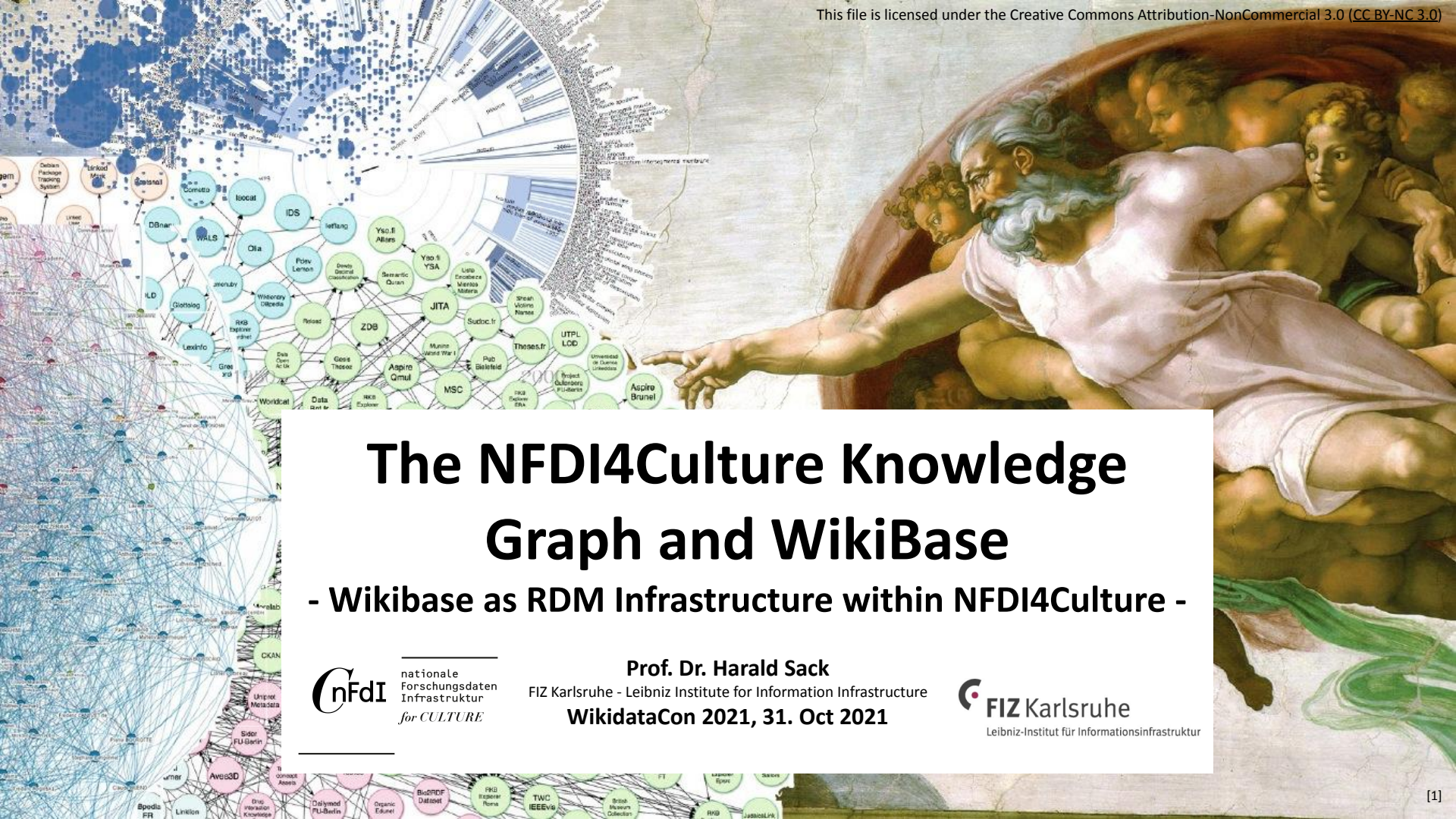


The NFDI4Culture Knowledge Graph and WikiBase

- Wikibase as RDM Infrastructure within NFDI4Culture -

Prof. Dr. Harald Sack
FIZ Karlsruhe - Leibniz Institute for Information Infrastructure

WikidataCon 2021, 31. Oct 2021

nationale Forschungsdaten Infrastruktur *for CULTURE*

FIZ Karlsruhe
Leibniz-Institut für Informationsinfrastruktur

The NFDI4Culture Knowledge Graph

CONTENT:

- Index and Metadata for NFDI4Culture Project Resources
  - 4Culture Data Resources
  - 4Culture Data Services
  - Persons and Institutions
  - Expertises
- Hub for/to NFDI4Culture Data Resources to enable federated access
- Connection to and federation with further NFDI domains

COVERAGE:

- NFDI4Culture Domains:
  - Architecture
  - Art History
  - Musicology
  - Performing Arts
  - Media Studies

[2]

Why do we need (formal) Semantics?

