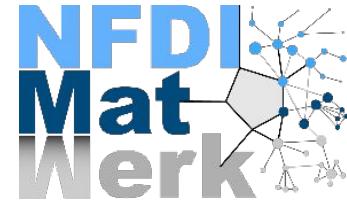


nationale  
Forschungsdaten  
Infrastruktur  
*for CULTURE*



17.04.2023

# Knowledge Graph based RDM Solutions for nfdi4culture, NFDI-MatWerk and beyond

Prof. Dr. Harald Sack

NFDI4DataScience Plenary, Berlin, 18.04.2023

# Knowledge Graphs @ FIZ Karlsruhe

Projects, Services & NFDIs



daily blog on science, tech & art in history



100 JAHRE  
WEIMARER  
REPUBLIK  
Weimar –  
Die erste deutsche  
Demokratie



2020



2021



2023



# Knowledge Graph based RDM Solutions for nfdi4culture, NFDI-MatWerk and beyond

## Outline

### 1. NFDI Ontologies

- Overview: nfdicore + modules
- Status and Future Plans

### 2. nfdi4Culture & NFDI-MatWerk Knowledge Graphs

- nfdi4Culture-KG v1.0 - Research Information Graph
- nfdi4Culture-KG v2.0 - Research Data Graph
- Data Integration Strategies

### 3. nfdi4Culture Portal

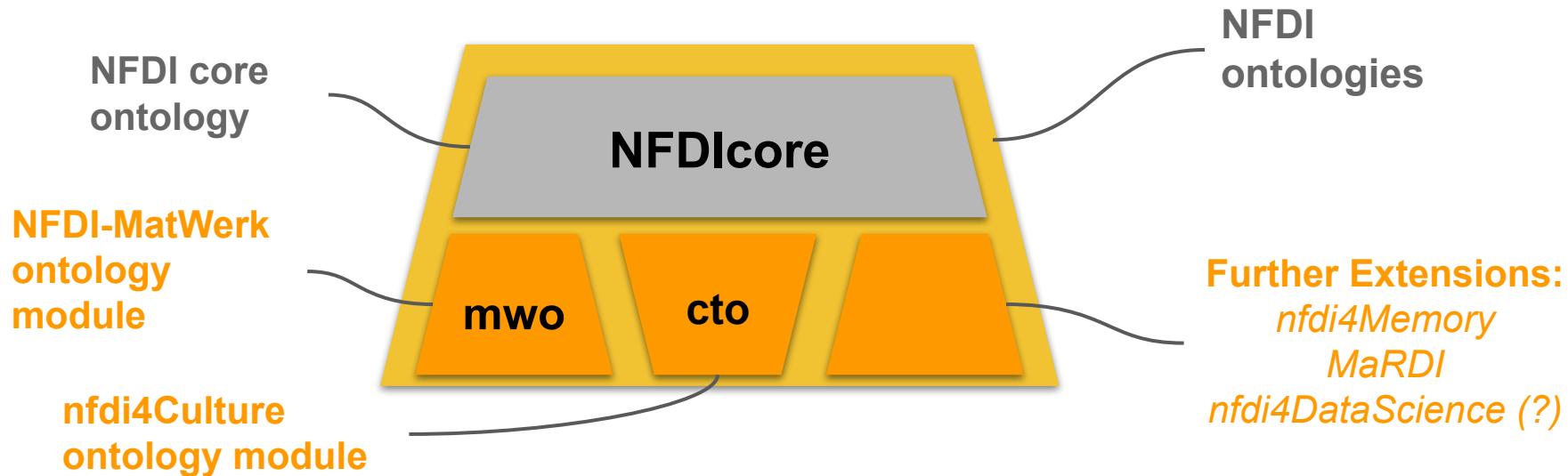
- Architecture
- Access & Curation Strategy

### 4. Next Steps

# 1. NFDI Ontologies

# Overview

- NFDI ontologies follow **modular approach**
- NFDIcore v1.1.0 represents **NFDI consortia in general**
- **NFDI modules** represent **specific consortia**, as e.g. NFDI-MatWerk
- Specific requirements of consortia lead to **ontology extension**

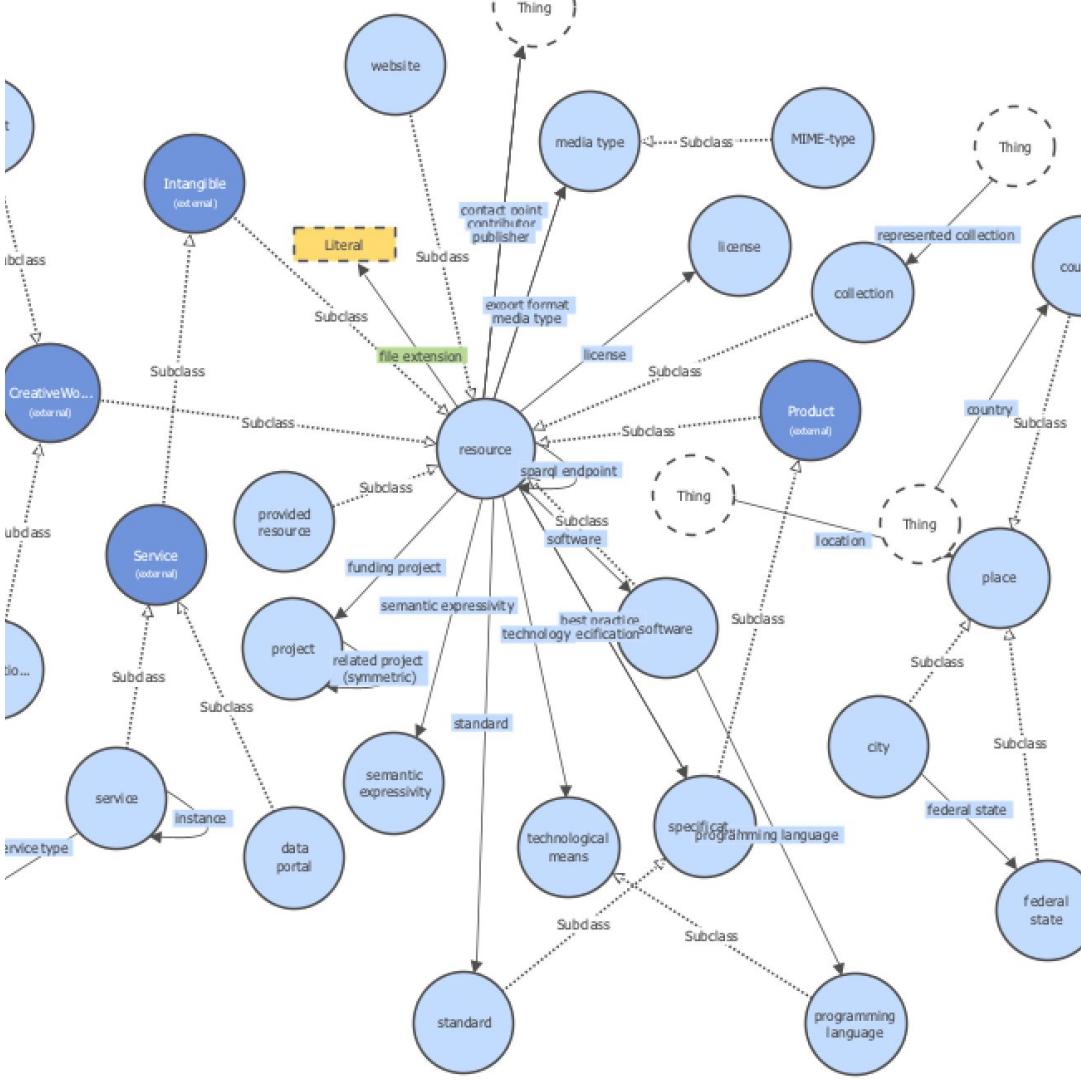


# NFDIcore Ontology

temporary namespace:

