

Semantic x-Lab Kick-Off

June 11, 2025



Knowledge Graph Infrastructure

Thomas Bauer - Service Steward



checkout last given KGI talk on Zenodo:
[10.5281/zenodo.15487695](https://zenodo.org/record/15487695/files/10.5281/zenodo.15487695)

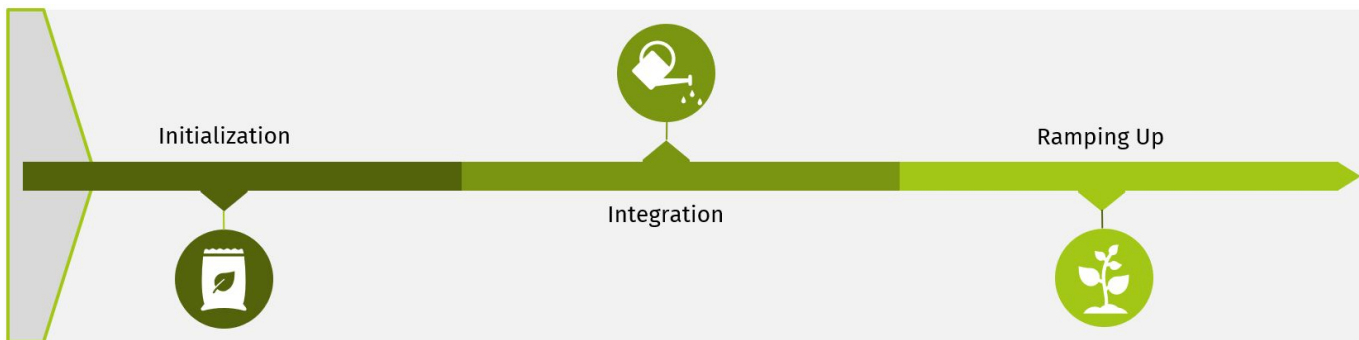
Structure of the talk

- Introduction
- Use cases
- Initialisation vs. Integration
- Getting involved

Introduction

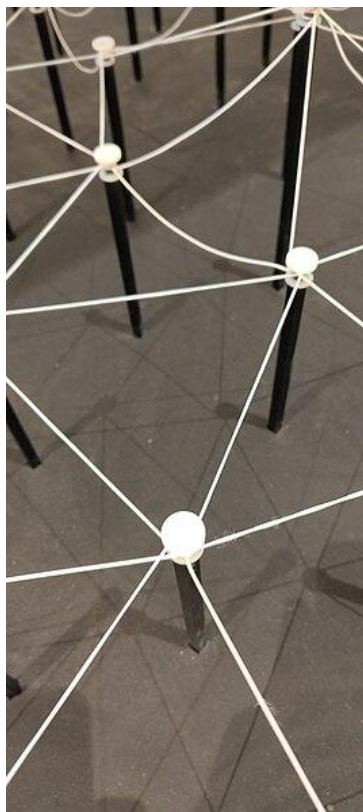
KGI4NFDI becomes a Basic Service in Base4NFDI

- Base4NFDI funds Basic Services that are deemed useful across multiple NFDI consortia
- KGI4NFDI - **K**nowledge **G**raph **I**nfrastructure
 - provides infrastructure and services around knowledge graphs (KG)
 - is currently in its Initialization Phase (extended until Dec 2025)
 - More info at: <https://kgi.services.base4nfdi.de/> and <https://base4nfdi.de/projects/kgi4nfdi>



Introduction

Why KGs and why KGI



KGs are an important technology supporting **interoperability** and enabling **data exchange**:

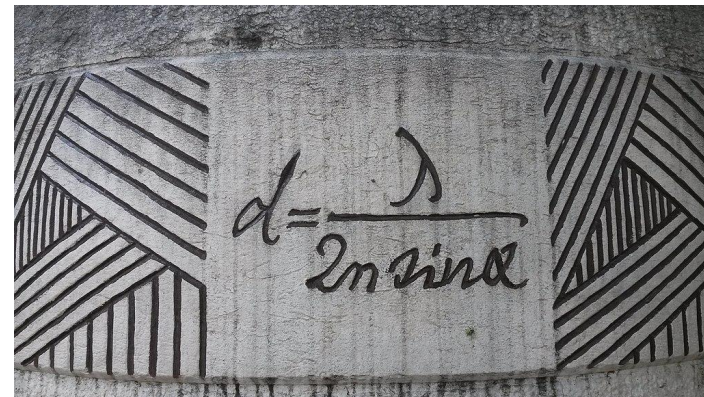
- A KG is a **graph-structured knowledge base** containing a terminology (vocabulary or ontology) and data entities interrelated via the terminology;
- KGs are based on **semantic web technologies** (RDF, SPARQL, etc.) and are often used for agile data integration;
- KGs are already **widely used** by research data producers and managers in Germany.

