web technologies. The modular implementation of a BibCard API aims for system stability by utilizing loosely coupled data sources implemented through web-based APIs.

# Hybrid Linked Data Projects

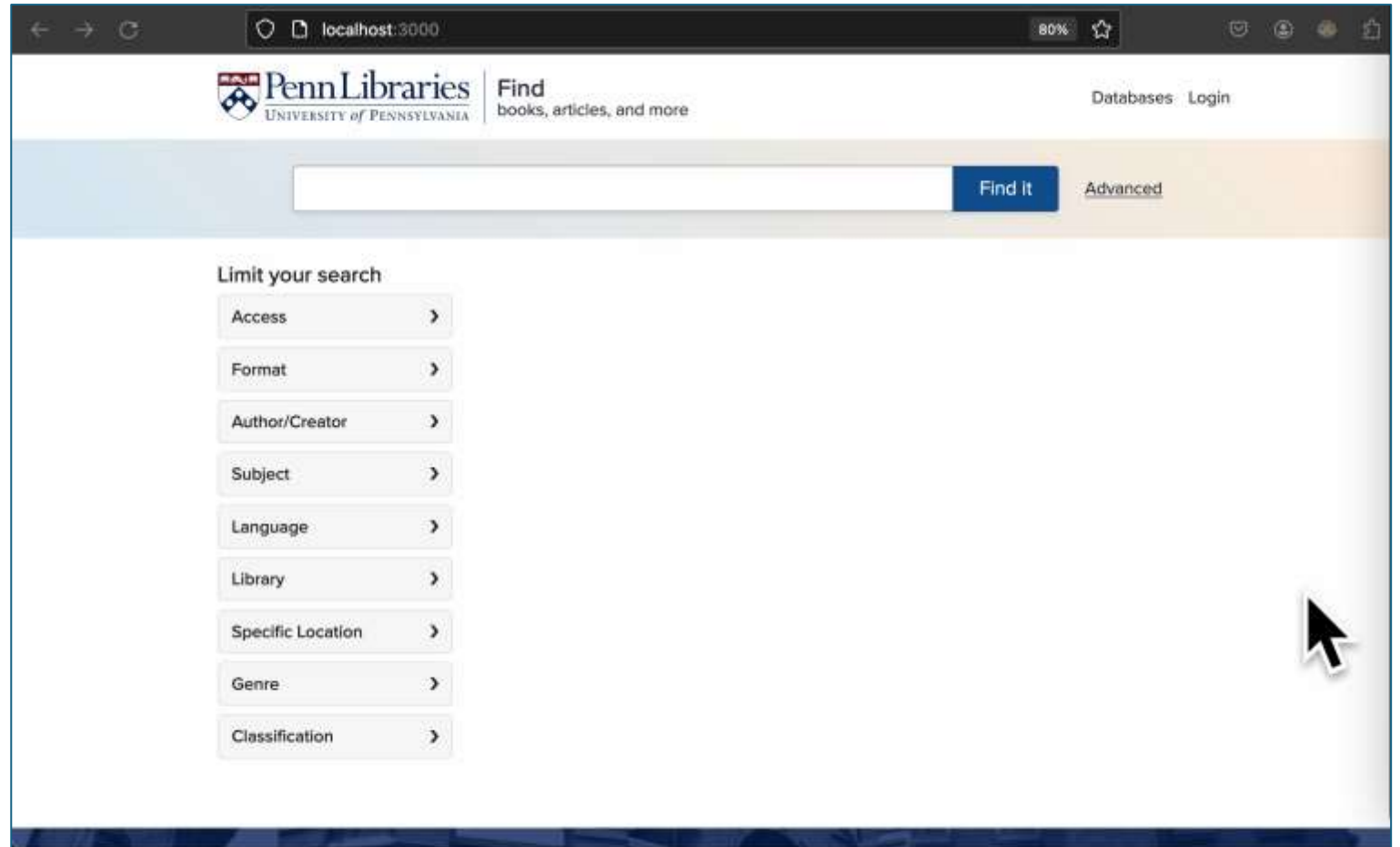
---

- Blacklight 8
  - New catalog at Penn
    - Include Knowledge Panels / other Linked Data sources
- Share-VDE
  - Integrating SVDE Linked Data into New Catalog



