



Akademie der Wissenschaften
und der Literatur | Mainz



FIZ Karlsruhe

Leibniz Institute for Information Infrastructure



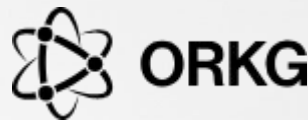
Sci-K Workshop, co-located with ISWC2025 | Nara, Japan | 02 November 2025

Knowledge Representation and Discovery for Cultural Heritage Research Data with CTO and SHMARQL

Tabea Tietz, Etienne Posthumus, Linnaea Söhn, Jonatan Jalle Steller,
Oleksandra Bruns, Joerg Waitelonis, Torsten Schrade and Harald Sack

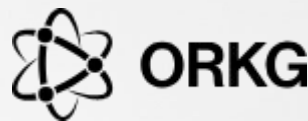
Research Knowledge Graphs

- Massive increase in **scientific publications** and **research artifacts**
- Need for structured, interoperable, and machine-understandable knowledge
- Traditional data management struggles with **heterogeneity** and **fragmentation**
- Drives the development of **Research Knowledge Graphs** (RKGs) and open data initiatives



Research Knowledge Graphs

- **Represent research (meta) data** in a semantically rich, interlinked way
- **Capture relationships** between datasets, collections, software, publications
- Enable **advanced search, analysis**, and **reuse** across disciplines
- Ensure **interoperability** and scalability for knowledge integration



Cultural Heritage Domain

- The **Cultural Heritage** (CH) domain illustrates both the need for and the potential of Research Knowledge Graphs (RKGs)
- CH research data are diverse, and often distributed across **heterogeneous** sources
- Each domain uses its own metadata standards, vocabularies, and formats