Introduction

Basic idea of Knowledge Graphs: link individual pieces of knowledge

- Which pieces of knowledge can be linked to and from this image, and how?

- It's a photograph of (a part of) a monument.
- The monument depicts Latin and Greek letters.
- The letters are part of a mathematical formula.
- The formula represents some physical concepts.
- The physical concepts describe a limit.
- The limit could be circumvented by engineering.
- The engineering got the Nobel Prize in Chemistry.
- The chemistry is widely used in biology and medicine.
- All of this has a history as well as material, social and data aspects, e.g.
 - the monument shares its pedestal with an earlier one
 - long before receiving the Nobel Prize, the laureates
 - faced resistance because the limit was “set in stone”
 - had been referring to the monument in their talks
 - journalists reused this image in their coverage



Introduction

Key ingredients for a Knowledge Graph

- Scope
 - use cases
 - types of nodes and edges
- Data
 - data sources
 - data quality
 - data access
 - data licensing
- Metadata
 - schemas for KG entries
 - domain knowledge
 - provenance
 - mapping to external vocabularies
- Storage, Security & Performance
- Workflows
 - importing
 - data discovery
 - entity & relation extraction / mapping
 - data integration
 - curating
 - duplicate detection / disambiguation
 - updates
 - exporting
 - entity and relation mapping
 - exploring
 - browsing / searching / querying / visualizations
- ...

Introduction

KGs in NFDI

Humanities and social sciences

- BERD@NFDI (KGs)
- KonsortSWD
- NFDI4Culture (KGs)
- NFDI4Memory (KGs)
- NFDI4Objects (KGs)
- Text+ (KG)

Engineering sciences

- NFDI4DataScience (KGs & KG Software)
- NFDI4Energy (KG)
- NFDI4Ing (KG Software)
- NFDI-MatWerk (KGs & KG Software)
- NFDIxCs

Life sciences

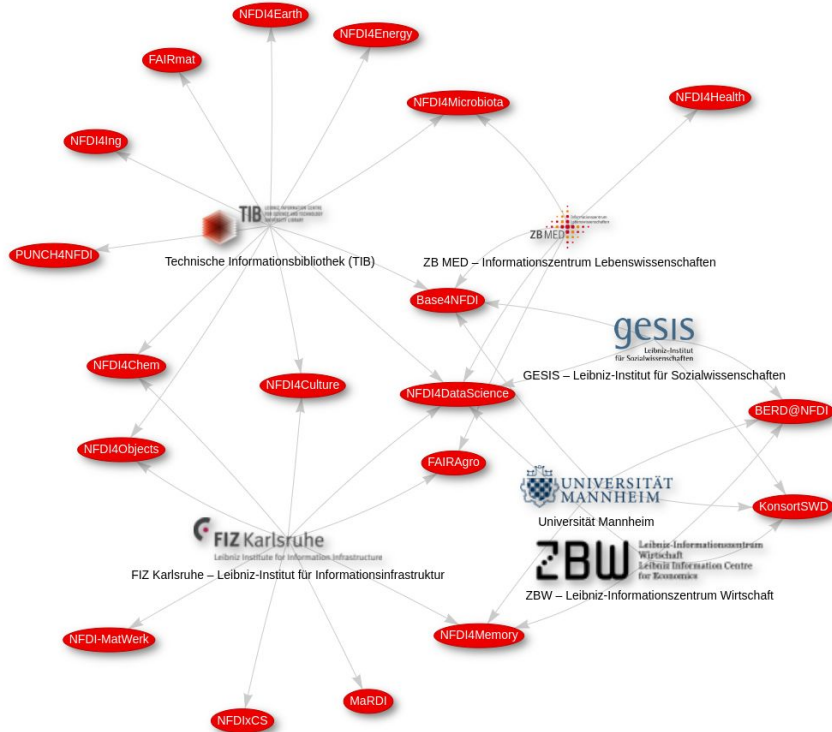
- DataPLANT
- FAIRagro
- NFDI4Immuno
- GHGA
- NFDI4Biodiversity
- NFDI4BIOIMAGE (KG)
- NFDI4Health
- NFDI4Microbiota (KG)