# Proof of Concept (Demo)

---



# Several Modules

Modified SQID for  
Display of Linked  
Data in Catalog

API Microservices

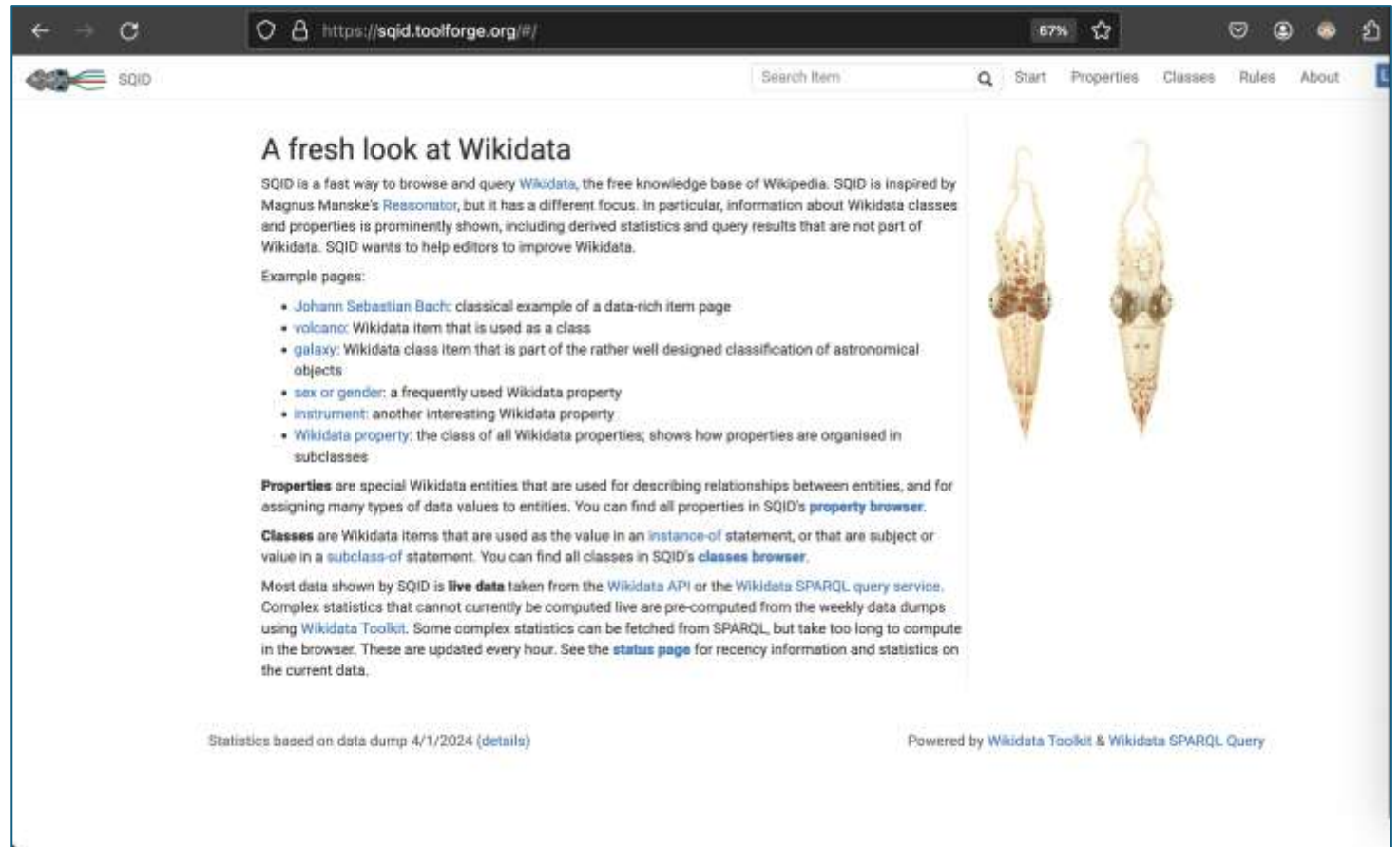
LD-Microservices

SVDE GraphQL  
middleware

# SQID

---

- <https://sqid.toolforge.org/#/>



The screenshot shows the SQID website homepage in a browser. The address bar displays <https://sqid.toolforge.org/#/>. The page features a navigation menu with links for Search Item, Start, Properties, Classes, Rules, and About. The main content area is titled "A fresh look at Wikidata" and includes an introductory paragraph, a list of example pages, and sections for Properties and Classes. A decorative image of two ice cream cones is on the right. The footer contains statistics and a power-by statement.

https://sqid.toolforge.org/#/

SQID

Search Item

Start Properties Classes Rules About

## A fresh look at Wikidata

SQID is a fast way to browse and query [Wikidata](#), the free knowledge base of Wikipedia. SQID is inspired by Magnus Manske's [Reasonator](#), but it has a different focus. In particular, information about Wikidata classes and properties is prominently shown, including derived statistics and query results that are not part of Wikidata. SQID wants to help editors to improve Wikidata.


Example pages:

- [Johann Sebastian Bach](#): classical example of a data-rich item page
- [volcano](#): Wikidata item that is used as a class
- [galaxy](#): Wikidata class item that is part of the rather well designed classification of astronomical objects
- [sex or gender](#): a frequently used Wikidata property
- [instrument](#): another interesting Wikidata property
- [Wikidata property](#): the class of all Wikidata properties; shows how properties are organised in subclasses