<https://nfdi.fiz-karlsruhe.de/ontology>

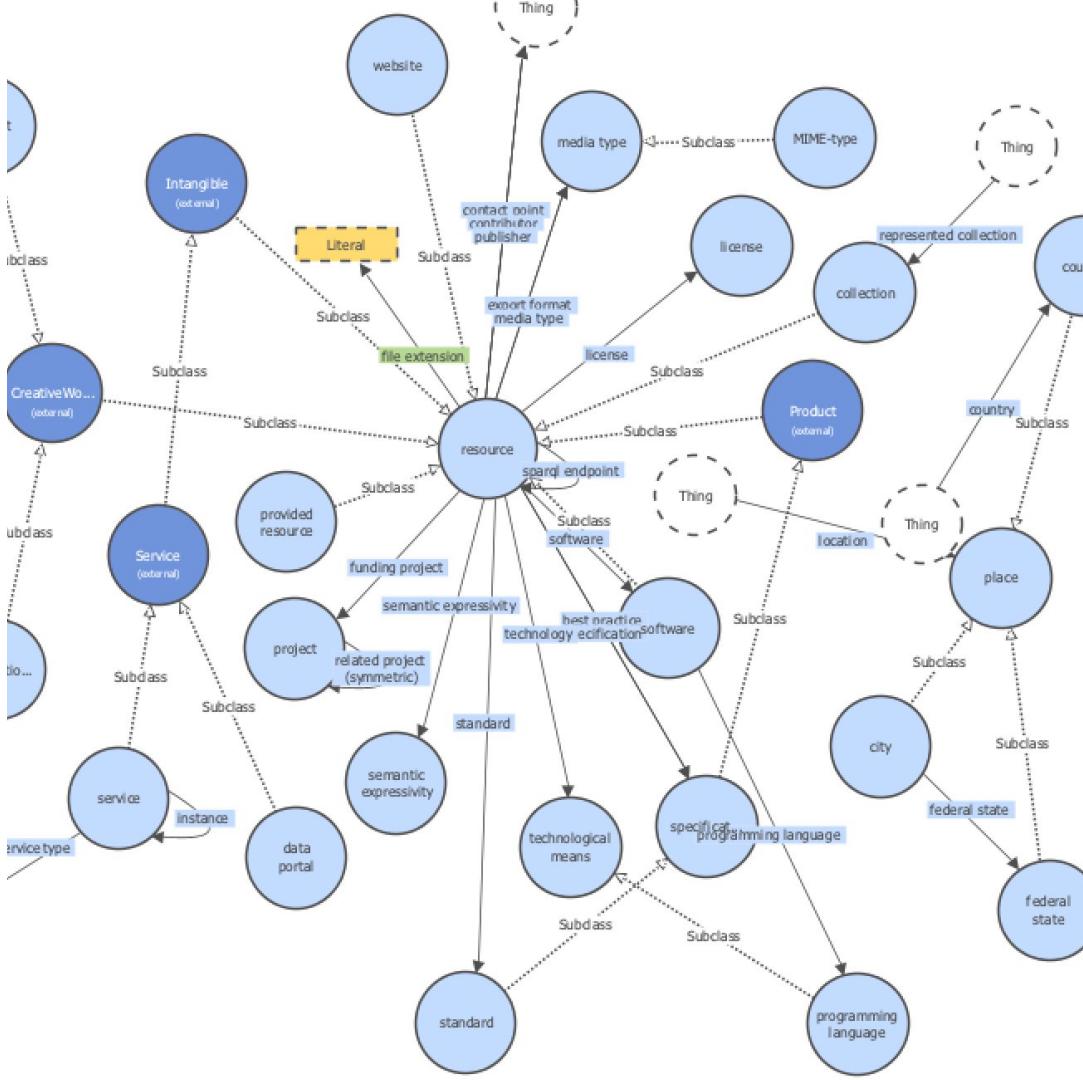
- Represents resources of NFDI consortia, e.g. services, data sets, guidelines, etc.
- Bottom-up approach and in constant feedback with NFDI4Culture and NFDI-MatWerk consortia  
→ in constant development
- Contains 33 classes and 60 properties



# NFDIcore Ontology

temporary namespace:  
<https://nfdi.fiz-karlsruhe.de/ontology>

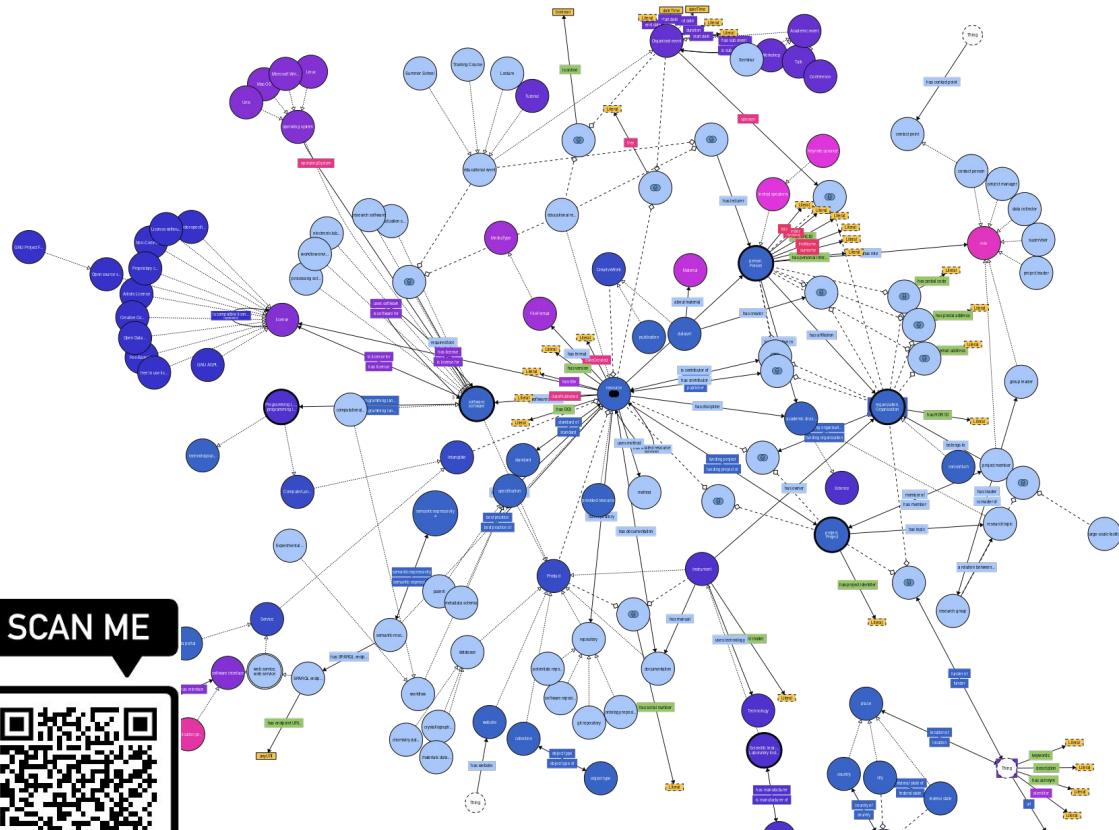
- Links to over 20 external vocabularies including schema, fabio, frapo, void, dcterms, foaf
- Available online  
<https://github.com/ISE-FIZKarlsruhe/nfdicore>  
Most recent release: version 1.1
- Documentation of the development version:  
[https://nfdi.fiz-karlsruhe.de/ontology\\_v1.1.0-alpha](https://nfdi.fiz-karlsruhe.de/ontology_v1.1.0-alpha)



# MatWerk Ontology (MWO)

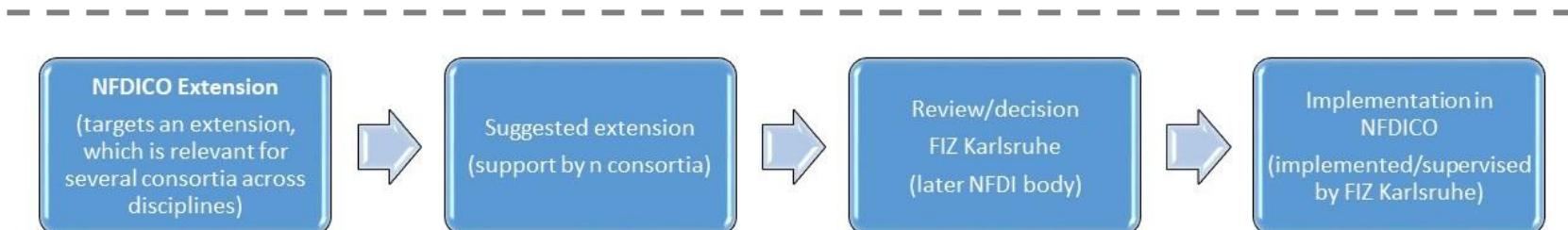
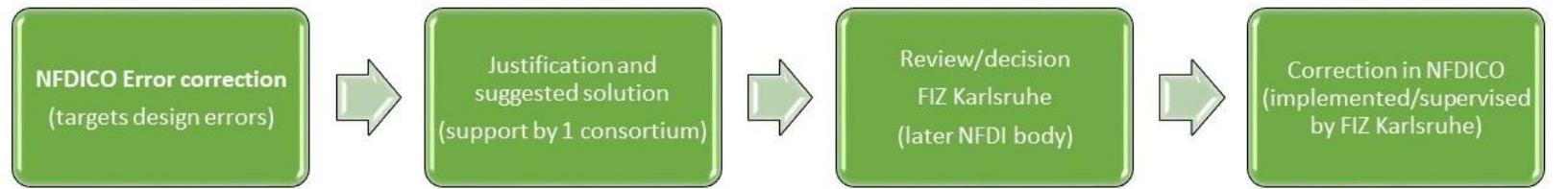
- MWO ontology - release: v1.0.0
- Based on NFDICORE v1.0.2
- Contains 115 classes and 103 object and data properties
- Mapped to 11 external vocabularies
- Documentation can be found at:  
<http://purls.helmholtz-metadaten.de/mwo>
- Development version:  
<https://git.rwth-aachen.de/nfdi-matwerk/ta-oms/mwo>
- Modular design based on  
NFDI Core  
+ Domain extension (MWO)