- NFDI4Culture is a **huge data integration effort** (research data resources from 70+ partner institutions)
- Using ontologies based on standardized metadata to **enable low effort research data integration**
- Enable **semantic and exploratory search** over federated research data resources
- Enable **intelligent recommendations** and **question answering** ability over federated data resources

How does (formal) semantics enable low(er) effort data integration?

a quantity ?
a point in time ?
a time span?
a length?
a weight?
a measurement?
a code ?
a character string?

42

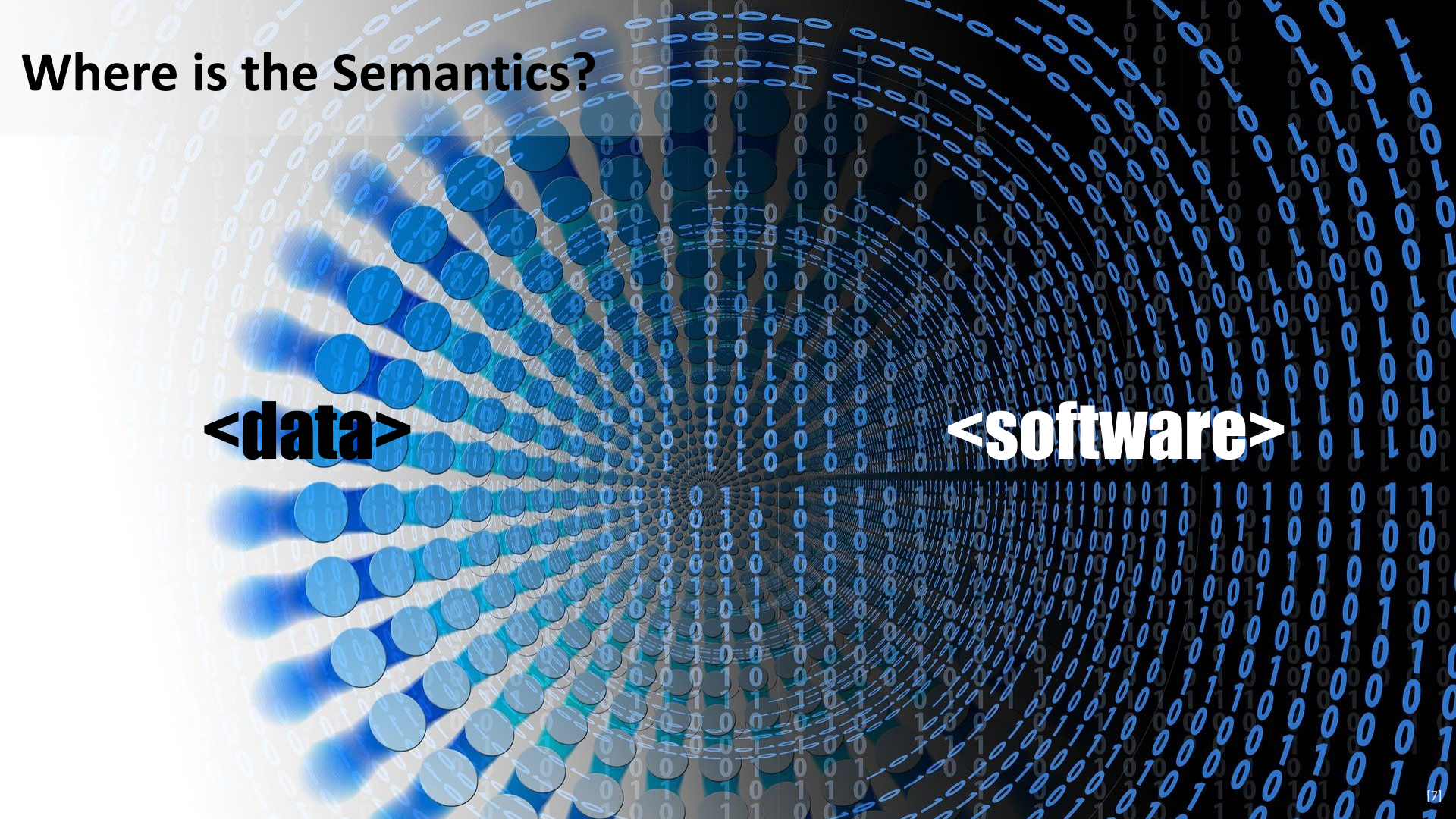
Alice in Wonderland has 42 illustrations [4]

Answer to the Ultimate Question of Life, the Universe, and Everything

Molybdenum has atomic number 42 [5]

Object 42 in the Messier catalogue, M42 [6]

Where is the Semantics?