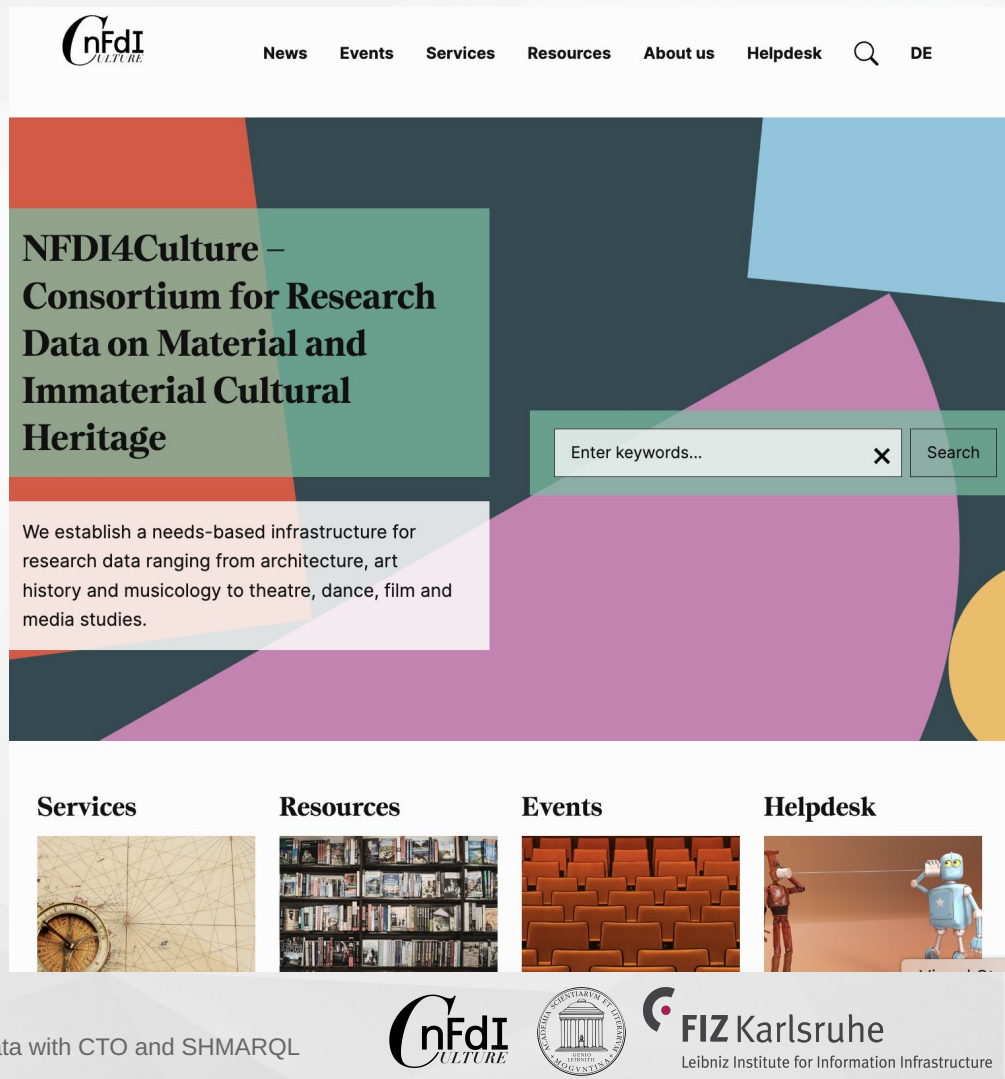
NFDI4Culture

- **NFDI4Culture** brings together tangible and intangible cultural heritage research data
- Part of the German National Research Data Initiative (**NFDI**)
- **Domains:** Architecture, Art History, Performing Arts, Media Studies, Musicology



NFDI4Culture

- **NFDI4Culture** brings together tangible and intangible cultural heritage research data
- **Goal:** NFDI has been creating a universal access point to decentralized and heterogeneous research data
- NFDI4Culture **Information Portal** and **Knowledge Graph**



NFDI4Culture Ontology (CTO)



RQ1:

How can cultural heritage research resources be represented to establish a **unified index** and **centralized access point** for **decentralized and heterogeneous research data** within the (German) cultural heritage landscape?

NFDI4Culture Ontology (CTO)



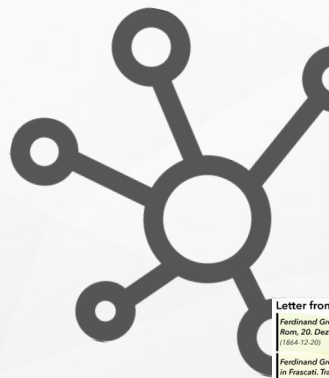
RQ2:

How can the requirements of the **cultural heritage domain** for the representation and discovery of research resources be met, while also providing **interoperability to other domains**, i.e. other NFDI consortia and beyond?

Contributions

1 NFDI4Culture Ontology (CTO) v.3.0

- Represents an **index of the decentralized CH research metadata**
 - Performing arts
 - Musicology
 - Art history
 - Media studies
 - Architecture
- Extends the **NFDIcore** mid-level ontology for cross-disciplinary interoperability.
- Aligns with the Basic Formal Ontology
- Builds the foundation of the **NFDI4Culture KG** and the **NFDI4Culture Portal**



NFDI4Culture Data Search

Find multimodal research data with our knowledge graph powered search

Explore Data Data Providers SPARQL Query Help

Data provider: Semantic Kompakt
Data feed: Metadata of 3D models in Semantic Kompakt

Central room

Permalink Provider Item Share Actions Graph Explorer

3D Model: "Central room" | CC-BY-SA 4.0 | Created by: Tim Pries. Go to provider item.

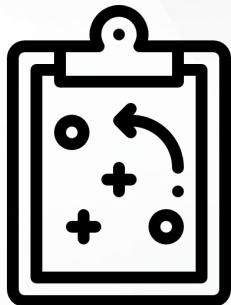
Data About:
http://www.cidoc-crm.org/cidoc-crm/E22_Human-Made_Object