Natural sciences

- DAPHNE4NFDI
- FAIRmat
- NFDI4Cat (KG)
- MaRDI (KGs)
- NFDI4Chem (KGs)
- NFDI4Earth (KG)
- PUNCH4NFDI

KGI4NFDI

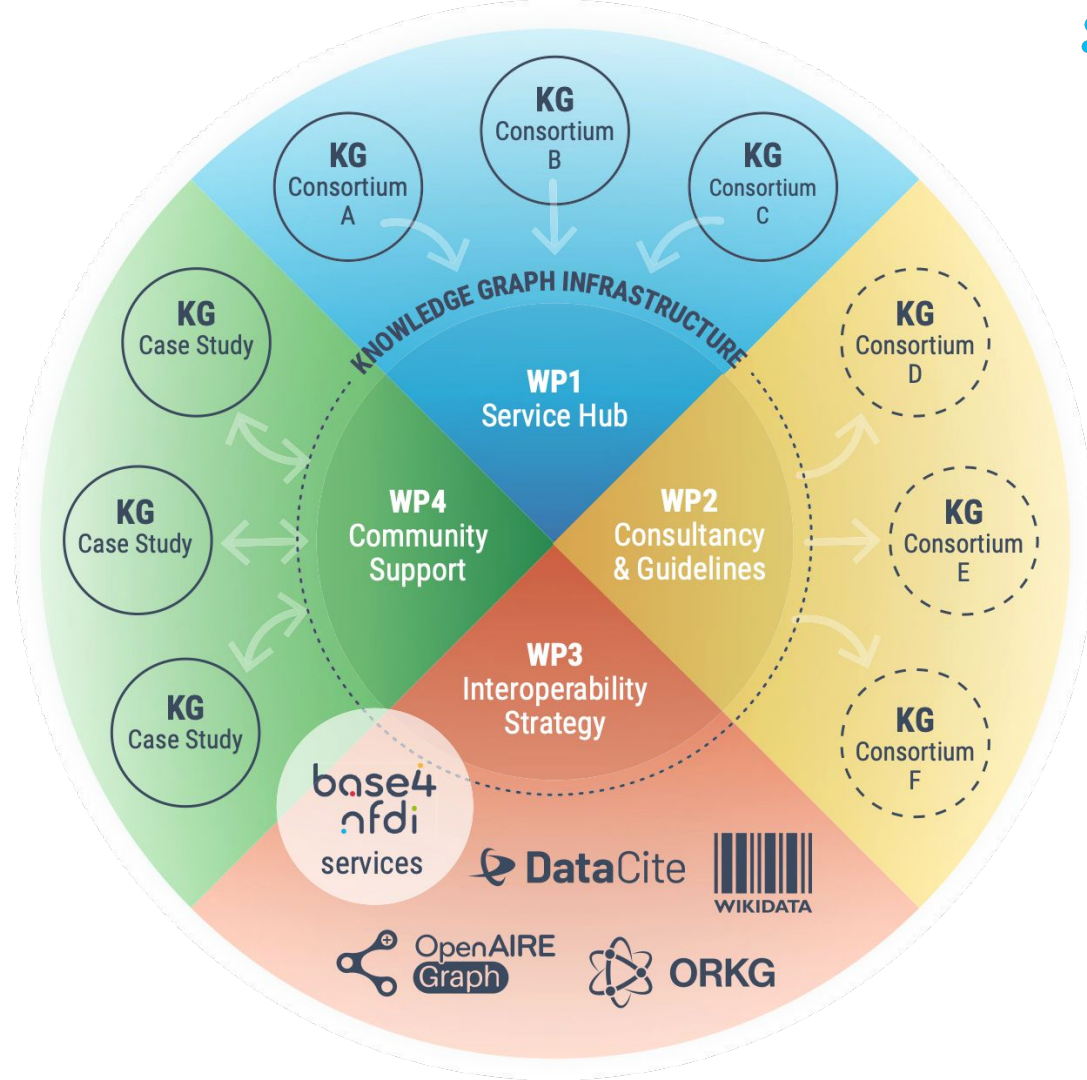
Project Team location



- FIZ Karlsruhe (MaRDI)
- GESIS (NFDI4DataScience)
- TIB (NFDI4Culture)
- Uni Mannheim Library (BERD@NFDI)
- ZB MED (NFDI4Microbiota)
- ZBW (KonsortSWD)