SCAN ME



# Work in Progress: NFDIcore Ontology Curation Practice

## Errors



## Extensions

## 2. nfdi4Culture & NFDI-MatWerk Knowledge Graph(s)

**Lessons learned from 3 years of NFDI:**

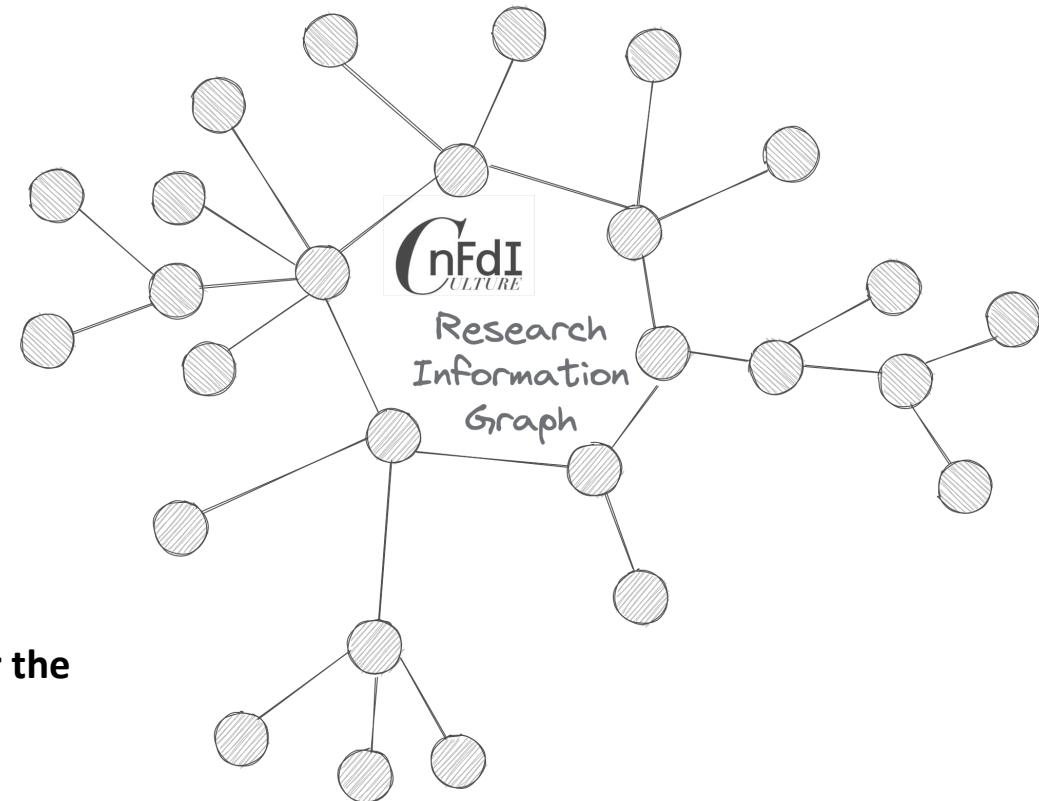
*“We will never succeed to convince all applicants and participants to actually convert their research data to fully fledged Linked (Open) Data ...”*

# nfdi4culture-KG v1.0 - Research Information Graph



## CONTENT:

- [Index and Metadata for nfdi4culture Project Resources](#)
  - nfdi4culture Data Resources
  - nfdi4culture Data Services
  - Persons and Institutions
  - Expertises



## PURPOSE:

- Backend (Meta-)Data Resource for the nfdi4culture Portal

# nfdi4culture-KG v2.0 - Research Data Graph

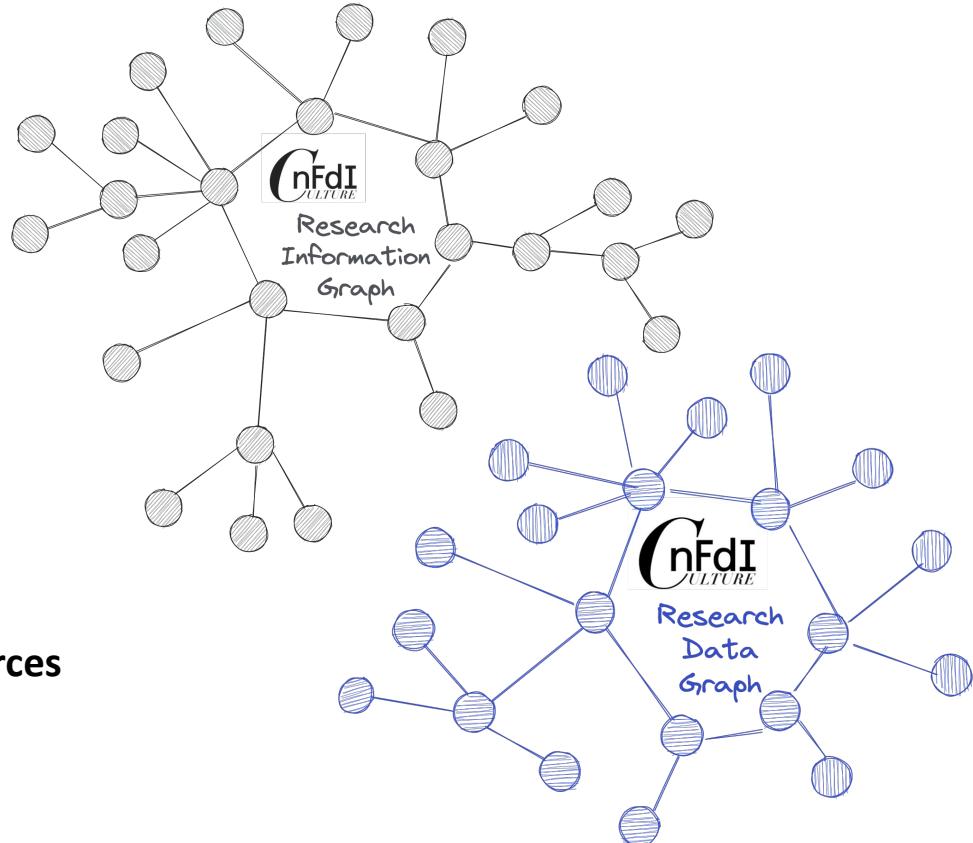


## CONTENT:

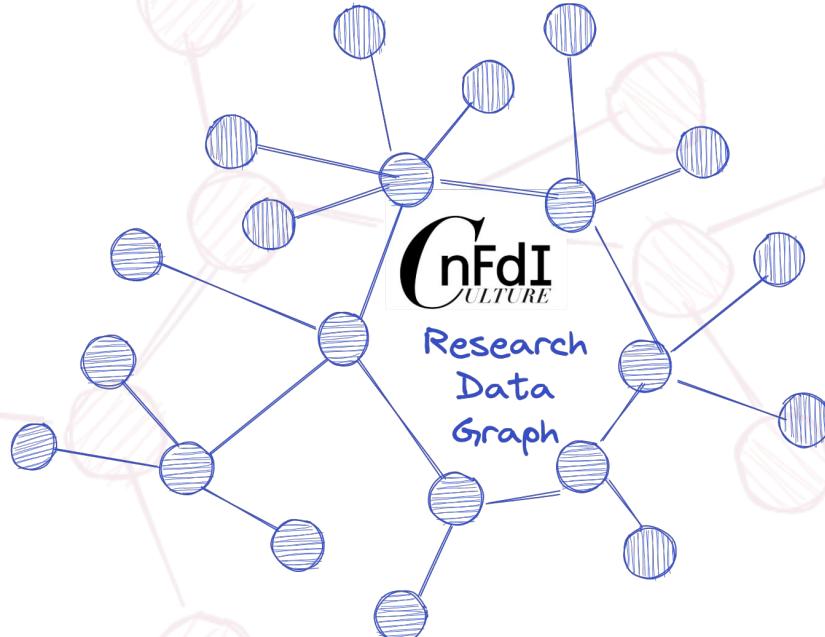
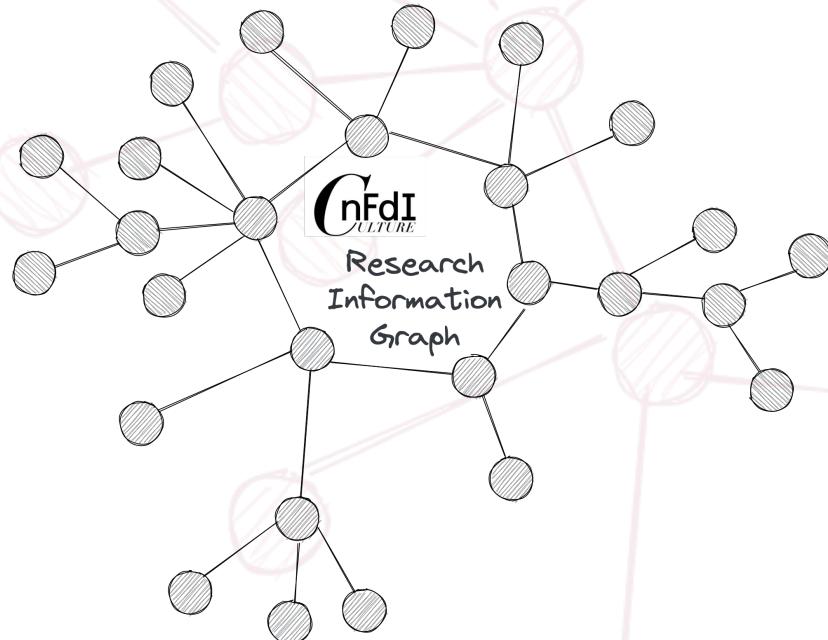
- All of nfdi4culture-KG v1.0
- plus metadata for nfdi4culture content as e.g., from
  - Fully fledged LOD resources
  - Structured Data resources (Legacy Data resources)
  - Wikibase-based resources

## PURPOSE:

- Metadata for retrieval, exploration, and analysis of nfdi4Culture research data resources