Classification:
<http://vocab.getty.edu/ia/300047753>

Letter from Gregorovius edition	Musician	Musical source from RISM
Ferdinand Gregorovius an Tommaso Gar in Neapel. Rom, 20. Dezember 1864 (1864-12-20)	Rossini, Gioacchino (1792-1868)	Fidelio-E-Dur: Druck; LbWVv op.72 (1792-1868)
Ferdinand Gregorovius an Terenzio Maniani della Rovere in Frascati, Traunstein, 15. Juli 1873 (1873-07-15)	Mendelssohn Bartholdy, Felix (1809-1847)	Guillaume Tell. Bouquet de mélodies (Bearbeitung): Abschrift; D-Mbs Mus.Schott.As.374 (1872-01-01T00:00:00-1873-01-01T00:00:00)
Ferdinand Gregorovius an Raffaele Mariano in Maloja, Traunstein, 26. August 1874 (1874-08-26)	Beethoven, Ludwig van (1770-1827)	Symphonies-E-Dur: Abschrift; LbWVv op.55, KInb.55; D-RU SH.46 (1860-01-01T00:00:00-1874-01-01T00:00:00)
Ferdinand Gregorovius an Mathilde von Humboldt in Ottmachau, Rom, 15. Februar 1880 (1880-02-15)		Massees-D-Dur (Ausschnitt); Abschrift; D-WRba Mus.ms. B.104 (1880-01-01T00:00:00-1900-01-01T00:00:00)
Ferdinand Gregorovius an Mathilde von Humboldt in Ottmachau, München, 6. August 1886 (1886-08-06)		Ave verum-g-Moll (Bearbeitung): Abschrift; Real. 500; D-Lsta 21069 C.F. Kahnt, Nr.96 (1885-01-01T00:00:00-1895-01-01T00:00:00)
Ferdinand Gregorovius an Hermann von Thile in Berlin, München, 12. März 1887 (1887-03-12)		Symphonies-d-Moll; Abschrift; MWV N.15; D-SHm Mus.C83.12 (1860-01-01T00:00:00-1885-01-01T00:00:00)
Ferdinand Gregorovius an Gertrud Gräfin von Werthern-Beichlingen in München, Rom, 28. April 1888 (1888-04-28)	Liszt, Franz (1811-1886)	Hunnenschlacht; Abschrift; Real. 422; D-SHm Mus.C80.14 (1885-01-01T00:00:00-1886-01-01T00:00:00)
		Der Doppelgänger-g-Moll (Bearbeitung): Abschrift; D1 957/12, Real. 651/5; D-Lsta 21069 C.F. Kahnt, Nr.113 (1860-01-01T00:00:00-1887-01-01T00:00:00)

NFDI4Culture Ontology (CTO)

Ontology Design



Competency Questions
extracted from user stories



Workshops with domain
experts from academia,
institutes, archives,
libraries, etc.



Integration into operational
environments
(NFDI4Culture Portal, KG)

NFDI4Culture Ontology (CTO)

Requirements v3.0

- **Complexity:** increase the expressivity of the ontology while still maintaining the overall goal of a lightweight and pragmatic design.
- **Domain Requirements and Granularity:** Individual domains like musicology and the performing arts required more expressive representations than initially anticipated.
- **Upper/mid-level Ontology Alignment:** Align CTO with the updated NFDIcore ontology/BFO 2020 [1,2] to ensure semantic consistency and interoperability across NFDI consortia.
- **Transparency, Collaboration, and Quality Control:** Establish a more transparent, reproducible, and quality-controlled ontology development process

NFDI4Culture Ontology (CTO)

A Modular Approach with NFDIcore

- **National Research Data Infrastructure** in Germany spans over all scientific disciplines and humanities
- Each consortium has **distinct domain-specific tasks**, needs, and modeling requirements.
- Consortia also **share common concepts** and structures, interoperability goals.



NFDI4Culture Ontology (CTO)

A Modular Approach with NFDIcore

Which resources from the 19th century related to Ludwig van Beethoven are annotated in MEI?

Which materials contribute to the preservation and restoration of cultural heritage objects?

