



Knowledge graph infrastructure for the research community

2024-12-13

Lozana Rossenova, Renat Shigapov, Moritz Schubotz, Fidan Limani,
Benjamin Zapilko, Till Sauerwein, Daniel Mietchen, Muhammad Elhossary,
Konrad U. Förstner

Introduction:

Why KGs and why KGI

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

- 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 often used for agile data integration;
- KGs are already **widely used** by research data producers and managers in Germany.



Overview NFDI



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

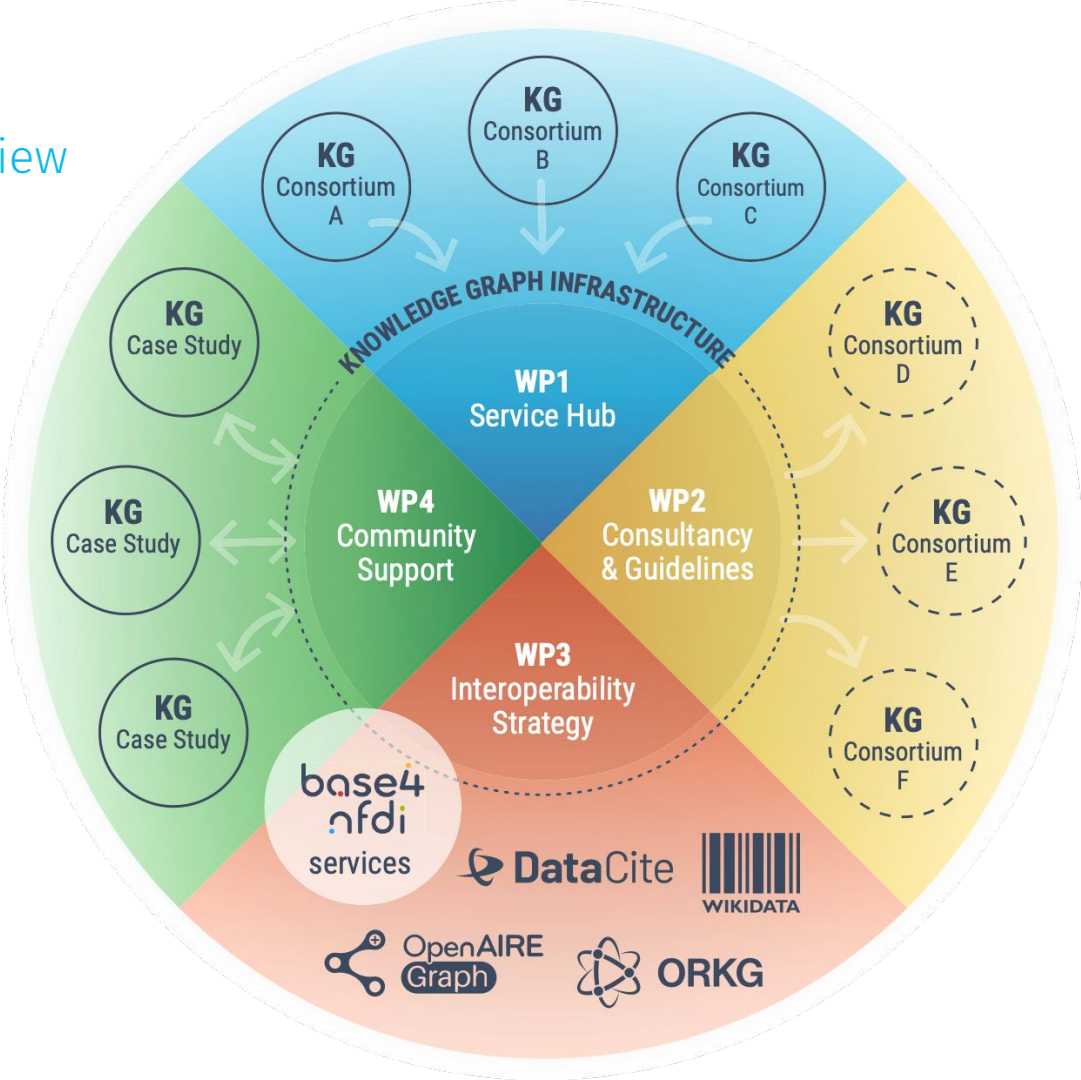
Life sciences

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

Natural sciences

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

KGI4NFDI:
Service overview



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



Aleen
a

When I

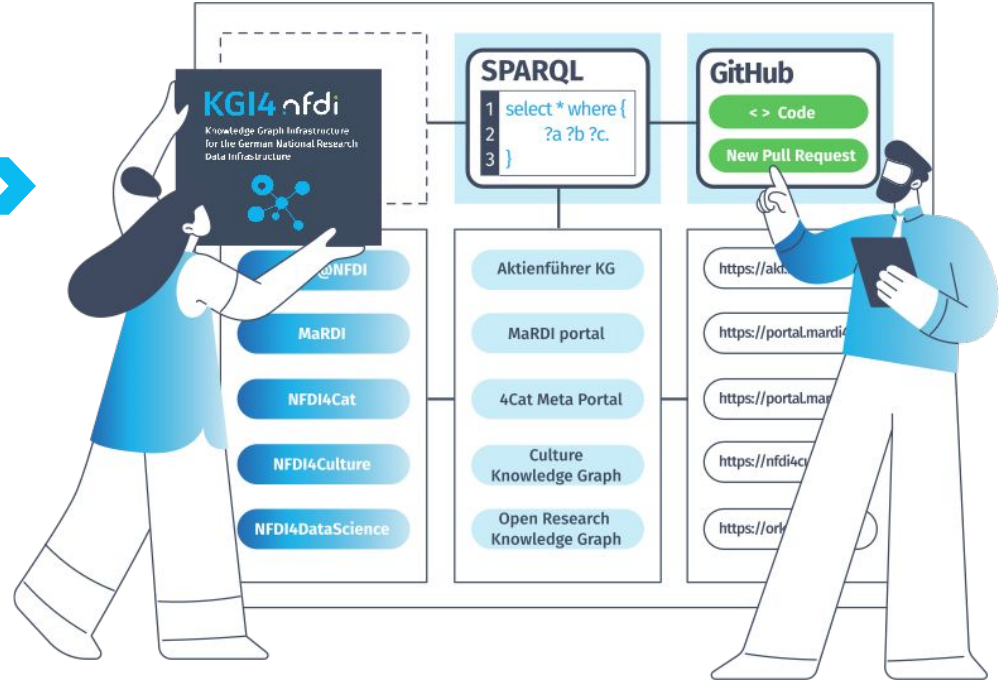
am looking for existing knowledge graphs in the domain of Bioinformatics,