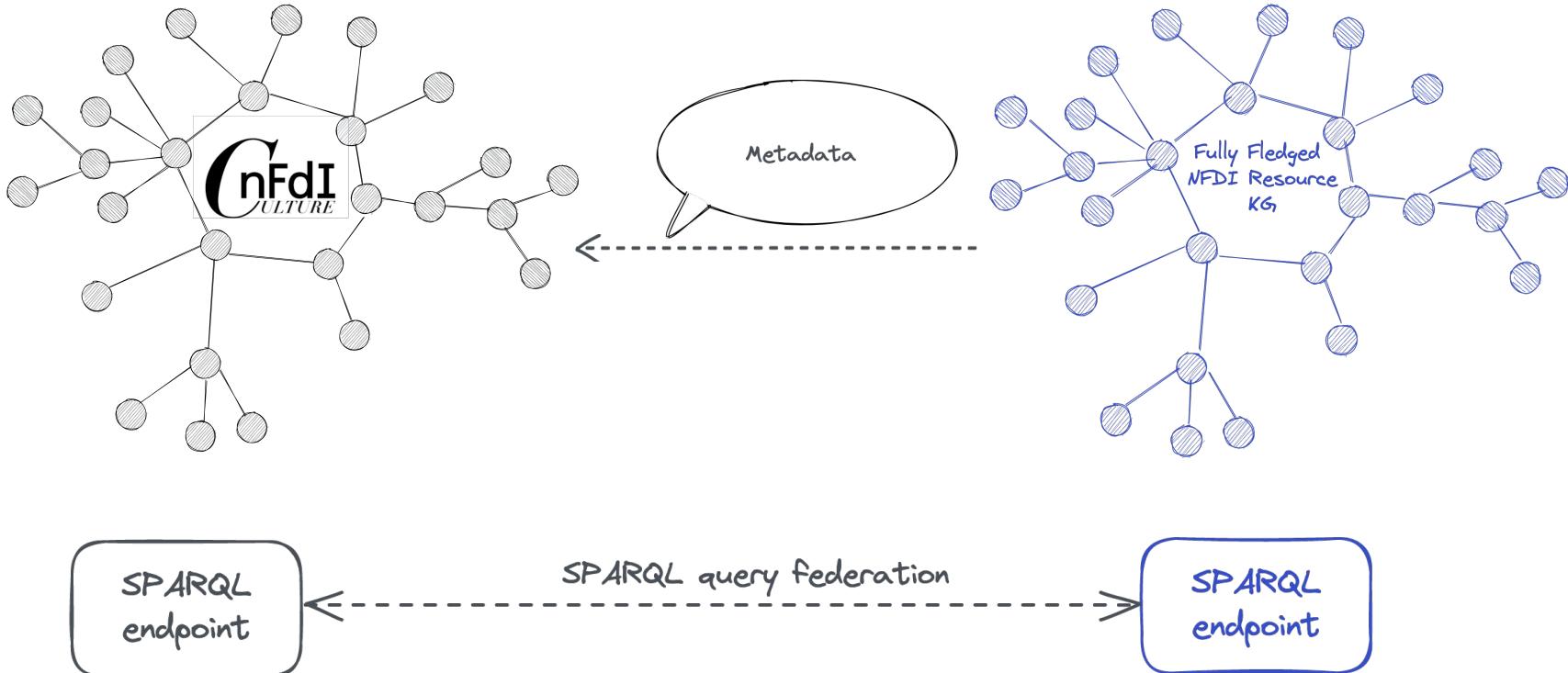
# nfdi4culture-KG v1.0 - Research Data Graph



**Research Information Graph + Research Data Graph = Culture KG v2.0**

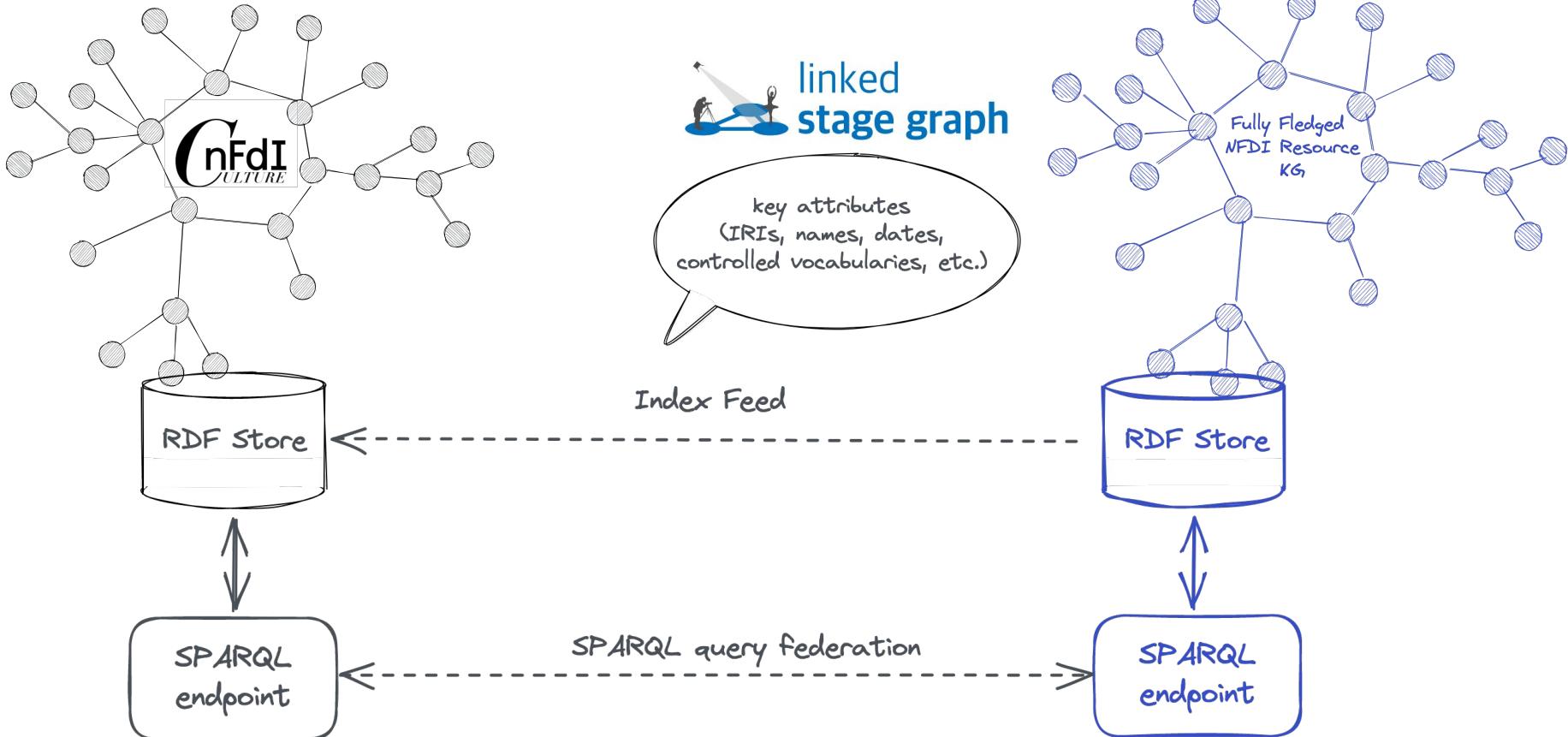
# nfdi4Culture-KG Content Integration

## (1) Fully Fledged LOD



# nfdi4Culture-KG Content Integration

(1) Fully Fledged LOD with SPARQL Endpoint - Use Case Linked Stage Graph



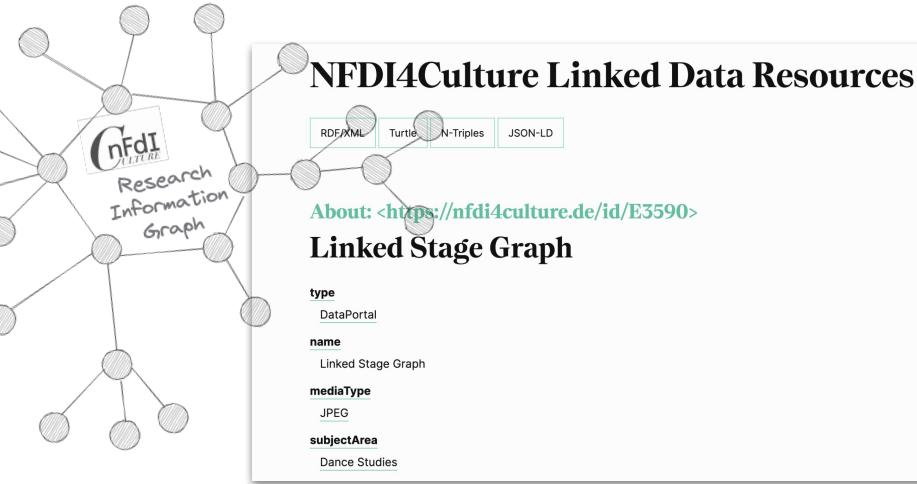
# nfdi4Culture-KG Content Integration

(1) Fully Fledged LOD with SPARQL Endpoint - Use Case Linked Stage Graph



- **Federation** of external RDF-based resources with nfdi4Culture-KG via **native SPARQL**
- Ingest of a **Metadata Index** from external RDF-based resources into nfdi4Culture-KG via lightweight **Culture Graph Interchange Format (CGIF)**
- Index enables **cross connection and discovery** of distributed data resources
- Query Support via nfdi4Culture Portal

# Linked Stage Graph @ nfdi4culture KG



```
PREFIX rdfs: <http://www.w3.org/2000/01/rdf-schema#>
PREFIX schema: <http://schema.org/>
PREFIX slod: <http://slod.fiz-karlsruhe.de/>
PREFIX foaf: <http://xmlns.com/foaf/0.1/>
```

```
SELECT DISTINCT ?resource ?basedOn (COUNT(?image) AS ?imagecount) WHERE {
  SERVICE <http://slod.fiz-karlsruhe.de/sparql> {
    ?resource schema:isBasedOn ?basedOn .
    ?resource foaf:depiction ?image .
  }
} GROUP BY ?resource ?basedOn
```