**Properties** are special Wikidata entities that are used for describing relationships between entities, and for assigning many types of data values to entities. You can find all properties in SQID's [property browser](#).

**Classes** are Wikidata Items that are used as the value in an [instance-of](#) statement, or that are subject or value in a [subclass-of](#) statement. You can find all classes in SQID's [classes browser](#).

Most data shown by SQID is **live data** taken from the [Wikidata API](#) or the [Wikidata SPARQL query service](#). Complex statistics that cannot currently be computed live are pre-computed from the weekly data dumps using [Wikidata Toolkit](#). Some complex statistics can be fetched from SPARQL, but take too long to compute in the browser. These are updated every hour. See the [status page](#) for recency information and statistics on the current data.



Statistics based on data dump 4/1/2024 ([details](#))

Powered by [Wikidata Toolkit](#) & [Wikidata SPARQL Query](#)

## About SQID

SQID is a project of the [Knowledge-Based Systems Group of TU Dresden](#). Developers include Markus Krötzsch, Michael Günther, Markus Damm, and Georg Wild.

SQID is programmed in JavaScript with additional server-side code in Java and Python to extract, query, and preprocess data.

## What does SQID mean?

The origin of the name is unknown, but it seems to be an abbreviation of something. Current hypotheses include:

- Searching, Querying, and Interacting with Data
- Sweet QIDs
- SPARQL Querying Isn't Difficult
- Surprisingly Quick Information Display
- See Quality In Data

Another possible theory is that the name is indeed inspired by the high intelligence, amazing social skills, and sharp vision that are characteristic for many [cephalopods](#) and Wikidata editors alike.

## Where can I make comments and contributions?

All code is found in our [public repository](#). Bugs and wishes are best reported by [submitting an issue there](#). Pull requests are welcome.



Please don't eat me.



# Frederick Douglass (Q215562)

Frederick Augustus Washington Bailey | Frederick Augustus Washington Baly | Fred Bailey | Freddie Bailey

African-American social reformer, writer, and abolitionist (c. 1818 – 1895)