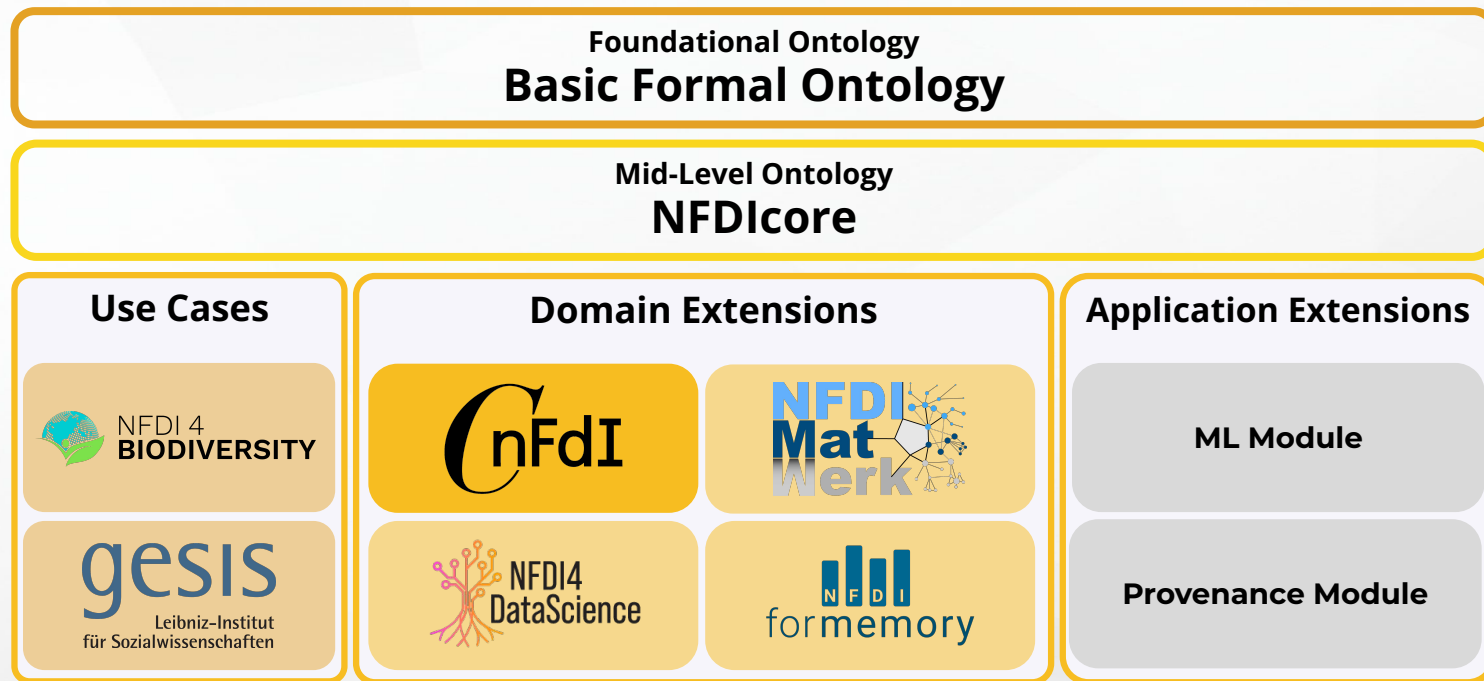
NFDI4Culture Ontology (CTO)

A Modular Approach with NFDIcore

Balancing the shared goals and structures of NFDI consortia with their individual disciplinary requirements.

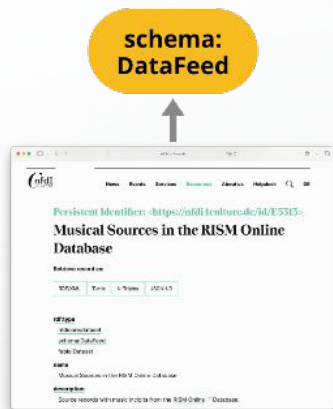
NFDI4Culture Ontology (CTO)