Initialisation Phase

Setting things up



Use Cases

Personas



Aleena

First year Postdoc
Researcher in
Bioinformatics

Aleena works remotely at the University of Cologne. Her scholarly interests are interdisciplinary and she enjoys contributing to citizen science projects like Wikidata. She has some experience using KGs but wants to learn more.



Andreas

Second year PhD
Researcher in
Sociology

Andreas is a PhD student at the University of Mannheim. He is interested in exploring large corpora of available sociological survey data. Andreas is comfortable using standard software for statistics and analysis, but wants to explore more innovative methods particularly in terms of discovery and data reuse.



Alexandra

Senior Data Steward
for Engineering Data
Domain

Alexandra works at a Leibniz Centre and is responsible for data collections in the Engineering domain. She is responsible for integrating the data in her institute to other national and international initiatives and is keen to explore the potential of KGs for this task.



Lisa

Junior Developer in
a research lab

Lisa is an MA graduate in CS, working in a research lab at a Leibniz Institute. She works closely with scientists and enjoys working on services that support innovation in science. She lacks industry experience, but is passionate about open source software and contributes to several projects in volunteer capacity.



Luca

Senior Developer in
a research lab

Luca is semantic web expert working in a research lab at a Fraunhofer Institute. He runs a big KG project and is eager to improve it and see how others use it. He is involved in various expert committees and wants to make knowledge and best practice sharing more effective.

Use Cases

The KGI Service Hub



Aleana

When I

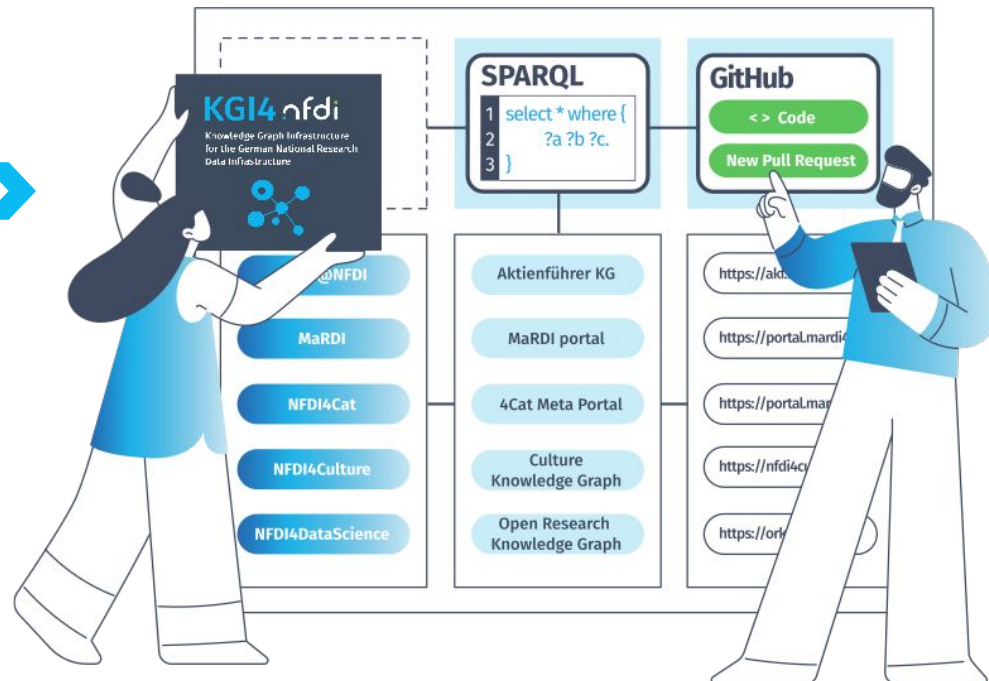
am looking for existing knowledge graphs in the domain of Bioinformatics,

I need / I wish

I need a centralised registry of knowledge graphs.