Human relationships	
Own statements	From related entities
<b>number of children</b>	5
<b>relative</b>	<p><b>Aunt Hester</b> (Gender: Female Life Dates: d. 1846 Types: African; American (US); Model of faith, piety; Servant; Slave, bondservant, former slave; Note: Entry refers to Aunt Hester as a "servant," but describes her "master" as buying and selling her (all told: sel)            kinship to subject : <b>aunt</b> (parent's sister or parent's sister-in-law)</p> <p><b>Fredericka Douglass Sprague Perry</b> (American social welfare activist)            kinship to subject : <b>granddaughter</b> (female grandchild. Avoid using with "relative" (P1038): add item for child instead and list there (both with "child" (P40)))</p> <p><b>Joseph Douglass</b> (American musician, grandson of Frederick Douglass)            kinship to subject : <b>grandson</b> (male grandchild. Avoid using with "relative" (P1038): create the missing generation instead)</p>
<b>spouse</b>	<p><b>Helen Pitts Douglass</b> (American suffragist (1838–1903))            start time : 1884            end time : 1895</p> <p><b>Anna Murray-Douglass</b> (American abolitionist)            start time : 1838            end time : 1882</p>
<b>child</b>	<b>Rosetta Douglass</b> (African-American teacher and activist)



### Links

[Wikidata page](#)

[Wikipedia article](#)

[Reasonator](#)

### Identifiers

**LOA ID** 215

**The Literary subject** 1302

**Encyclopedia named as : person ID** Freder...uglass

**CCAB ID** 000181635

**Open Library subject ID** person...-1895)



Find it

[Advanced](#)

Start Over

Access > At the library



### Limit your search

« Previous | 1 - 10 of 95 | Next »

Sort by Relevance ▾

10 per page ▾

Access ▾

At the library ✖ 95

Format >

Author/Creator >

Subject >

Language >

Library >

1. [The hypothalamus of the cat; a cytoarchitectonic atlas in the Horsley-Clarke co-ordinate system.](#)

Bleier, Ruth, 1923-  
Baltimore : John Hopkins Press, [1961]  
Book

Available for checkout

**Oversize QL937 B646 1961**

Book - Veterinary: Atwood Library (Campus) - Reserve Oversize



Bookmark

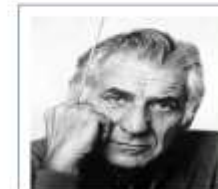


2. [Leonard Bernstein at Harvard. Vol. 3 : Musical semantics.](#)

Bernstein, Leonard, 1918-1990.  
Charles Eliot Norton lectures ; 1973.  
The Norton lectures, 1973: The unanswered question  
Masterwork



Bookmark



# First level of search results is using LD Microservice API to concatenate associations



A screenshot of a web browser window. The address bar shows the URL `localhost:5001/mmsid/9913203433503681`. The page content displays a JSON object with the following fields:

```
{  
  id: "9913203433503681",  
  qid: "Q18148474",  
  svde_agent: "961654264330943"  
}
```

