

Knora

A Generic Storage System and API for Accessing and
Reusing Digital Qualitative Research Data in the
Humanities

B. Geer T. Schweizer

Digital Humanities Lab
University of Basel

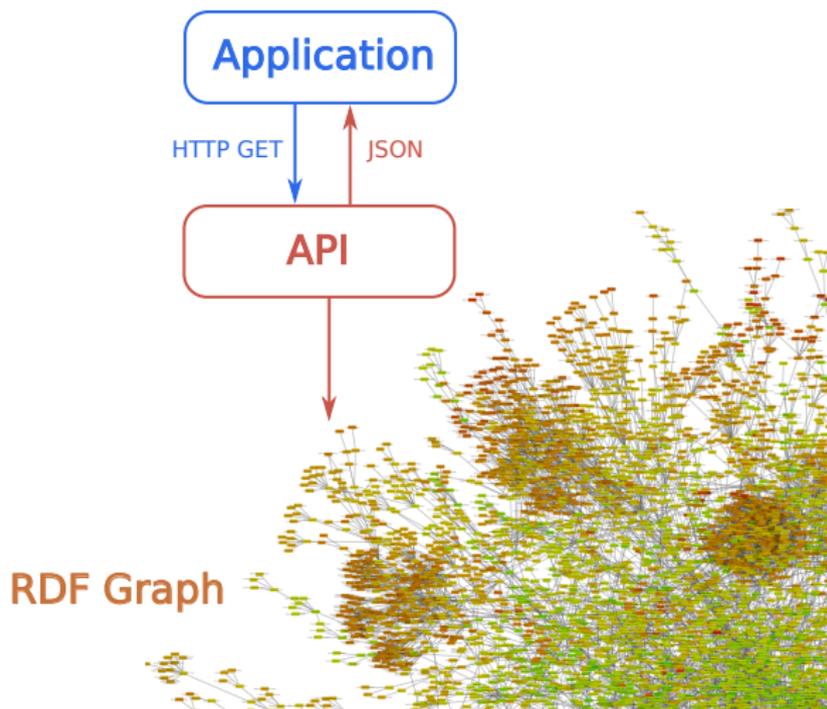
Forschungsdaten in den Geisteswissenschaften, 2015

No Unified Way to Work with Repositories



A collage of four different search interfaces, illustrating the lack of a unified way to work with repositories. The interfaces are: 1. TextGrid Repository: A search bar with a 'Search' button and radio buttons for 'Search Fulltext', 'Search Metadata', and 'Search Both'. 2. iDAI.objects arachne: A search bar with a magnifying glass icon and a 'Search' button, with a 'Total 3,044,870 Entries' indicator. 3. isidore: A search bar with a magnifying glass icon and a 'Search' button, with a 'Your search' label and a 'Social Sciences and Humanities' category. 4. NARCIS: A search bar with a magnifying glass icon and a 'SEARCH' button, with the text 'Search in NARCIS, the gateway to scholarly information in the Netherlands'.

Knora: A Common, Flexible Storage Model and API





TextGrid
Virtuelle Forschungsumgebung
für die Geisteswissenschaften

Search Contact Deutsch

Registration Download Community Support About TextGrid

Digital editing– research– archiving



Laboratory

Open source tools and services support humanistic scholars during the whole process of research, especially in the entire work flow of producing digital editions.



Repository

The long-term archive allows safe storing, publishing and researching of versatile digital material – e.g. XML/TEI formatted text, images and databases.



Community

The online help, mailing lists, a bug tracker and the source code enable users and developers to quickly start working with TextGrid.