I am frustrated when

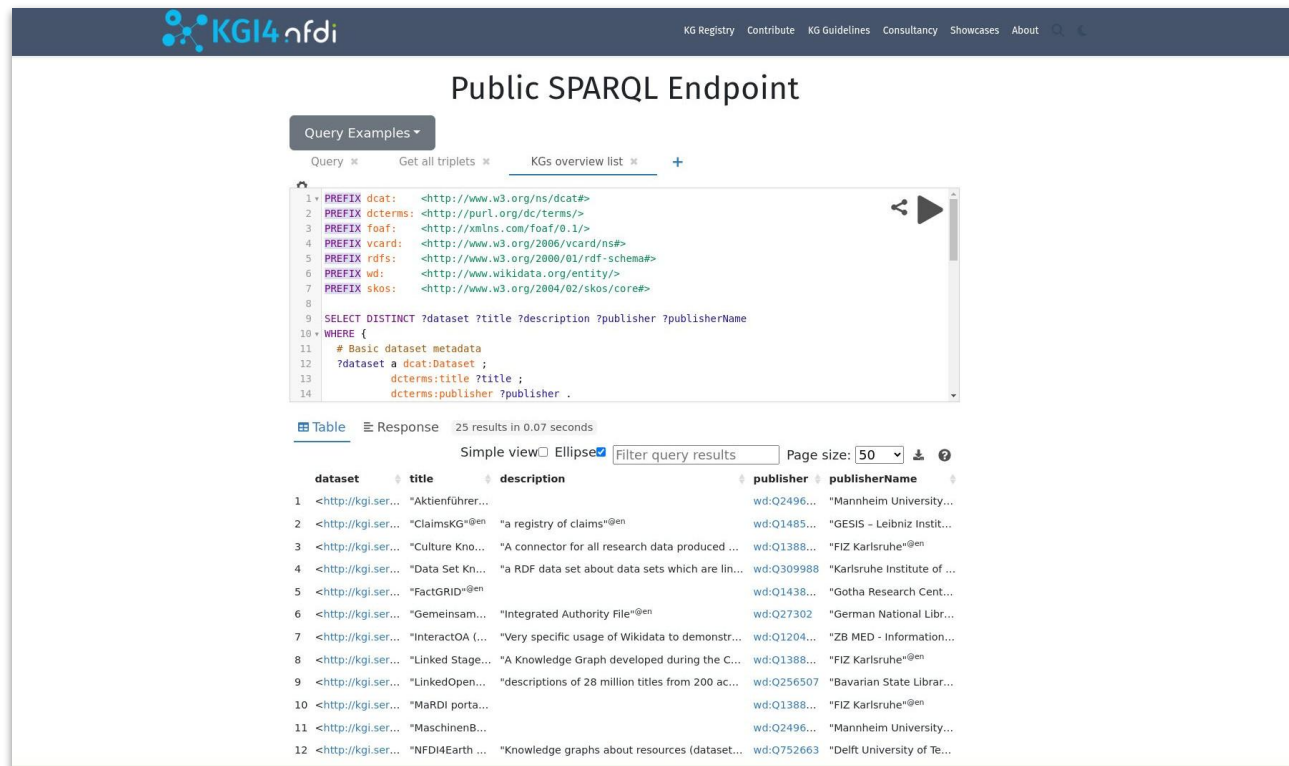
I don't know what question to ask in the query interface and how best to formulate it.



- ★ Registry of KGs in NFDI
- ★ Search and query platform
- ★ Query examples to get started