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

/kgreg-service

query [add data](#) [edit](#) [info](#)

SPARQL Query

To try out some SPARQL queries against the selected dataset, enter your query here.

Example Queries

[Selection of triples](#) [Selection of classes](#)

Prefixes

[rdf](#) [rdfs](#) [owl](#) [xsd](#)

SPARQL Endpoint

/kgreg-service/query

Content Type (SELECT)

JSON

Content Type (GRAPH)

Turtle

```

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 ?publisherNameWikidata
10 WHERE {
11   ?dataset a dcat:Dataset ;
12           dcterms:title ?title ;
13           dcterms:publisher ?publisher .
14

```

Table Response 25 results in 0.895 seconds

Simple view Ellipse Filter query results Page size: 50 ↓ ⌵ ?

dataset	title	description	publisher	publisherNameWikidata
1<http://kgi.services.base4nfdi...	"Aktienführer Knowledge Graph" ^{@en}		<http://www.wikidata.org/entit...	"Mannheim University Library" ^{@en}
2<http://kgi.services.base4nfdi...	"ClaimsKG" ^{@en}	"a registry of claims" ^{@en}	<http://www.wikidata.org/entit...	"GESIS - Leibniz Institute for the Social Sci...
3<http://kgi.services.base4nfdi...	"Culture Knowledge Graph" ^{@en}	"A connector for all research data produced within the NFDI4Culture research landscape, improving the findability, a...	<http://www.wikidata.org/entit...	"FIZ Karlsruhe" ^{@en}
4<http://kgi.services.base4nfdi...	"Data Set Knowledge Graph (DSKG)"...	"a RDF data set about data sets which are linked to publications that mention the data sets" ^{@en}	<http://www.wikidata.org/entit...	"Karlsruhe Institute of Technology" ^{@en}
5<http://kgi.services.base4nfdi...	"FactGRID" ^{@en}		<http://www.wikidata.org/entit...	"Gotha Research Centre of the University ...
6<http://kgi.services.base4nfdi...	"Gemeinsame Normdatei (GND)" ^{@en}	"Integrated Authority File" ^{@en}	<http://www.wikidata.org/entit...	"German National Library" ^{@en}
7<http://kgi.services.base4nfdi...	"InteractOA (using Wikidata)" ^{@en}	"Very specific usage of Wikidata to demonstrate referencing to sources (started 2018 already)" ^{@en}	<http://www.wikidata.org/entit...	"ZB MED - Information Centre for Life Scie...
8<http://kgi.services.base4nfdi...	"Linked Stage Graph" ^{@en}	"A Knowledge Graph developed during the Coding da Vinci Süd 2019 hackathon using a dataset by the National Arch..."	<http://www.wikidata.org/entit...	"FIZ Karlsruhe" ^{@en}
9<http://kgi.services.base4nfdi...	"LinkedOpenData-Service des B3Kat"...	"descriptions of 28 million titles from 200 academic libraries in Bavaria, Berlin and Brandenburg" ^{@en}	<http://www.wikidata.org/entit...	"Bavarian State Library" ^{@en}
10<http://kgi.services.base4nfdi...	"MaRDI portal" ^{@en}		<http://www.wikidata.org/entit...	"FIZ Karlsruhe" ^{@en}

<https://sparql.kgi.services.base4nfdi.de/> (WIP!)

Use cases: 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



Guide

Introduction

ApacheJena

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

Apache Jena Fuseki

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

Virtuoso

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

Tutorial

- Introduction

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 tools for data handling, ontology management, and SPARQL querying.

On this page

Overview

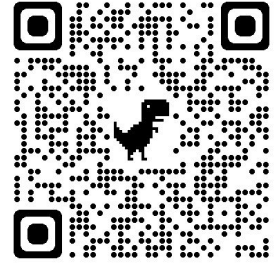
Brief Introduction to Knowledge Graphs

Importance and Applications of Knowledge Graphs

Overview of the Software Solutions

Purpose and Scope of the Guidelines

Get involved!



★ **Participate in our survey!**

Survey open until:
2024-12-31

★ **Visit the service hub**

<https://kgi.services.base4nfdi.de/>

★ **Contribute to the federation show cases**

Get in touch at: kgi4nfdi@lists.nfdi.de

Thank you!
Questions?



-  kgi4nfdi@lists.nfdi.de
-  base4nfdi-servicestewards@lists.nfdi.de for general inquiries
-  <https://github.com/KGI4NFDI>