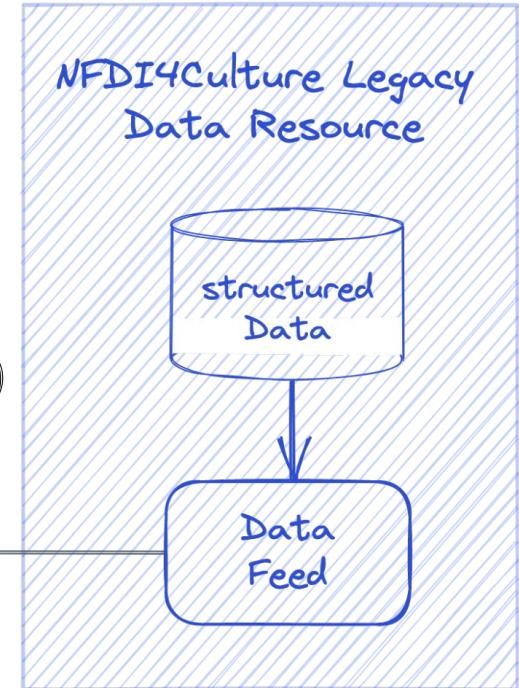
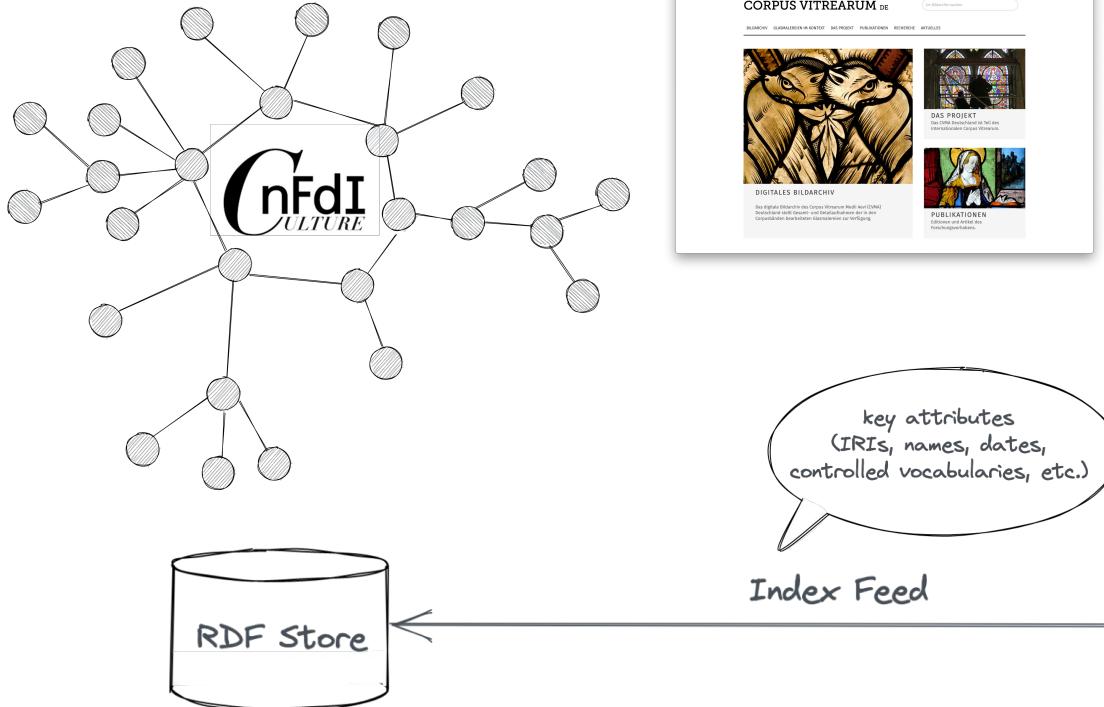
## SPARQL Query Federation:

Show me all performances in Linked Stage Graph data portal, the original work they are based on and show me how many photographs are available for each performance.

resource	basedOn	imagecount
1 slod:labw-2-2601070	<http://www.wikidata.org/entity/Q1774386>	"72"^^<http://www.w3.org/2001/XMLSchema#integer>
2 slod:labw-2-2601341	<http://d-nb.info/gnd/4099338-3>	"56"^^<http://www.w3.org/2001/XMLSchema#integer>
3 slod:labw-2-2601341	<http://www.wikidata.org/entity/Q425837>	"56"^^<http://www.w3.org/2001/XMLSchema#integer>
4 slod:labw-2-2602552	<http://d-nb.info/gnd/300156405>	"56"^^<http://www.w3.org/2001/XMLSchema#integer>
5 slod:labw-2-2602552	<http://www.wikidata.org/entity/Q1165496>	"56"^^<http://www.w3.org/2001/XMLSchema#integer>
6 slod:labw-2-2601340	<http://d-nb.info/gnd/4099338-3>	"52"^^<http://www.w3.org/2001/XMLSchema#integer>
7 slod:labw-2-2601340	<http://www.wikidata.org/entity/Q425837>	"52"^^<http://www.w3.org/2001/XMLSchema#integer>
8 slod:labw-2-2601353	<http://d-nb.info/gnd/300162677>	"52"^^<http://www.w3.org/2001/XMLSchema#integer>
9 slod:labw-2-2601353	<http://www.wikidata.org/entity/Q657595>	"52"^^<http://www.w3.org/2001/XMLSchema#integer>
10 slod:labw-2-2599395	<http://d-nb.info/gnd/300132514>	"48"^^<http://www.w3.org/2001/XMLSchema#integer>

# nfdi4Culture-KG Content Integration

## (2) Structured Data (Web-based) - Use Case: Corpus Vitrearum



# Data Integration with CGIF



Culture Graph Interchange Format Specification

Home  
Introduction  
Specification  
Implementation

Search

NFDI4Culture Knowledge Base  
NFDI4Culture Portal

Gefördert durch die Deutsche Forschungsgemeinschaft (DFG)  
Projektnummer 441958017

Impressum Datenschutz

Page last updated: 2023-09-11 10:15:00

Edit this Page

NFDI4Culture Specification

## Culture Graph Interchange Format

Many potential contributors in the domain of NFDI4Culture provide websites or online databases, but dedicated REST APIs or SPARQL endpoints are rare. For the federated acquisition of data for the *Culture Knowledge Graph* a **lightweight interchange format** for harvesting of resources from data collections with key attributes (IRIs, names, dates, terms from controlled vocabularies) needed to be defined. Getting data into the *Culture Knowledge Graph* is as easy as adding a bit of markup to a website and submitting the entry-point URL. By applying schema.org properties, the *Culture Graph Interchange Format* (CGIF) has the added benefit of automatically making the data eligible for Google Dataset Search and other search providers, including image searches. The Google bot and other web crawlers will recognise the markup as soon as they next visit any website embedding the CGIF.

### Autor:innen

Sasha Bruns  
Etienne Posthumus  
Harald Sack  
Linnaea Söhn  
Torsten Schrade  
Jonatan Jalle Steller  
Tabea Tietz  
Julia Tolksdorf

### Lizenz, Identifier, Version

Open Educational Resource, CC BY 4.0  
Permalink: [LINK](#)  
Metadaten: [LINK](#)  
Datenpublikation: [LINK](#)  
Version: DRAFT

Lightweight format based on Schema.org for harvesting resources with **key attributes** (IRIs, names, dates, (key)terms)

**No API required** - integration into the Culture KG just by adding a bit of markup to a website

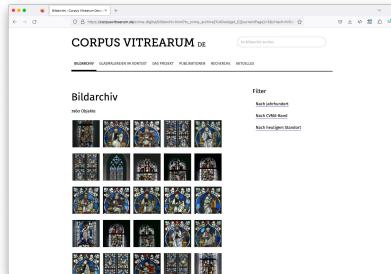
Added benefit of automatically making the website / data eligible for **Google Dataset Search**

Mapping for **Fair Digital Objects**

Low effort implementation for **legacy data resources**

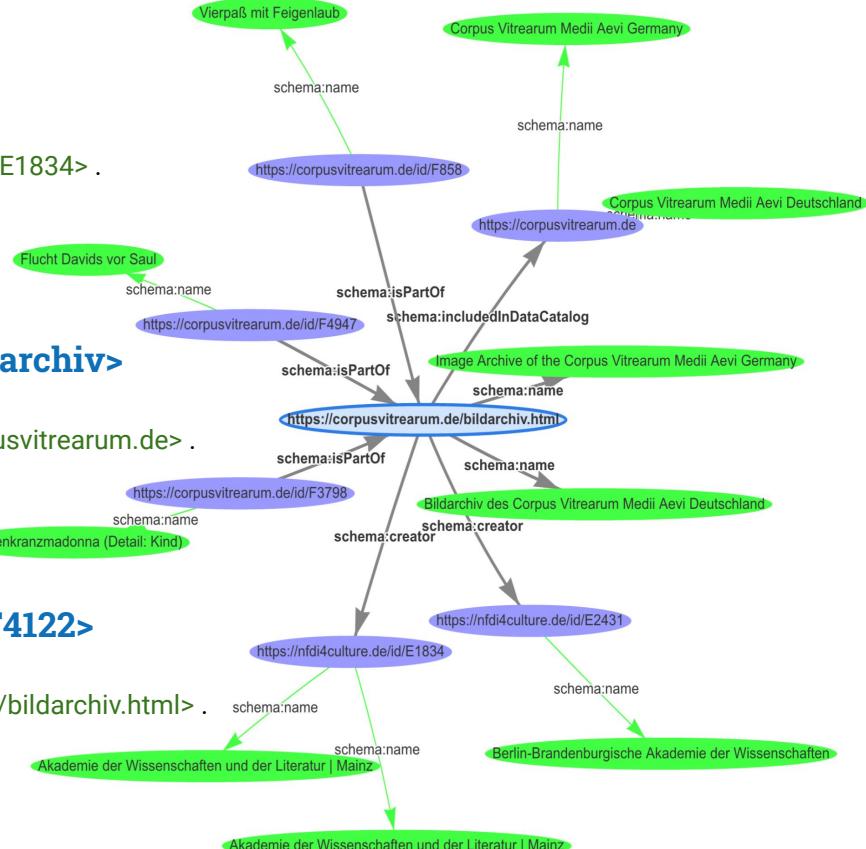
Version 1.0: <https://nfdi4culture.de/go/E3712>

# Application of CGIF Model in nfdi4culture



<https://corpusvitrearum.de>

a schema:DataCatalogue ;  
schema:publisher <<https://nfdi4culture.de/id/E1834>> .