Search ▾

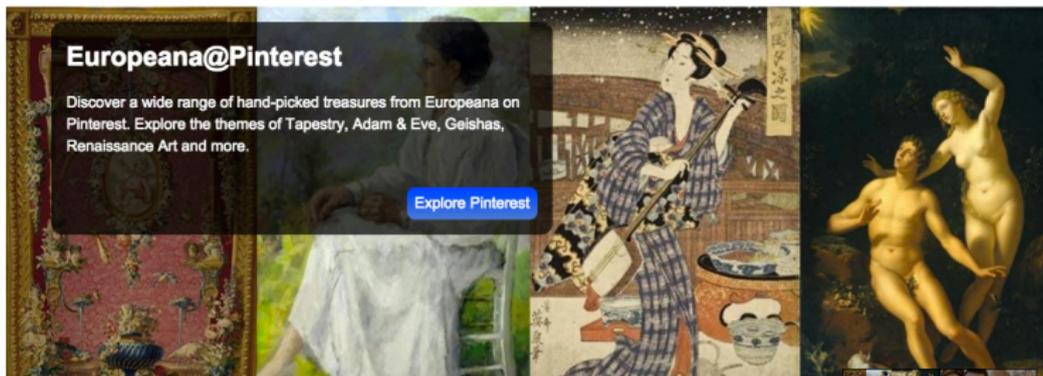
Search

Help

Europeana@Pinterest

Discover a wide range of hand-picked treasures from Europeana on Pinterest. Explore the themes of Tapestry, Adam & Eve, Geishas, Renaissance Art and more.

Explore Pinterest



Data Archiving and Networked Services (DANS)

Data Archiving and Networked Services

Nederlands English Contact Search this website

HOME DEPOSIT SEARCH TRAINING AND CONSULTANCY PROJECTS ABOUT DANS NEWS AND EVENTS

DANS promotes **sustained access** to digital research data files and encourages researchers to **archive** and **reuse** data.

ARCHIVING
Deposit your datasets in EASY or send research data and publications to NARCIS.
DEPOSIT

REUSE
Find datasets, publications, researchers, projects and institutions via NARCIS and EASY.
SEARCH

TRAINING & CONSULTANCY
Let DANS advise you on data management and certification of digital archives.
ADVICE FROM DANS

The screenshot displays the Isidore website interface. At the top left is the Isidore logo. In the top right corner, there are language selection buttons for 'Bêta', 'fr', 'en', and 'es'. Below the logo, the breadcrumb 'You are here : Home' is visible. The main search area features a search box with the placeholder 'Your search' and an 'OK' button. To the right of the search box are two blue buttons labeled 'Social' and 'Sciences and Humanities'. A large blue decorative graphic is on the left side of the page.

NUMBERS

Collections	129
Sources	3 126
Documents	3 710 932

BROWSE

- By novelties
- By document type
- By categories
- By historic periods
- By disciplines
- By nature of data sources
- By collections and organizations

SOURCES

Click the link below to consult the list of Isidore's sources

ZOOM IN...

Religions

7 lettres de Daniel Hymans (22 janvier 1965, 30 janvier 1965, 12 février 1965, 28 avril 1965, 13 octobre 1966, 17 novembre 1966, s. d.)
See more information | View the document card

Source: Calames

CONTRIBUTE

Libraries, A&H labs and research projects, you can help rechercheisidore.fr with your metadata and data.

MOBILE DEVICES

Explore « M », a ISIDORE website for mobiles devices: m.rechercheisidore.fr

NEWS

Tweets Follow

Isidore @rech_isidore 13h
ISIDORE Suggestions est

Fedora

[Home](#) [Contact](#)

[About](#) [Download](#) [Features](#) [Community](#) [Resources](#)



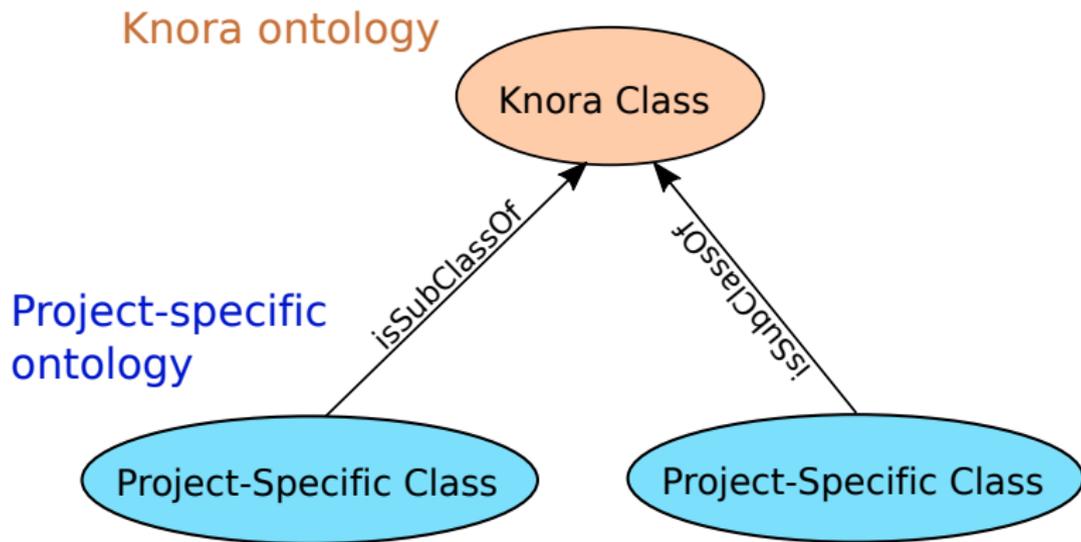
Robert Cartolano
Associate Vice President, Digital Programs and
Columbia University