# LD Microservice API

---

- Run through Docker Compose:
  - Flask API
  - Postgresql DB

```
ld-microservices
/Users/jimhahn/Documents/GitLab-Penn/ld-microservices

ld-microservices-...
ld-microservices-wel
Running
5001.5000

ld-microservices-...
ld-microservices-db
Running

2024-04-05 09:11:49 db-1 | PostgreSQL Database directory appears to contain a database; Skipping
Initialization
2024-04-05 09:11:49 db-1 |
2024-04-05 09:11:50 db-1 | 2024-04-05 14:11:50.070 UTC [1] LOG: starting PostgreSQL 13.14 (Debi
an 13.14-1.pgdg120+2) on aarch64-unknown-linux-gnu, compiled by gcc (Debian 12.2.0-14) 12.2.0, 64-
bit
2024-04-05 09:11:50 db-1 | 2024-04-05 14:11:50.076 UTC [1] LOG: listening on IPv4 address "0.0.
0.0", port 5432
2024-04-05 09:11:50 db-1 | 2024-04-05 14:11:50.076 UTC [1] LOG: listening on IPv6 address "::",
port 5432
2024-04-05 09:11:50 db-1 | 2024-04-05 14:11:50.078 UTC [1] LOG: listening on Unix socket "/var/
run/postgresql/.s.PGSQL.5432"
2024-04-05 09:11:50 db-1 | 2024-04-05 14:11:50.085 UTC [27] LOG: database system was shut down
at 2024-04-02 21:04:27 UTC
2024-04-05 09:11:50 db-1 | 2024-04-05 14:11:50.095 UTC [1] LOG: database system is ready to acc
ept connections
2024-04-08 14:23:20 db-1 |
2024-04-08 14:23:20 db-1 | PostgreSQL Database directory appears to contain a database; Skipping
Initialization
2024-04-08 14:23:20 db-1 |
2024-04-08 14:23:21 db-1 | 2024-04-08 19:23:21.004 UTC [1] LOG: starting PostgreSQL 13.14 (Debi
an 13.14-1.pgdg120+2) on aarch64-unknown-linux-gnu, compiled by gcc (Debian 12.2.0-14) 12.2.0, 64-
bit
2024-04-08 14:23:21 db-1 | 2024-04-08 19:23:21.004 UTC [1] LOG: listening on IPv4 address "0.0.
0.0", port 5432
2024-04-08 14:23:21 db-1 | 2024-04-08 19:23:21.004 UTC [1] LOG: listening on IPv6 address "::",
port 5432
2024-04-08 14:23:21 db-1 | 2024-04-08 19:23:21.006 UTC [1] LOG: listening on Unix socket "/var/
run/postgresql/.s.PGSQL.5432"
2024-04-08 14:23:21 db-1 | 2024-04-08 19:23:21.008 UTC [27] LOG: database system was interrupte
d; last known up at 2024-04-05 14:16:50 UTC
2024-04-08 14:23:21 db-1 | 2024-04-08 19:23:21.008 UTC [27] LOG: database system was not proper
ly shut down; automatic recovery in progress
2024-04-08 14:23:21 db-1 | 2024-04-08 19:23:21.001 UTC [27] LOG: redo starts at 0/1705C510
2024-04-08 14:23:21 db-1 | 2024-04-08 19:23:21.002 UTC [27] LOG: invalid record length at 0/170
5C5FB: wanted 24, got 0
2024-04-08 14:23:21 db-1 | 2024-04-08 19:23:21.002 UTC [27] LOG: redo done at 0/1705C5C0
2024-04-08 14:23:21 db-1 | 2024-04-08 19:23:21.006 UTC [1] LOG: database system is ready to acc
ept connections
```

- ID from response creates association for QID or SVDE.
  - demo is using ID to QID.
- The QID from Wikidata is used in the SQID Module.

The screenshot shows the Penn Libraries search interface. The browser address bar displays 'localhost:3000/?f[access\_facet][]=At+the+library'. The page header includes the Penn Libraries logo, a search bar, and navigation links for 'Databases' and 'Login'. The search results are filtered by 'Access > At the library' and show 95 results. The first result is '1. The hypothalamus of the cat; a cytoarchitectonic atlas in the Horsley-Clarke co-ordinate system.' by Bleier, Ruth, 1923-. The second result is '2. Leonard Bernstein at Harvard. Vol. 3 : Musical semantics.' by Bernstein, Leonard, 1918-1990. The interface includes filters for 'Access', 'Format', 'Author/Creator', 'Subject', 'Language', and 'Library', and options to 'Sort by Relevance' and '10 per page'.



<> \_author\_knowledge\_card.html.erb M X



find > app > views > catalog > <> \_author\_knowledge\_card.html.erb

```
You, 19 hours ago | 1 author (You)
1 <%= creator_ss = @document['creator_ss'].first if @document['creator_ss'] %>
2 <%= qid = fetch_qid_from_api(@document.id) %>
3 <%= iframe_src = "http://localhost:8080/entity/" + qid unless qid.nil? %> You, 1
4
5 <!-- Display button if qid is present -->
6 <%= if creator_ss.present? && qid.present? %>
7   <!-- Button trigger modal -->
8   <button type="button" class="btn btn-secondary mt-3 mb-3" data-bs-toggle="modal" d
9     | Author/Creator: <%= creator_ss %>
10  </button>
11
12 <!-- Modal -->
13 <div class="modal fade" id="knowledgeCardModal" tabindex="-1" aria-labelledby="exa
14   | <div class="modal-dialog modal-dialog-centered">
15     | <div class="modal-content">
16     | | <div class="modal-header">
```

# Modified SQID Module

---

- The SQID module will be containerized in Docker.