<https://corpusvitrearum.de/bildarchiv>

a schema:Dataset, schema:Datafeed ;  
schema:includedInDataCatalog <<https://corpusvitrearum.de>> .

<https://corpusvitrearum.de/id/F4122>

a schema:ImageObject ;  
schema:isPartOf <<https://corpusvitrearum.de/bildarchiv.html>> .

# CVMA: Data Federation via Culture KG

Find depictions with ICONCLASS "wild man" in Culture KG, Wikidata and Rijksmuseum

```
SELECT ?resource ?resourceLabel ?dataCatalogue
WHERE {
  {
    SELECT ?resource ?resourceLabel ?dataCatalogue WHERE {
      ?resource schema:keywords iconclass:31A44411 .
      ?resource schema:name ?resourceLabel .
      BIND("Corpus Vitrearium / NFDI4Culture" AS ?dataCatalogue) .
    } LIMIT 2
  } UNION {
    SERVICE <https://query.wikidata.org/sparql> {
      SELECT ?resource ?resourceLabel ?dataCatalogue WHERE {
        ?resource wdt:P1257 "31A44411" .
        BIND("Wikidata" AS ?dataCatalogue) .
        SERVICE wikibase:label { bd:serviceParam wikibase:language "en" . }
      } LIMIT 2
    } UNION {
      SERVICE <https://api.data.netwerkdigitaalergoed.nl/datasets/Rijksmuseum>
      SELECT ?resource ?resourceLabel ?dataCatalogue WHERE {
        ?resource dc:subject <http://iconclass.org/31A44411> .
        ?resource dc:title ?resourceLabel .
        BIND("Rijksmuseum Amsterdam" AS ?dataCatalogue) .
      } LIMIT 2
    }
  }
}
```

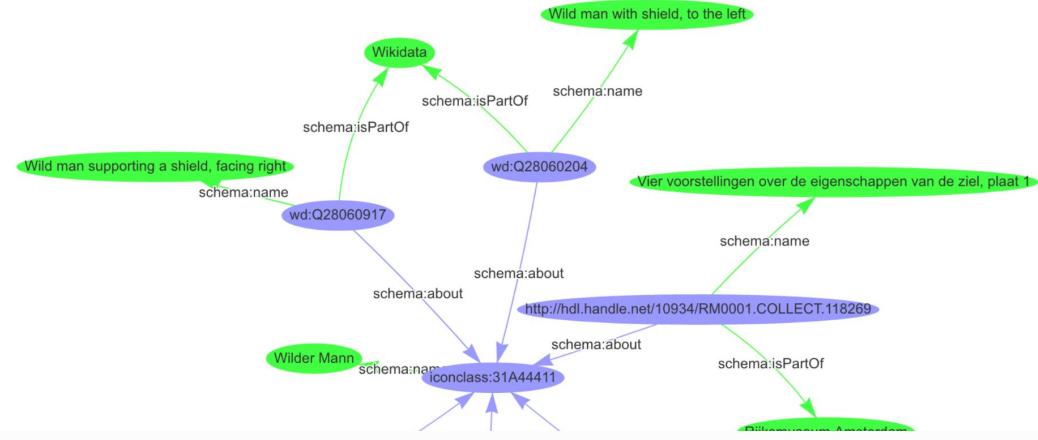


Table Response 6 results in 4.654 seconds

Simple view  Ellipse  Filter query results Page size: 50

resource	resourceLabel	dataCatalogue
1 < <a href="http://hdl.handle.net/10934/RM0001.COLLECT.118269">http://hdl.handle.net/10934/RM0001.COLLECT.118269</a> >	"Vier voorstellingen over de eigenschappen van de ziel, plaat 1"	"Rijksmuseum Amsterdam"^^< <a href="http://www.w3.org/2001/XMLSchema#string">http://www.w3.org/2001/XMLSchema#string</a> >
2 < <a href="http://hdl.handle.net/10934/RM0001.COLLECT.118269">http://hdl.handle.net/10934/RM0001.COLLECT.118269</a> >	"Peter Wild Boy"@nl	"Rijksmuseum Amsterdam"^^< <a href="http://www.w3.org/2001/XMLSchema#string">http://www.w3.org/2001/XMLSchema#string</a> >
3 wd:Q28060204	"Wild man with shield, to the left"@en	"Wikidata"^^< <a href="http://www.w3.org/2001/XMLSchema#string">http://www.w3.org/2001/XMLSchema#string</a> >
4 wd:Q28060917	"Wild man supporting a shield, facing right"@en	"Wikidata"^^< <a href="http://www.w3.org/2001/XMLSchema#string">http://www.w3.org/2001/XMLSchema#string</a> >
5 < <a href="https://corpusvitrearum.de/id/F13046">https://corpusvitrearum.de/id/F13046</a> >	"Wappen Blume/Lamm mit wilden Leuten als Wappen..."	"Corpus Vitrearium / NFDI4Culture"^^< <a href="http://www.w3.org/2001/XMLSchema#string">http://www.w3.org/2001/XMLSchema#string</a> >
6 < <a href="https://corpusvitrearum.de/id/F4122">https://corpusvitrearum.de/id/F4122</a> >	"Wilder Mann und Wappen mit Hausmarke"@de	"Corpus Vitrearium / NFDI4Culture"^^< <a href="http://www.w3.org/2001/XMLSchema#string">http://www.w3.org/2001/XMLSchema#string</a> >

Showing 1 to 6 of 6 entries

< 1 >

Try it out: <https://nfdi4culture.de/go/kg-query-cvma>

# CVMA: Enhanced SEO / Google Dataset Search



The screenshot shows the Google Dataset Search interface with the query "medieval stained glass". It displays several search results, each with a thumbnail, title, and a snippet of information. One result is from the "Bildarchiv des Corpus Vitrearum Medii Aevi Deutschland" (Image Archive of the Corpus Vitrearum Mediævæ Germany), which is highlighted in red.

- Bildarchiv des Corpus Vitrearum Mediævæ Deutschland  
Image Archive of the Corpus Vitrearum Mediævæ Germany
- Cathedral of Beauvais  
openheritagefdi.org
- Sie sehen nicht das erwartete Ergebnis?  
Hier erfahren Sie, wie Sie unserem Index neue Datensätze hinzufügen können.

The screenshot shows a Google search results page for the query "ulm münster glasmalerei". The top result is a link to a page titled "Glasmalereien im Kontext: Corpus Vitrearum Deu...". Below the link are several thumbnail images of medieval stained glass windows, all of which are labeled "Alamy". To the right of the thumbnails, there is a sidebar with the text "Read the use case report and CGIF integration tutorial:" followed by a link: <https://nfdi4culture.de/go/cgif-cvma>.

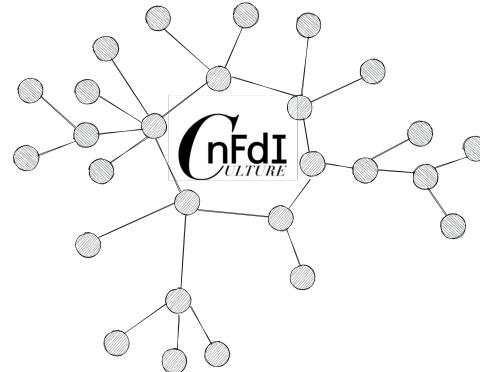
Integration of CGIF has significantly improved findability in Google dataset & image search

Read the use case report and CGIF integration tutorial:

<https://nfdi4culture.de/go/cgif-cvma>

# nfdi4Culture-KG Content Integration

## (3) Further Variants - e.g., Wikibase-based repositories



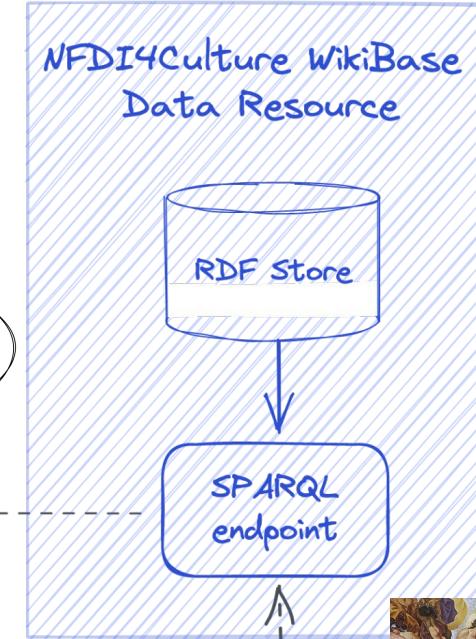
key attributes  
(IRIs, names, dates,  
controlled vocabularies, etc.)