"Fedora is the foundation of Columbia's digital preservation storage system. We chose Fedora in order to preserve and make accessible our digital collections because we can leverage the flexibility of the system for a wide variety of uses, and because we can work...

[READ MORE](#)



Generic, Extensible Ontologies



Generic, Extensible API

HTTP Request

```
GET /resources/http%3A%2F%2Fdata.knora.org%2Fc5058f3a HTTP/1.1
```

HTTP Response

```
{
  "resinfo": {
    "restype_id": "http://www.knora.org/ontology/incunabula#book",
    "restype_label": "Book",
    "person_id": "http://data.knora.org/users/91e19f1e01",
    "project_id": "http://data.knora.org/projects/77275339"
  },
  "props": {
    "http://www.knora.org/ontology/incunabula#title": {
      "pid": "http://www.knora.org/ontology/incunabula#hasTitle",
      "label": "Title",
      "occurrence": "0-n",
      "valuetype_id": "http://www.knora.org/ontology/knora-base#TextValue",
      "value_ids": [
        "http://data.knora.org/c5058f3a/values/c3295339"
      ],
      "values": [
        "Zeitglöcklein des Lebens und Leidens Christi"
      ]
    }
  }
}
```


SPARQL Is Hard

Traverses a tree from a leaf node upwards.

```
PREFIX rdfs: <http://www.w3.org/2000/01/rdf-schema#>
PREFIX knora-base: <http://www.knora.org/ontology/knora-base#>

SELECT DISTINCT ?node ?nodeName ?label ?child
WHERE {
    BIND(IRI("http://data.knora.org/lists/c7f07a3fc1") as ?leafNode)

    ?node knora-base:hasSubListNode* ?leafNode .

    MINUS {
        ?node knora-base:isRootNode "true"^^xsd:boolean .
    }

    OPTIONAL {
        ?node knora-base:hasSubListNode ?child .
        ?child knora-base:hasSubListNode* ?leafNode .
    }

    OPTIONAL {
        ?node knora-base:listNodeName ?nodeName .
    }

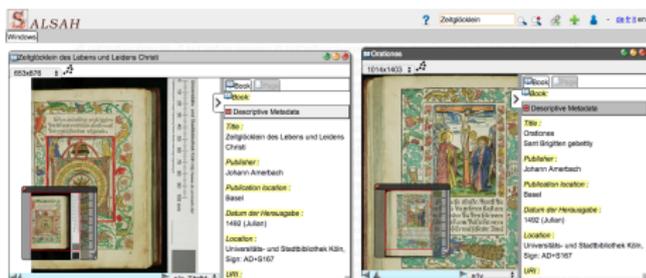
    OPTIONAL {
        ?node rdfs:label ?label .
    }
}
```



SPARQL Doesn't Have...