Traditional Scenario

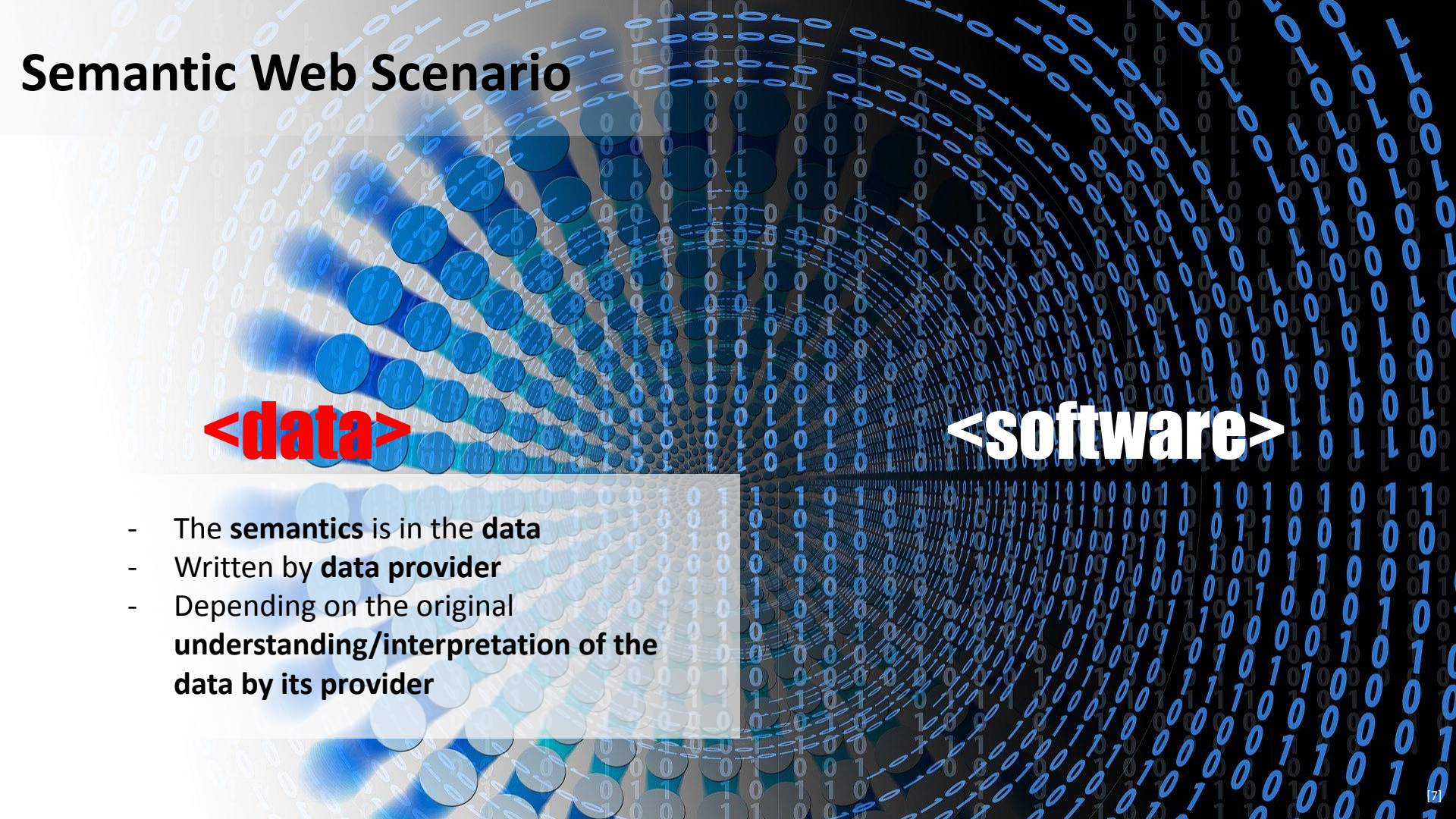
- The **semantics** is in the **programme code**
- Written by **programmers**
- Depending on the **understanding/interpretation of the individual programmer**