Index Feed

SPARQL query federation

SPARQL  
endpoint

semantic  
Mapping



Use Case 3: (planned)  
Corpus der barocken  
Deckenmalerei in  
Deutschland (CbDD)

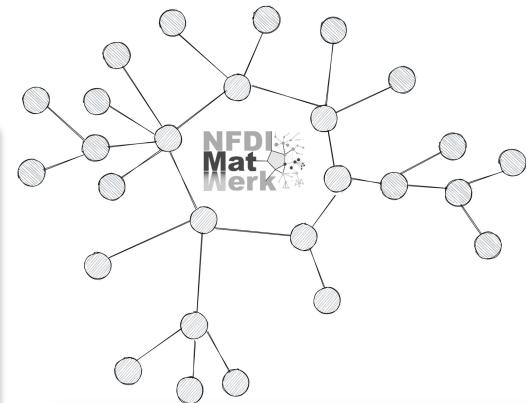
## Entity Resolution via LodView and SPARQL endpoint

LodView

Prof. Dr. Harald Sack  
<http://demo.fiz-karlsruhe.de/matwerk/E5518>

<<http://nfdi.fiz-karlsruhe.de/ontology/Person>>

rdfs:label	Prof. Dr. Harald Sack
foaf:surname	Sack
mwo:emailAddress	harald.sack@fiz-karlsruhe.de
mwo:hasORCID	<a href="https://orcid.org/0000-0001-7069-9804">https://orcid.org/0000-0001-7069-9804</a>
foaf:firstName	Harald
foaf:title	Prof. Dr.
rdf:type	< <a href="http://nfdi.fiz-karlsruhe.de/ontology/Person">http://nfdi.fiz-karlsruhe.de/ontology/Person</a> >
mwo:hasAffiliation	< <a href="http://demo.fiz-karlsruhe.de/matwerk/E1011">http://demo.fiz-karlsruhe.de/matwerk/E1011</a> > ↳ Leibniz-Institut für Informationsinfrastruktur
mwo:hasExpertiseIn	< <a href="http://demo.fiz-karlsruhe.de/matwerk/E14269">http://demo.fiz-karlsruhe.de/matwerk/E14269</a> > ↳ Information Service Engineering



Entity Resolution: <https://demo.fiz-karlsruhe.de/matwerk/E5518.html> and SPARQL Endpoint: <https://demo.fiz-karlsruhe.de/sparql>

### 3. nfdi4Culture Portal

# Culture Portal: “Things not Strings”



The figure displays four screenshots of the NFDI4Culture website, illustrating its structure and content:

- Communities**: A page featuring a large black and white photograph of architectural structures. It includes a section on "Architecture" and "Art History".
- Task Area 5 – Technical, Ethical & Legal**: A page titled "Overarching technical, ethical and legal activities" featuring two portraits of people.
- Resources**: A page titled "Finding Aids" showing a server room and a computer lab. It includes sections for "Curated Repository List" and "Software & Dataservice Registry".
- Knowledge Base**: A page titled "Knowledge Base" with a large diagram of concentric hexagons and a list of seven numbered items: 1. Capture & Enrichment, 2. Standards, Quality & Curation, 3. Tools & Services, 4. Publication & Availability, 5. Technical, Ethical & Legal, 6. Qualification & Training, and 7. Governance & Administration.

<https://nfdi4culture.de/>

# Culture Portal: “Things not Strings”



[<https://nfdi4culture.de/id/E1925>](https://nfdi4culture.de/id/E1925)

The screenshot shows a web browser window for the NFDI4Culture portal. The URL is https://nfdi4culture.de/resource/E1925/about.html. The page title is "About: <https://nfdi4culture.de/id/E1925>". The main content area displays a person named Harald Sack. The "type" field is listed as "Person". Below it is a thumbnail image of a man with glasses and a beard. The "memberOf" field shows he is a member of the "Culture Spokespersons Committee" and holds the role "Member".

[<https://nfdi4culture.de/id/E1825>](https://nfdi4culture.de/id/E1825)

The screenshot shows a web browser window for the NFDI4Culture portal. The URL is https://nfdi4culture.de/resource/E1825/about.html. The page title is "About: <https://nfdi4culture.de/id/E1825>". The main content area displays an organization named "Overarching technical, ethical and legal activities". The "type" field is listed as "Organization".

The screenshot shows a web browser window for the NFDI4Culture portal. The URL is https://nfdi4culture.de/resource/E2391/about.html?\_id\_api=query#knowledge. The page title is "About: <https://nfdi4culture.de/id/E2391>". The main content area displays the "Culture Knowledge Graph". The "type" field is listed as "Service". The "name" field is "Culture Knowledge Graph". The "identifier" field is "TASM3". The "description" field contains a detailed paragraph about the purpose and development of the knowledge graph. The "keywords" field is "Linked Data & Semantics".

[<https://nfdi4culture.de/id/E2391>](https://nfdi4culture.de/id/E2391)

The screenshot shows a web browser window for the NFDI4Culture portal. The URL is https://nfdi4culture.de/resource/E3625/about.html?\_id\_api=query#FAIR. The page title is "About: <https://nfdi4culture.de/id/E3625>". The main content area displays a guideline titled "Guideline for a FAIR Cultural Studies Research Data Management". The "type" field is listed as "Contribution Guide". The "name" field is "Guideline for a FAIR Cultural Studies Research Data Management". The "url" field is "https://nfdi4culture.de/go/E3625". The "version" field is "1.0.1". The "description" field contains a paragraph explaining the principles of FAIR data management. The "keywords" field is "Digital Humanities".

[<https://nfdi4culture.de/id/E3625>](https://nfdi4culture.de/id/E3625)

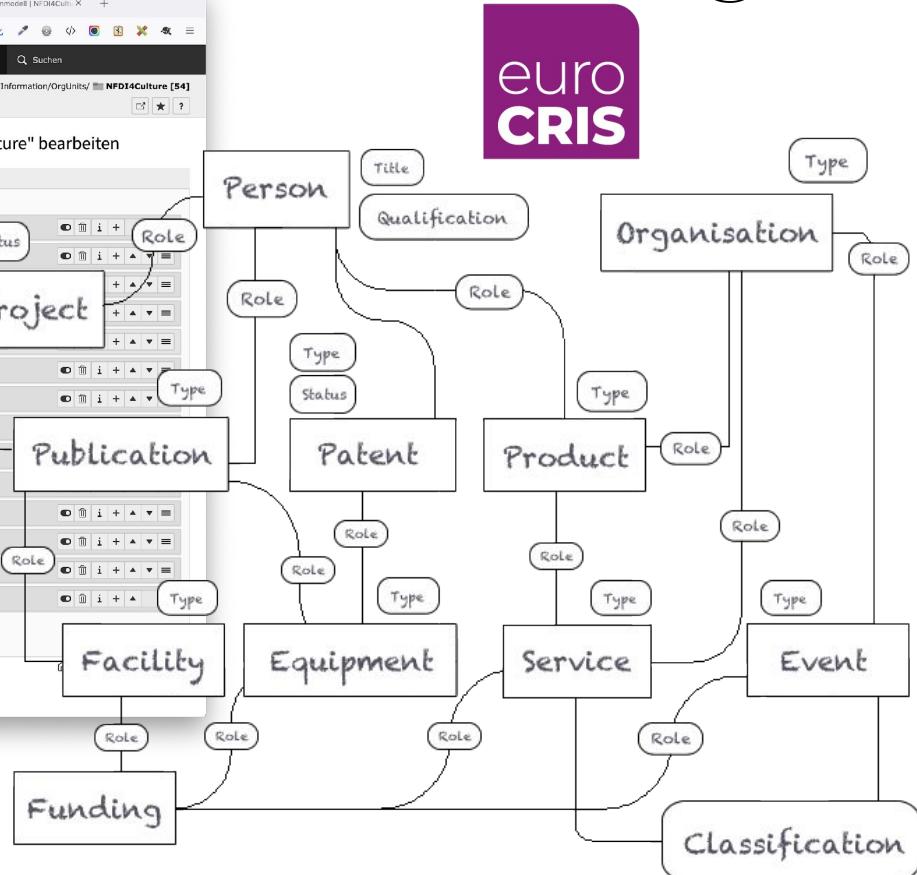
# Culture Research Information System



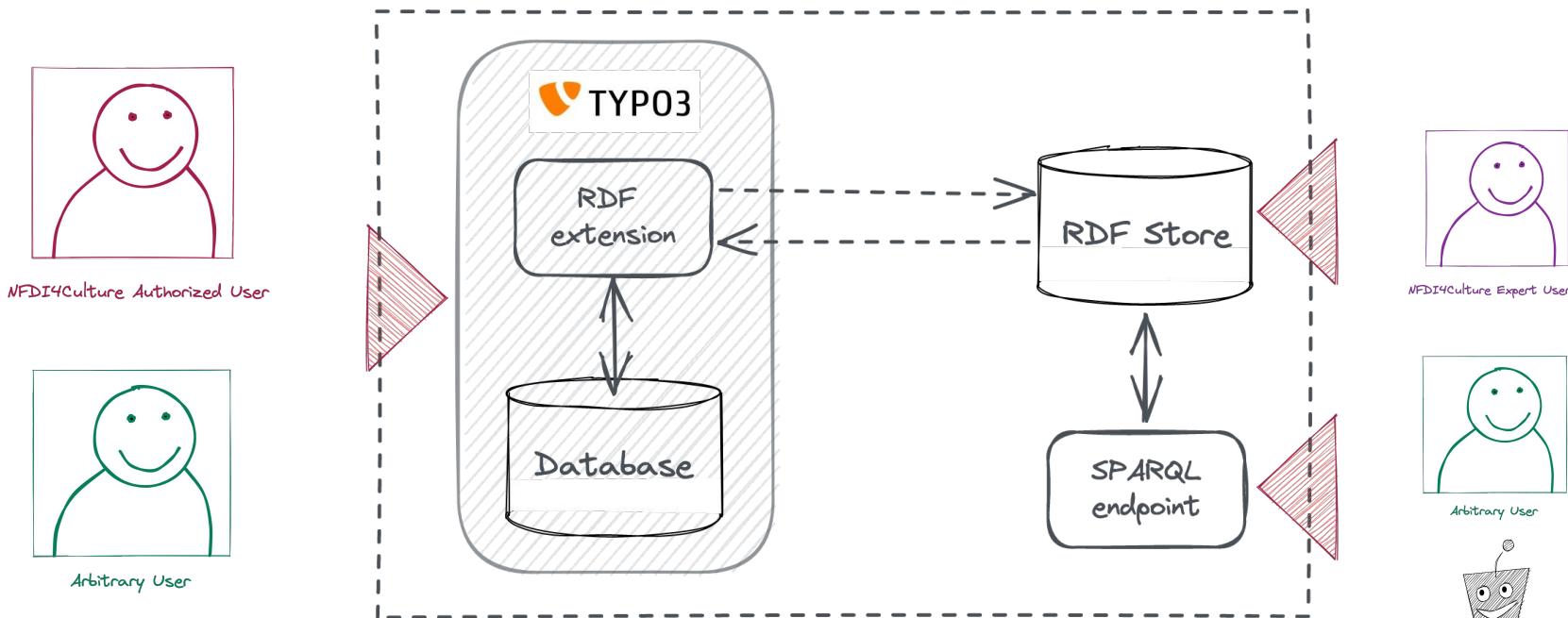
Screenshot of the TYPO3 CMS interface for the nFDI4Culture extension. The left sidebar shows a tree structure of research information categories. The main content area displays a list of organization units under "Standards, data quality and curation". Each entry includes a "Details" icon, a "Delete" icon, and a "New" button. The list contains entries such as "Co-Spokesperson: Prof. Dr. Barbara Wiermann, Standards, data quality and curation" and "Member: Dr. Melanie Gruß, Standards, data quality and curation".