- ▶ Access control
- ▶ Versioning of data

Knora API Server



Virtual
Research
Environment

GET /resources/http%3A%2F%2Fdata.knora.org%2Fc5058f3a HTTP/1.1

Knora API Server

Generated SPARQL

```
SELECT ?p ?o
WHERE {
  BIND(IRI("http://data.knora.org/c5058f3a")
    as ?resource)
  ?resource ?p ?o .
  MINUS {
    ?p rdfs:subPropertyOf+ knora-base:hasValue .
  }
}
```

RDF Triplestore

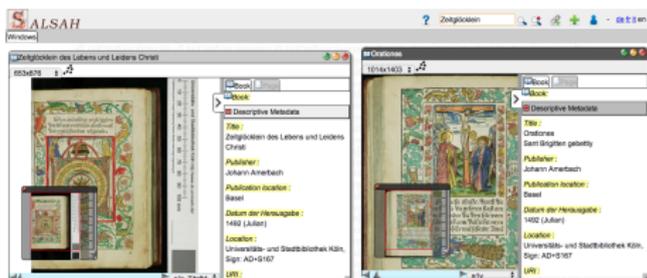


JSON Response

```
{
  "resinfo": {
    "restype_id":
      "http://www.knora.org/ontology/vincunabula#book",
    "restype_label": "Book",
    "person_id": "http://data.knora.org/users/91e19f1e01",
    "project_id": "http://data.knora.org/projects/77275339"
  }
}
```

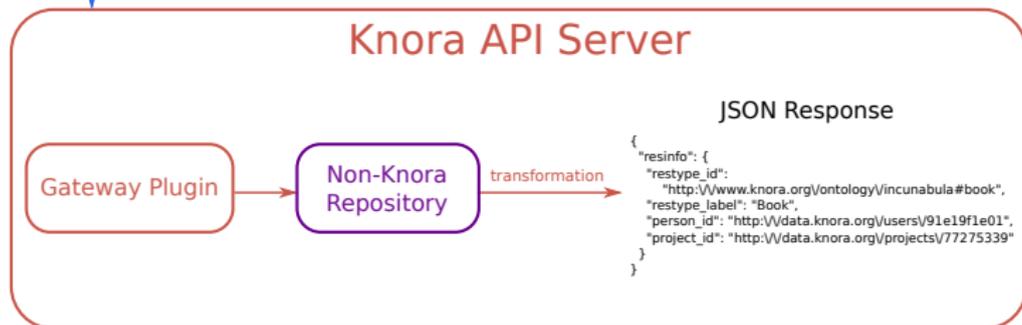
transformation

Knora Gateway to a Non-Knora Repository



Virtual
Research
Environment

GET /resources/http%3A%2F%2Fdata.knora.org%2Fc5058f3a HTTP/1.1



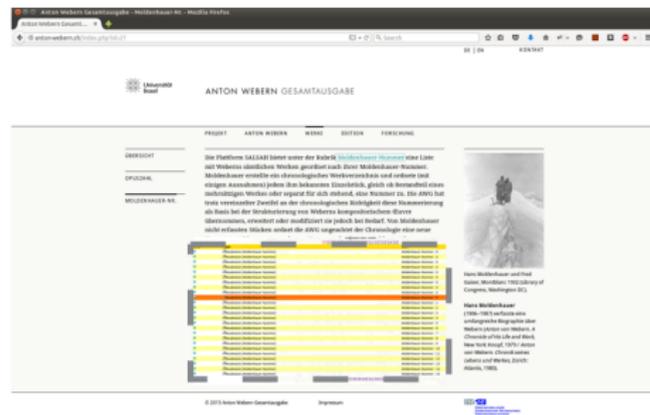
Use Case: Anton Webern Gesamtausgabe

- ▶ Currently, the project web site references Knora resources via hyperlinks
- ▶ This solution is technically insufficient
- ▶ The Knora API should do this job instead

The image displays two browser windows side-by-side. The left window shows the homepage of the 'Anton Webern Gesamtausgabe' project. It features a navigation menu with 'PROJEKT', 'ANTON WEBERN', 'BÜCHER', 'DIGITAL', and 'THERAPIEN'. Below the menu, there is a main text block with a paragraph that includes a Knora resource link: [https://knora.org/urn:nbn:de:hbz:5:1-64862-p0001-9](#). To the right of the text is a small image of a person. The right window shows a table of Knora resources, with columns for 'ID' and 'NAME'. The table lists 25 rows of resources, each with a unique ID and the name 'Anton Webern Gesamtausgabe'. An arrow points from the Knora link in the left window to the first row of the table in the right window.