Traditional Scenario

- **Extensive software maintenance** necessary (for any data structure changes)
- **Error prone** (understanding / interpretation of the programmer)

- The **semantics** is in the **programme code**
- Written by **programmers**
- Depending on the individual **understanding/interpretation of the data by the programmer**

Semantic Web Scenario



- The **semantics** is in the **data**
- Written by **data provider**
- Depending on the original **understanding/interpretation of the data by its provider**

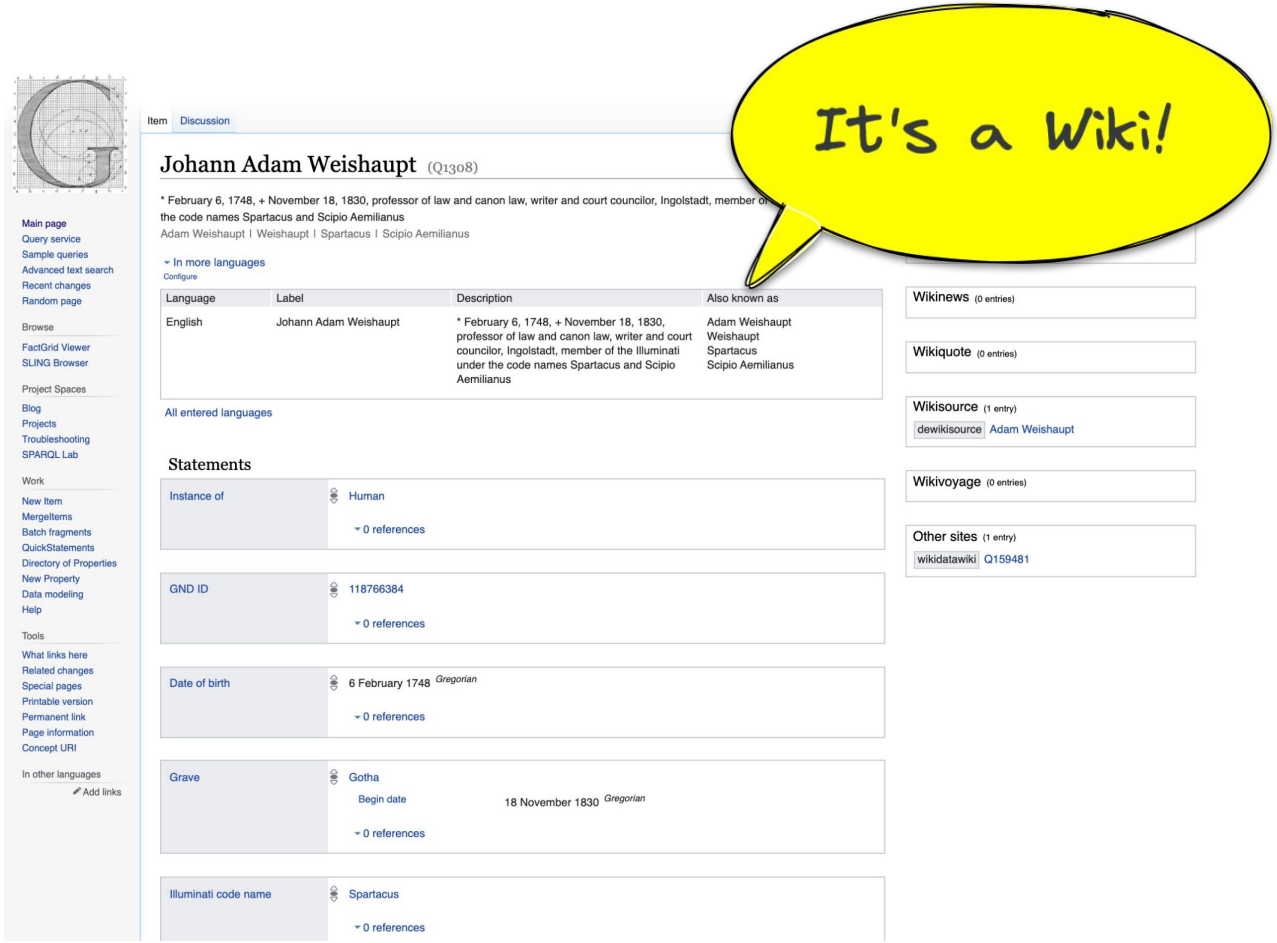
Semantic Web Scenario

- **Less software maintenance** (regarding changes in the data structure)
- **Less error prone**
- **Simpler data integration**
- **Simpler data reuse**

- The **semantics** is in the **data**
- Written by **data provider**
- Depending on the original **understanding/interpretation of the data by its provider**

Why WikiBase?