TYPO3 CRIS extension on GitHub:

<https://github.com/digicademy/academy>



# Culture Portal: Linked Data Architecture



TYPO3 LOD extension on GitHub:

<https://github.com/digicademy/lod>

Automated Services

# SPARQL Endpoint Data Exploration



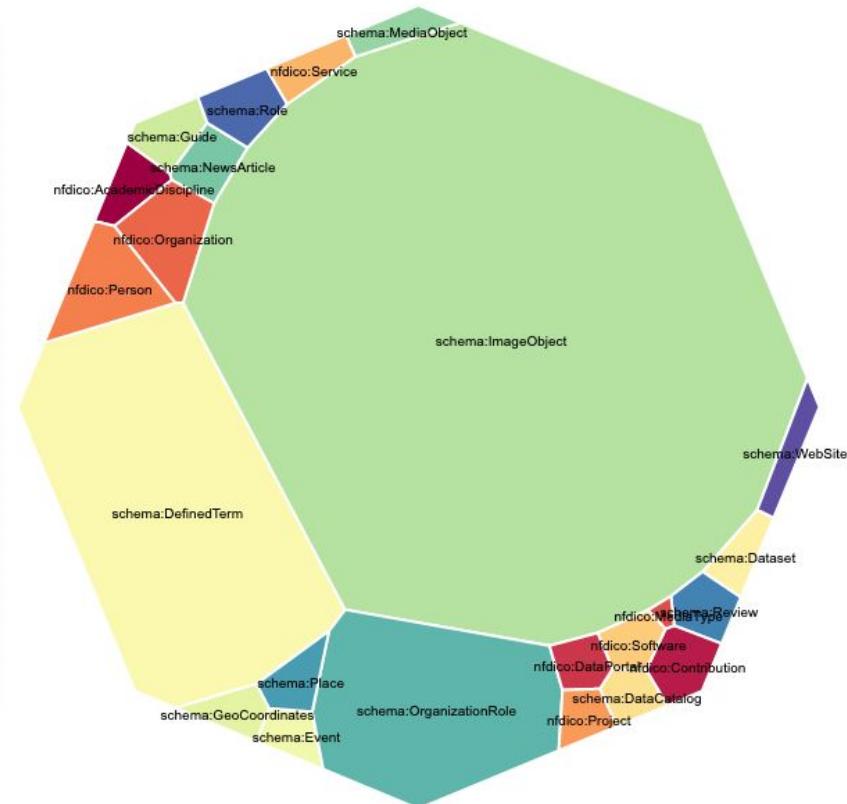
Screenshot of the NFDI4Culture Knowledge Graph interface. The page title is "NFDI4Culture Knowledge Graph". Below the title, a paragraph describes the Culture Knowledge Graph as a connector for research data collections, software tools, infrastructures, and services within the NFDI4Culture subject areas. It integrates resources from partners, research communities, and beyond, and allows querying through a SPARQL interface.

Key features shown include:

- Navigation menu: What we do, News & Events, About us, Resources (highlighted), Contact, DE.
- Search bar.
- Four buttons at the top: Count triples, Show all types, Count instances per type, Show properties for types.
- A query editor with tabs for Query 1, Query 2, Query 3, Query 4, Query 5, Query 6, and Query 7. The current tab is Query 4, which contains the following SPARQL code:

```
PREFIX rdf: <http://www.w3.org/1999/02/22-rdf-syntax-ns#>
PREFIX rdfs: <http://www.w3.org/2000/01/rdf-schema#>
PREFIX schema: <http://schema.org/>
PREFIX nfdico: <https://nfdi4culture.de/ontology#>
SELECT ?type (COUNT (?s) AS ?instanceCount)
WHERE {
    ?s a ?type .
}
GROUP BY ?type
ORDER BY ?type
```
- Output area showing results of the query.
- Page footer with links for Table, Response, Simple view, Ellipse, Filter query results, Page size: 50, and a help icon.

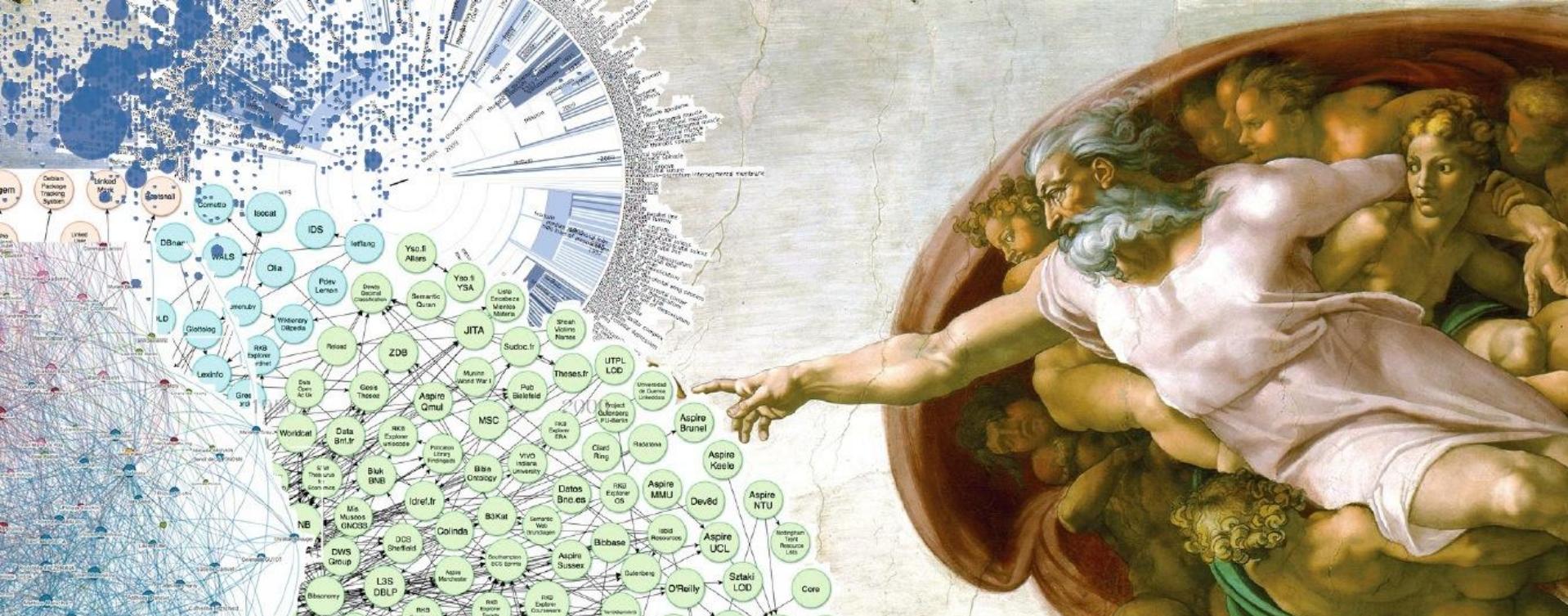
<https://nfdi4culture.de/resources/knowledge-graph.html>



## 4. Next Steps

# Next Steps

- Enable **interactive KG extension & registration/curation of datasets** via Culture Portal
- Further extension(s) of NFDIcore (4Memory, MaRDI, ...) (what about NFDI4DataScience?)
- Extension of nfdi4culture Research Data KG for **WissKI** (CIDOC-CRM-based CMS) based data repositories and **further legacy data variants**
- General **adoption/implementation of CGIF** by nfdi4culture partners
- Federation with Wikidata, DDB, Europeana and other NFDI KGs



CnFdi  
CULTURE

NFDI  
Mat  
Werk

Thank you very much  
for your attention!

 FIZ Karlsruhe  
Leibniz-Institut für Informationsinfrastruktur

Mitglied der  
  
Leibniz  
Gemeinschaft

 Akademie der  
Wissenschaften  
und der Literatur  
Mainz

 JÜLICH  
Forschungszentrum