A Modular Approach with NFDIcore



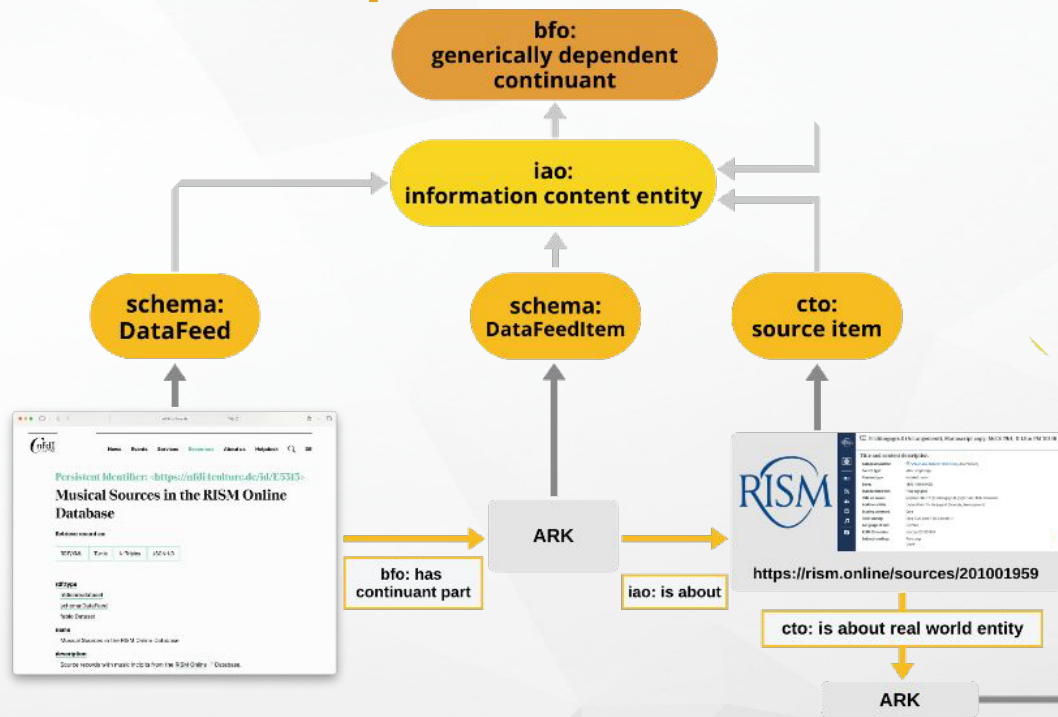
NFDI4Culture Ontology (CTO)

Domain Example



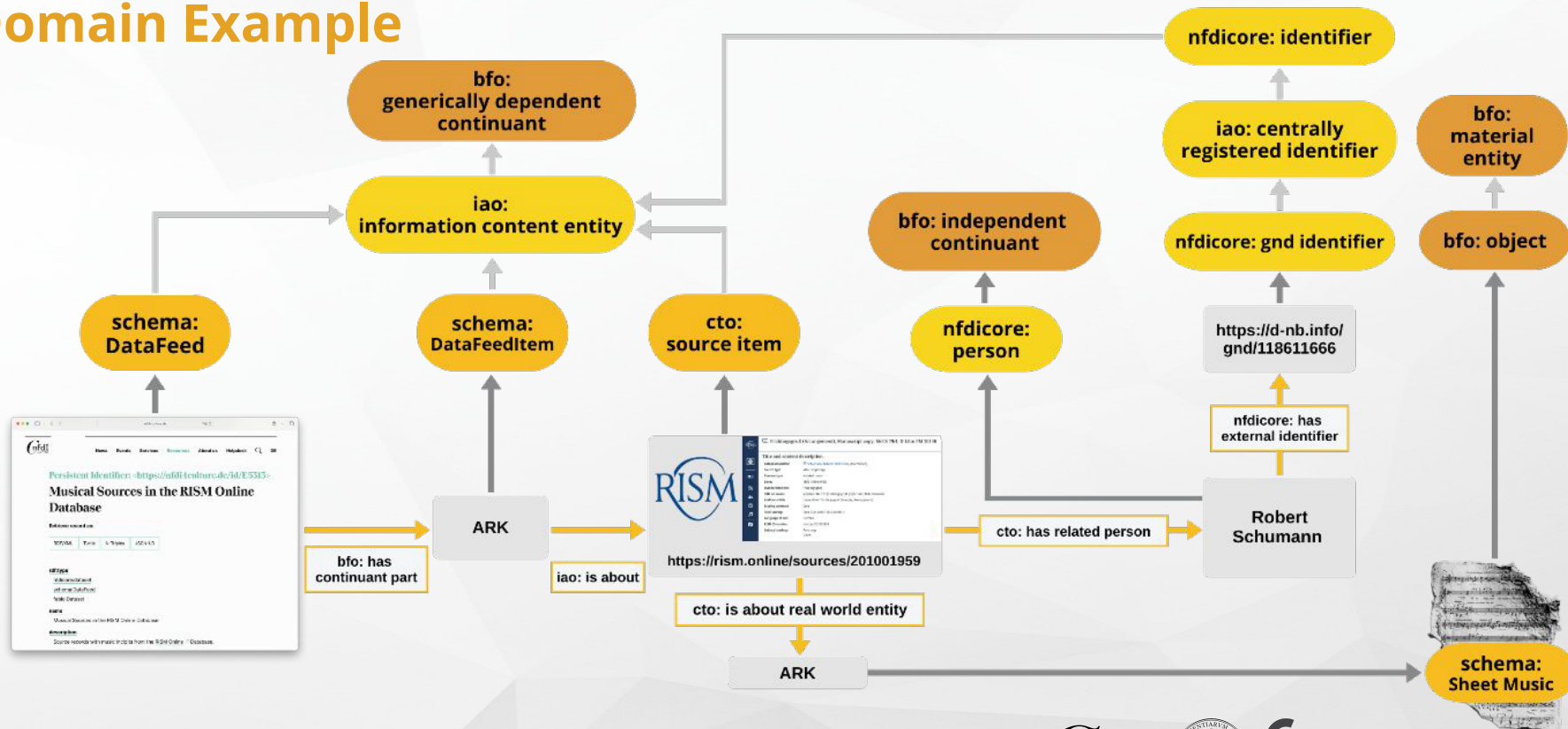
NFDI4Culture Ontology (CTO)