localhost:8080/entity/Q215562 80% ☆

## Frederick Douglass

Frederick Augustus Washington Bailey | Frederick Augustus Washington Baly | Fred Bailey | Freddie Bailey

African-American social reformer, writer, and abolitionist (c. 1818 – 1895)

### Selected Works

- North star (Rochester, N.Y.)
- My bondage and my freedom
- American slavery
- Voice ringing o'er the gale!
- One of the noblest, wisest and best men
- Correspondence between the Rev. Samuel H. Cox, D.D., of Brooklyn, L.I. and Frederick Douglass, a fugitive slave
- Frederick Donglass' paper
- DOUGLASS, F.: Story of an American Slave (The) (Unabridged)
- Frederick Douglass papers. Correspondence. Series three
- Oration by Frederick Douglass, delivered on the occasion of the unveiling of the Freedmen's Monument in memory of Abraham Lincoln, in Lincoln Park, Washington, D.C., April 14th, 1876



### Links

- [Wikidata page](#)
- [Reasonator \(Wikidata-app\)](#)

<https://github.com/jimfhahn/SQID>

You, 1 hour ago | 1 author (You)

```
1 // apiService.ts
2 import axios from 'axios';
3
4 export async function fetchSvdeAgentId(entityId: string) {
5   try {
6     const response = await axios.get('http://localhost:5000/qid2svde', {
7       params: {
8         qid: entityId,
9       },
10    });
11    return response.data.svde_agent_id;
12  } catch (error) {
13    console.error(error);
14  }
15 }
16
17 export async function fetchWorks(svdeAgentId: string) {
18   try {
19     const response = await axios.get('http://localhost:5000/agent2opus', {
20       params: {
21         svde_agent: svdeAgentId,
22       },
23    });
24    return response.data;
25  } catch (error) {
26    console.error(error);
27  }
28 }
```

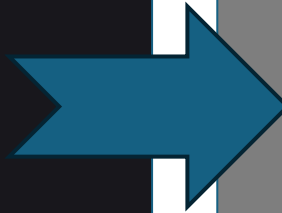
## Modifications to SQID

- Adding paths for API middleware service (Share-VDE GraphQL Middleware)



# Share-VDE Middleware


```
localhost:5000/agent2opus?svde_agent=961654264330943
JSON Raw Data Headers
Save Copy Collapse All Expand All Filter JSON
- data:
  - person:
    - opuses:
      - resources:
        - 0:
          title: "Science and gender"
          uri: "https://svde.org/opuses/1351654890883684"
        - 1:
          title: "Hypothalamus of the guinea pig"
          uri: "https://svde.org/opuses/1361654891359586"
        - 2:
          title: "Hypothalamus of the rhesus monkey"
          uri: "https://svde.org/opuses/1371654890921131"
        - 3:
          title: "Feminist approaches to science"
          uri: "https://svde.org/opuses/1451654887673856"
        - 4:
          title: "Hypothalamus of the cat"
          uri: "https://svde.org/opuses/1501654883645209"
```



Author Knowledge Card

## Ruth Bleier

Ruth Harriet Bleier  
American neurophysiologist and activist



**Selected Works**

- Science and gender
- Hypothalamus of the guinea pig
- Hypothalamus of the rhesus monkey
- Feminist approaches to science
- Hypothalamus of the cat

**Links**

- [Wikidata page](#)
- [Reasonator \(Wikidata-app\)](#)

Close

### Author Knowledge Card



## Ruth Bleier

Ruth Harriet Bleier

American neurophysiologist and activist



#### Selected Works



- Science and gender
- Hypothalamus of the guinea pig
- Hypothalamus of the rhesus monkey
- Feminist approaches to science
- Hypothalamus of the cat

#### Links



[Wikidata page](#)

[Reasonator \(Wikidata-app\)](#)