Use Cases

Example query: List of NFDI KGs registered in the KG Registry



The screenshot shows the KGI4 nfdi Public SPARQL Endpoint interface. The top navigation bar includes links for KG Registry, Contribute, KG Guidelines, Consultancy, Showcases, and About. The main heading is "Public SPARQL Endpoint". Below this, there's a "Query Examples" dropdown menu. The active tab is "KGs overview list". The query editor shows a SPARQL query that selects distinct datasets with their titles, descriptions, and publishers. The query is as follows:

```

1 PREFIX dcat: <http://www.w3.org/ns/dcat#>
2 PREFIX dcterms: <http://purl.org/dc/terms/>
3 PREFIX foaf: <http://xmlns.com/foaf/0.1/>
4 PREFIX vcard: <http://www.w3.org/2006/vcard/ns#>
5 PREFIX rdfs: <http://www.w3.org/2000/01/rdf-schema#>
6 PREFIX wd: <http://www.wikidata.org/entity/>
7 PREFIX skos: <http://www.w3.org/2004/02/skos/core#>
8
9 SELECT DISTINCT ?dataset ?title ?description ?publisher ?publisherName
10 WHERE {
11   # Basic dataset metadata
12   ?dataset a dcat:Dataset ;
13           dcterms:title ?title ;
14           dcterms:publisher ?publisher .

```

The results are displayed in a table format, showing 25 results in 0.07 seconds. The table has columns for dataset, title, description, publisher, and publisherName. The results list various datasets registered in the KG Registry, including "Aktienführer...", "ClaimsKG@en", "Culture Kno...", "Data Set Kn...", "FactGRID@en", "Gemeinsam...", "InteractOA (...", "Linked Stage...", "LinkedOpen...", "MaRDI porta...", "MaschinenB...", and "NFDI4Earth ...".

https://kgi.services.base4nfdi.de/kg_registry/

Use Cases

How can you interact with KGs?

- Find existing KGs
 - within NFDI
 - within your fields
 - in general
 - starting from a specific piece of knowledge
- Use existing KGs
 - alone
 - in combination
 - with each other
 - with other resources
 - including other Basic Services
- Deploy your own KG
- Increase the interoperability of your KG with other KGs
- ...



Use Case

Deploy your own KG



Lisa

When I

am deciding which KG technology to deploy,

I need / I wish

an overview of existing tools with clear indication of advantages & disadvantages of each tool.

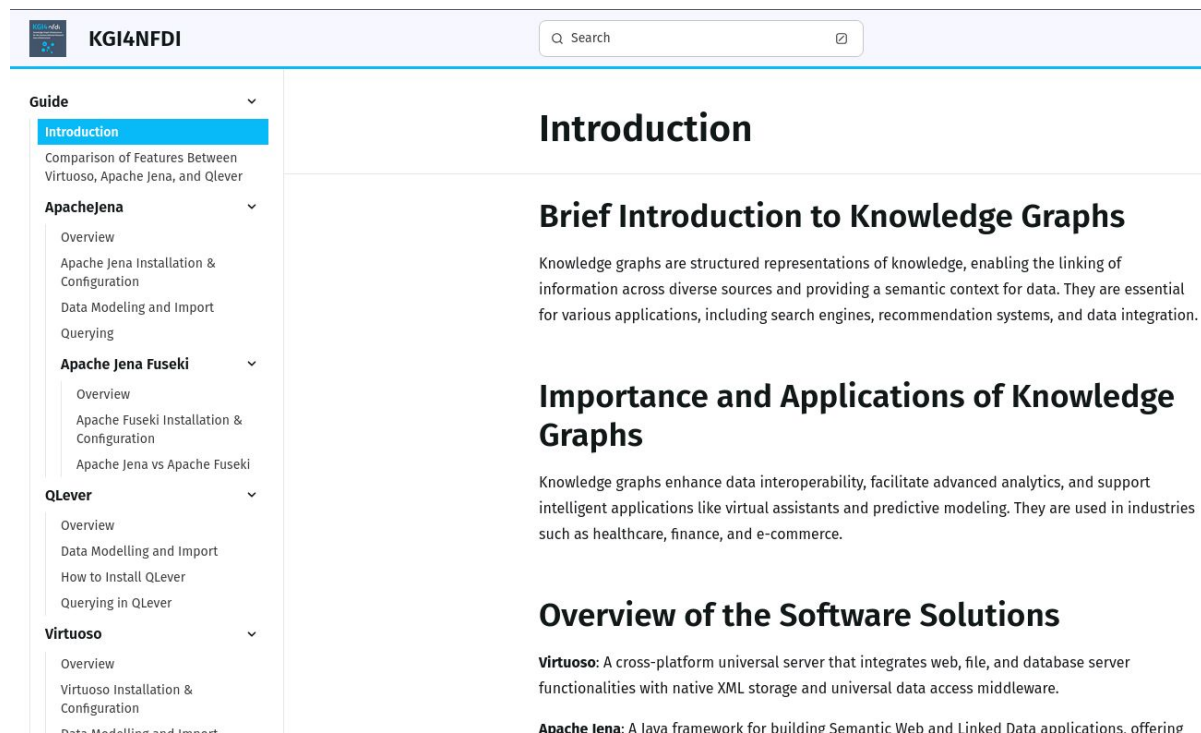
I am frustrated when

there is not enough documentation to support less experienced users of these technologies.