Domain Example



NFDI4Culture Ontology (CTO)

Domain Example



NFDI4Culture Ontology (CTO)

Summary - Representation

- CTO provides a **lightweight** yet **expressive** framework for integrating decentralized CH metadata into a centralized, queryable infrastructure within NFDI4Culture
- Currently interconnects more than 18mio resources within NFDI4Culture
- Reused in the NFDI4Memory consortium



SHMARQL



RQ3:

How can cultural heritage research resources be **published** and **queried** effectively to enable the **discovery of decentralized research resources** through a unified and centralized access point?

Contributions



2 SHMARQL

- Lightweight web user interface for scholars to intuitively **explore existing SPARQL endpoints**
- Create queries, visualize results without deep technical expertise
- Semantic web **application framework** on which experienced practitioners can build custom projects and disseminate data effectively
- Originally created for NFDI4Culture to support data publishing → **generalized** into a versatile platform adopted by multiple projects
- Deployable with a single Docker command

SHMARQL

Features

Prefixes

```
1 select ?s ?p ?o ?olabel where {  
2   values ?s { <https://nfdi4culture.de/id/E3590> }  
3   ?s ?p ?o .  
4   optional {  
5     ?o rdfs:label ?olabel .  
6   }  
7 }  
8 } limit 999
```

33 results in 269ms

#	s	p	o	olabel
1	nfdi4culture:E3590	nfdicore:publisher	nfdi4culture:E1835	Leibniz Institute for Information Infrastructure
2	nfdi4culture:E3590	nfdicore:publisher	nfdi4culture:E1835	FIZ Karlsruhe – Leibniz Institut für Informationsinfrastruktur



- Interactive browsing of Knowledge Graphs
- **Navigate through triples** by clicking IRIs in subject or object positions
- Eliminates manual SPARQL query creation when exploring new graphs

SHMARQL

Features - Fuzzysearch

- Built-in SPARQL re-writer **Fuzzysearch**
- Extra search options to an existing triplestore:
 - Flexible **fulltext** searches
 - **RDF2Vec** searches

Fulltext

When an **FTS index has been configured**, you can do queries that look like this:

```
1 | select ?s where {  
2 |   ?s fuzzy:fts "something" .  
3 | }
```

This will then do a fulltext query looking for all triples that have a literal string matching the word "something". It is also possible to use wildcards and boolean queries.

RDF2vec

When a **RDF2Vec index has been configured**, you can do queries that look like this:

```
1 | select distinct ?s (STR(?o) AS ?oLabel) where {  
2 |   ?s fuzzy:rdf2vec <https://example.com/entity/xyz> .  
3 |   ?s rdfs:label ?o .  
4 | }
```

This searches for all entities similar to `<https://example.com/entity/xyz>`, and then selects the `rdfs:label` for display.