Use Case: Anton Webern Gesamtausgabe Version 2

- ▶ The project website directly queries the Knora API
- ▶ The results are embedded in the project website
- ▶ Every project can build its own interface to its resources stored in Knora



The screenshot shows a web browser displaying the website for the Anton Webern Gesamtausgabe. The browser's address bar shows the URL antonwebern.ch/index.php?id=2. The website header includes the University of Basel logo and the title "ANTON WEBERN GESAMTAUSGABE". Below the header, there are navigation tabs for "PROJEKT", "ANTON WEBERN", "WERKE", "EDITION", and "FORSCHUNG".

The main content area is divided into several sections:

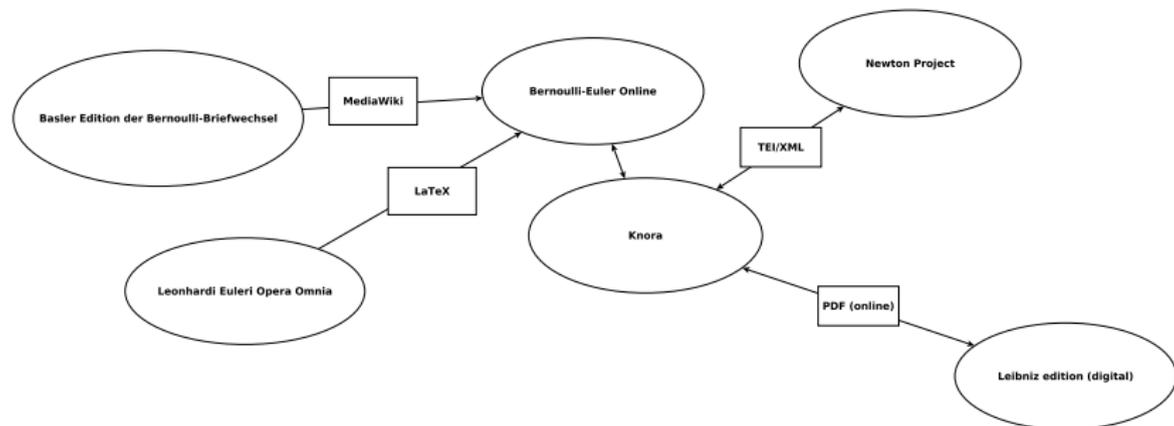
- ÜBERSICHT:** A short introductory paragraph about the project.
- EDITION:** A list of works with columns for title, date, and location. The list is partially obscured by a yellow highlight.
- BIODATEN:** A section containing a photograph of Anton Webern and a biographical text.

The biographical text reads: "Anton Weibern (1868-1948) war ein Schweizer Komponist und Pianist. Er war einer der bedeutendsten Schweizer Komponisten der ersten Hälfte des 20. Jahrhunderts. Er war einer der bedeutendsten Schweizer Komponisten der ersten Hälfte des 20. Jahrhunderts. Er war einer der bedeutendsten Schweizer Komponisten der ersten Hälfte des 20. Jahrhunderts." (Note: The text in the image is partially obscured and appears to be a placeholder or a very short biography).

At the bottom of the page, there is a copyright notice: "© 2013 Anton Webern Gesamtausgabe" and a logo for the University of Basel.

Use Case: Bernoulli-Euler Online (BEOL)

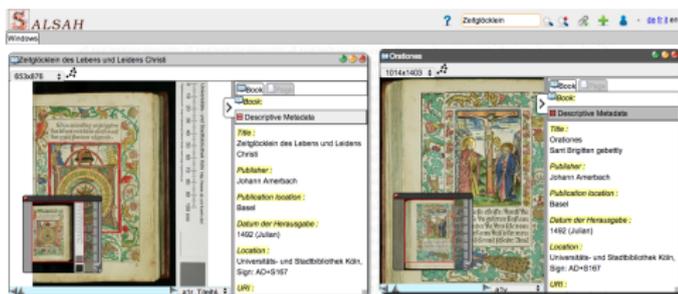
- ▶ Unification of two editions in one digital platform
- ▶ BEOL is an edition and a virtual research environment
- ▶ Knora allows to integrate third party repositories (proxy resources containing external ID, offer our data to them via the API)



Knora API Version 2



Conclusion: A Distributed Architecture



Virtual
Research
Environment