- ★ Several tool options to choose from
- ★ Guidance docs and deployment pipelines
- ★ Consulting service

KGI Guidelines



The screenshot shows the KGI4NFDI website interface. The header includes the KGI4NFDI logo and a search bar. The left sidebar contains a navigation menu with categories: Guide, ApacheJena, Apache Jena Fuseki, QLever, and Virtuoso. The 'Introduction' page is selected under the 'Guide' category. The main content area displays the title 'Introduction' and a sub-section 'Brief Introduction to Knowledge Graphs'. Below this, there is a paragraph explaining knowledge graphs. Another sub-section 'Importance and Applications of Knowledge Graphs' follows, with another explanatory paragraph. The final sub-section is 'Overview of the Software Solutions', which includes descriptions for 'Virtuoso' and 'Apache Jena'.

KGI4NFDI Search

Guide

- Introduction**
- Comparison of Features Between Virtuoso, Apache Jena, and QLever

ApacheJena

- Overview
- Apache Jena Installation & Configuration
- Data Modeling and Import
- Querying

Apache Jena Fuseki

- Overview
- Apache Fuseki Installation & Configuration
- Apache Jena vs Apache Fuseki

QLever

- Overview
- Data Modelling and Import
- How to Install QLever
- Querying in QLever

Virtuoso

- Overview
- Virtuoso Installation & Configuration
- Data Modeling and Import

Introduction

Brief Introduction to Knowledge Graphs

Knowledge graphs are structured representations of knowledge, enabling the linking of information across diverse sources and providing a semantic context for data. They are essential for various applications, including search engines, recommendation systems, and data integration.

Importance and Applications of Knowledge Graphs

Knowledge graphs enhance data interoperability, facilitate advanced analytics, and support intelligent applications like virtual assistants and predictive modeling. They are used in industries such as healthcare, finance, and e-commerce.

Overview of the Software Solutions

Virtuoso: A cross-platform universal server that integrates web, file, and database server functionalities with native XML storage and universal data access middleware.

Apache Jena: A Java framework for building Semantic Web and Linked Data applications, offering

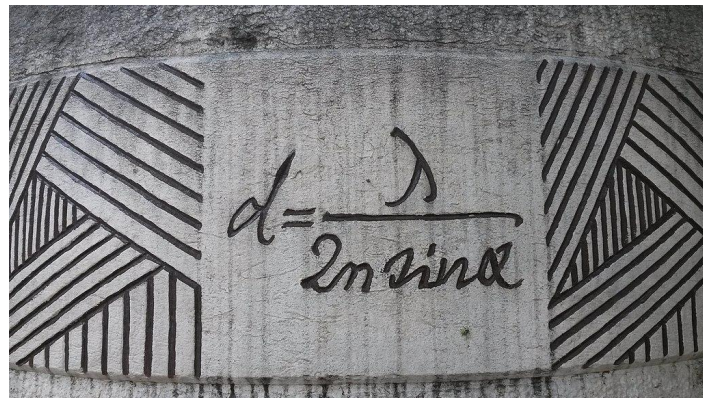
- <https://kgi4nfdi.github.io/Guidelines/>
- <https://kgi4nfdi.github.io/Guidelines/tutorial/>

Facets of Cross-Disciplinarity in KGs

Types of relevant questions

- Which KGs should (or do) hold which knowledge about an entity?
- How should the details be modeled within and across KGs?
- What if different KGs have conflicting information on “the same” entity?
- How do we identify “the same” concept or relationship across KGs?
- How can KGI leverage other Basic Services, and vice versa?