Use Cases: CTO and SHMARQL

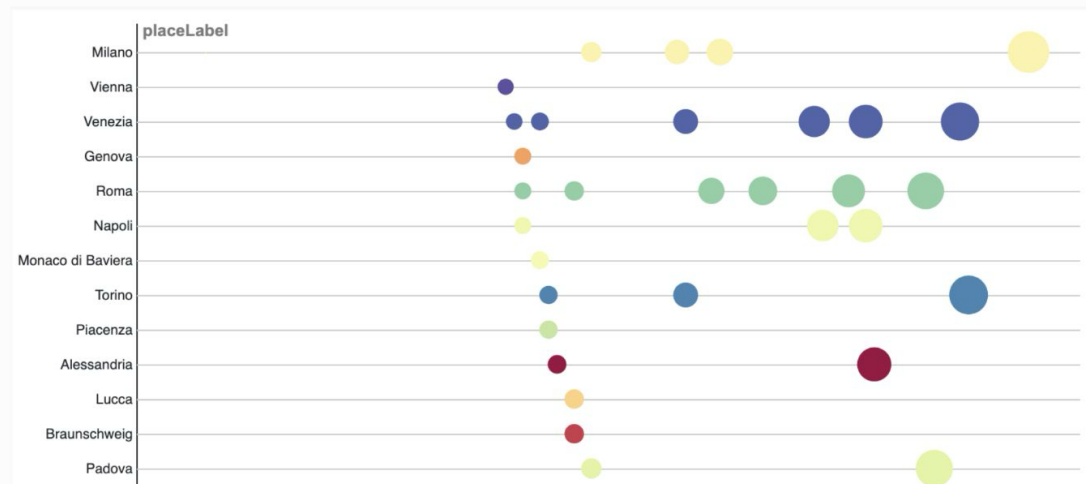
Use Cases: CTO and SHMARQL

Data Stories

- SHMARQL and CTO form the semantic backbone of the NFDI4Culture **Data Stories** platform based on the data in the NFDI4Culture-KG
- Combines semantic technologies with **storytelling** to foster data literacy and to provide a starting point for domain experts

Example 10: Staging history of the opera Demofonte in the 18th century

Show SPARQL query 10



Use Cases: CTO and SHMARQL

Data Stories

- Uses narratives, visualizations, and interactive SPARQL queries within a MKdocs environment
- Enables the exploration of semantically connected CH datasets within the NFDI4Culture Knowledge Graph



NFDI4Culture Data Stories

Search

Home

Stories

01 An Italian Data Journey

Contribute

SPARQL Exploration

Gefördert durch die Deutsche
Forschungsgemeinschaft (DFG)
Projektnummer 441958017

Portal Imprint Privacy

An Italian Data Journey

Analysing research data about 18th century Italian opera using the Culture Knowledge Graph and federated European research infrastructures

Author: Torsten Schrade

Persistent Identifier: <https://nfdi4culture.de/id/E6263>

Metadata: <https://nfdi4culture.de/id/E6263/about.html>

License: CC-BY 4.0



Daniel de Lafeuille, Nouvelle Carte D'Italie - Nieuwe Kaart van Italien, 1706, Wikimedia Commons, Public Domain

Abstract: This data story illustrates a digital exploration of research data on opera holdings in the Doria Pamphilj Archive by the *Partitura* project of the German Historical Institute in Rome (DHI Rome). By enriching the *Partitura* dataset with established authority sources such as Wikidata, RISM, GeoNames, and transforming it to LOD, new analytical insights into the



Meet us at
ISWC!

Thank you!

Tabea Tietz, Etienne Posthumus, Linnaea
Söhn, Jonatan Jalle Steller,
Oleksandra Bruns, Joerg Waitelonis,
Torsten Schrade and Harald Sack

tabea.tietz@fiz-karlsruhe.de

SHMARQL



CnFdI
CULTURE



Akademie der Wissenschaften
und der Literatur | Mainz



FIZ Karlsruhe

Leibniz Institute for Information Infrastructure