Close

```

You, 41 minutes ago | 2 authors (Maximilian Marx and others)
1 <template>
2   <sqid-bars>
3     <template v-slot:mainbar>
4       <div v-if="banner" style="overflow: hidden;">
5         <sqid-image :file="banner" width="850" />
6       </div>
7       <h1><span>{{ label }}</span></h1>
8       <div id="aliases">
9         <ul class="list-inline">
10          <li class="list-inline-item" v-for="alias in aliases" :key="alias.value">{{ alias }}</li>
11        </ul>
12      </div>
13      <div id="description">{{ description }}</div>
14      <sqid-collapsible-card :header="'Selected Works'" :id="'yourId'" narrow>
15        <b-card-body>
16          <ul>
17            <li v-for="(work, index) in works" :key="index" v-if="work.title">
18              {{ work.title }}
19            </li>
20          </ul>
21        </b-card-body>
22      </sqid-collapsible-card>
23    </template>
24    <template v-slot:sidebar>
25      <sqid-image :file="images[0]" width="260" v-if="images && images[0]" />
26      <sqid-collapsible-card :header="'$t('entity.links')'" :id="links" narrow>
27        <b-card-body>
28          <table class="table table-striped table-sm narrow">
29            <tbody>
30              <tr v-for="(link, lidx) in links" :key="lidx">
31                <th><a :href="link.url" target="_blank">{{ link.label }}</a></th>

```

# Adding templating for Works in SQID

<https://github.com/jimfhahn/SQID>

Filter JSON

```
"https://api-gw.staging.svde.org/works/13516548908837"
```

```
"https://api-gw.staging.svde.org/forms/11654262831406"
```

```
"https://api-gw.staging.svde.org/works/13516548908837"
```

```
"https://api-gw.staging.svde.org/languages/2165426284"
```

```
"https://api-gw.staging.svde.org/opuses/1351654890883"
```

```
"https://api-gw.staging.svde.org/works/13516548908837"
```

```
"https://api-gw.staging.svde.org/opuses/1351654890883  
ovenances"
```

```
"https://api-gw.staging.svde.org/works/13516548908837"
```

```
"https://api-gw.staging.svde.org/opuses/1351654890883"
```

```
"https://api-gw.staging.svde.org/works/13516548908837"
```

```
"https://api-gw.staging.svde.org/opuses/1351654890883"
```

```
"https://api-gw.staging.svde.org/works/20041656498124"
```

```
"https://api-gw.staging.svde.org/forms/11654262831461"
```

```
"https://api-gw.staging.svde.org/works/20041656498124"
```

```
"https://api-gw.staging.svde.org/languages/2165426284"
```

- Use Opus2Works service to gather Works associated with Opus

- Get subjects from Works

```
localhost:5000/work2subjects?work_id=1351654890883715
Raw Data Headers
Copy Collapse All Expand All Filter JSON
Added:
sourceList:
  links:
  provenances:
  - href: "https://api-gw.staging.svde.org/subjects/132165487398457"
  self:
  - href: "https://api-gw.staging.svde.org/subjects/132165487398457"
  subjectType:
    href: "https://api-gw.staging.svde.org/subjectTypes/top"
  heading:
  0:
  - _links:
    - concept:
        href: "https://svde.org/topics/1321654873984569"
        text: "Human biology"
  1:
  - _links:
    - concept:
        href: "https://svde.org/topics/1211654873780745"
        text: "Philosophy"
  links:
  provenances:
  - href: "https://api-gw.staging.svde.org/subjects/324416581386563"
  self:
  - href: "https://api-gw.staging.svde.org/subjects/324416581386563"
  subjectType:
    href: "https://api-gw.staging.svde.org/subjectTypes/top"
  heading:
  0:
  - _links:
    - concept:
        href: "https://svde.org/topics/32441658138656309"
        text: "Sexism"
  links:
  provenances:
```

# Microservices Architecture

Catalog

LD-Microservices (API)

Share-VDE Middleware (API)

SQID

Wikidata API / Wikipedia API  
Wikicommons

SVDE GraphQL  
+ SVDE API



# Still working on:

---

- Develop Airflow process to systematically curate PostgreSQL DB of IDs, QIDs, Share-VDE Agents.
- Design interactions for relationships among BF Works to what is available in the catalog
- Connecting Subjects to Works
- Subject Search
  - w/ SVDE API



# Linked Data Strategy

<https://guides.library.upenn.edu/linked-data>