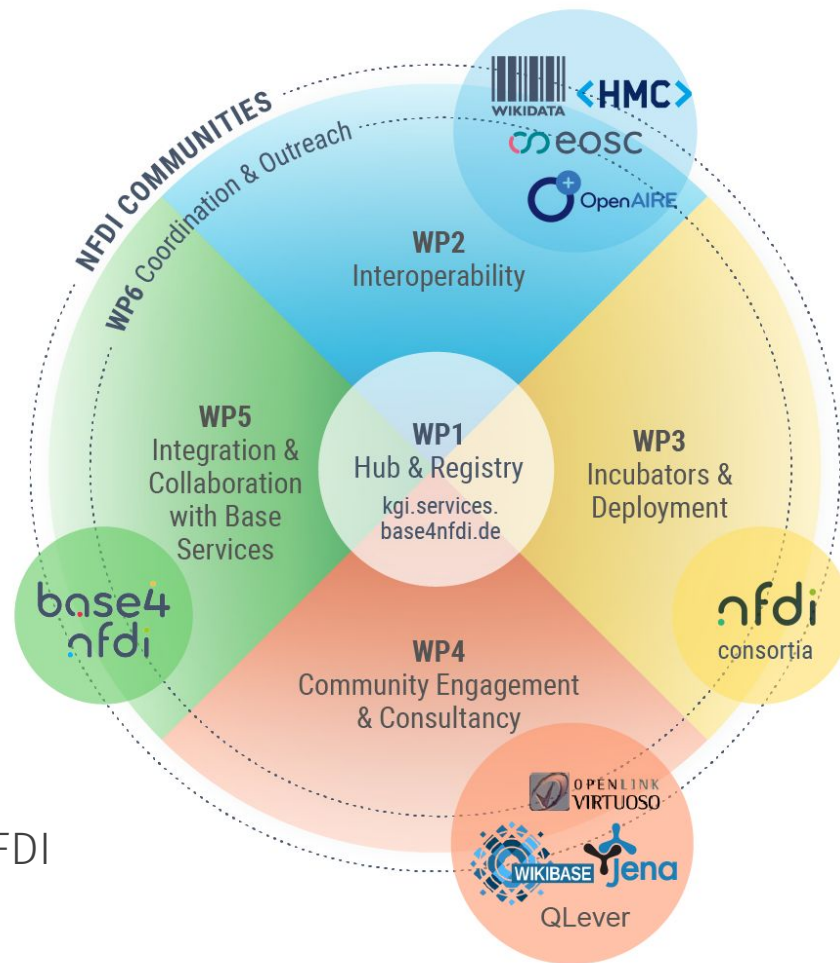
- It's a photograph of (part of) a monument.
- The monument depicts Latin and Greek letters.
- The letters are part of a mathematical formula.
- The formula represents some physical concepts.
- The physical concepts describe a limit.
- The limit could be circumvented by engineering.
- The engineering got the Nobel Prize in Chemistry.
- The chemistry is widely used in biology and medicine.
- All of this has a history as well as material, social and data aspects ...



Integration Phase

Provisional Plans

- Run incubator projects
- Extend the registry
- Enrich the hub
- Benchmark KG technologies
- Harmonize metadata
- Streamline deployment
- Integration with Base Services
- Community engagement
- Consultancy
- Provide training
- Documentation and reporting
- Coordination within and beyond NFDI



Get involved



More info at: <https://kgi.services.base4nfdi.de/> and <https://base4nfdi.de/projects/kgi4nfdi>

- ★ Share KG-related use cases
 - integration needs
 - existing KGs in your domains ([example](#))
 - data models for KG entries
 - ...
- ★ Take-advantage of consulting opportunities
 - Plan an incubator project
 - Attend open consulting sessions (third Wednesday of the month, 3pm)
 - Access training resources
 - ...
- ★ Explore KGI resources and provide feedback
 - [KGs in the registry](#)
 - [Virtuoso & Jena KG Guidelines](#)
 - [MaRDI meets 4Memory](#)
 - [FAIRJupyter](#)

Thank you!
Questions?



-  [**kgi4nfdi@lists.nfdi.de**](mailto:kgi4nfdi@lists.nfdi.de)
-  base4nfdi-servicestewards@lists.nfdi.de for general inquiries
-  [**base4nfdi.de/projects/kgi4nfdi**](https://base4nfdi.de/projects/kgi4nfdi)