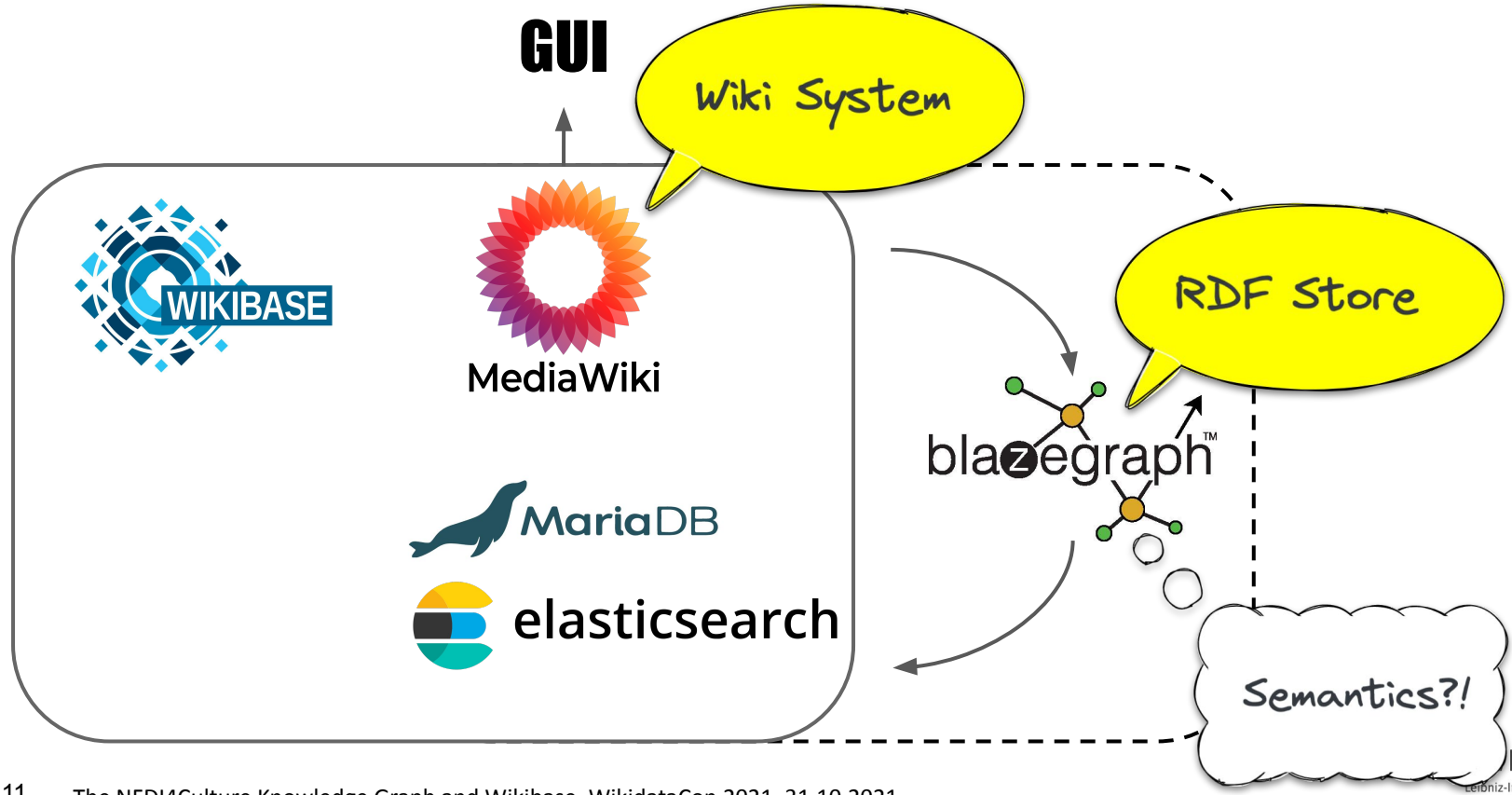
- **GUI** for interactive collaborative editing of **structured data**
- **Convenient way to enable/engage Communities**
- **Connects to a Semantic Data store**



https://database.factgrid.de/wiki/Item:Q1308

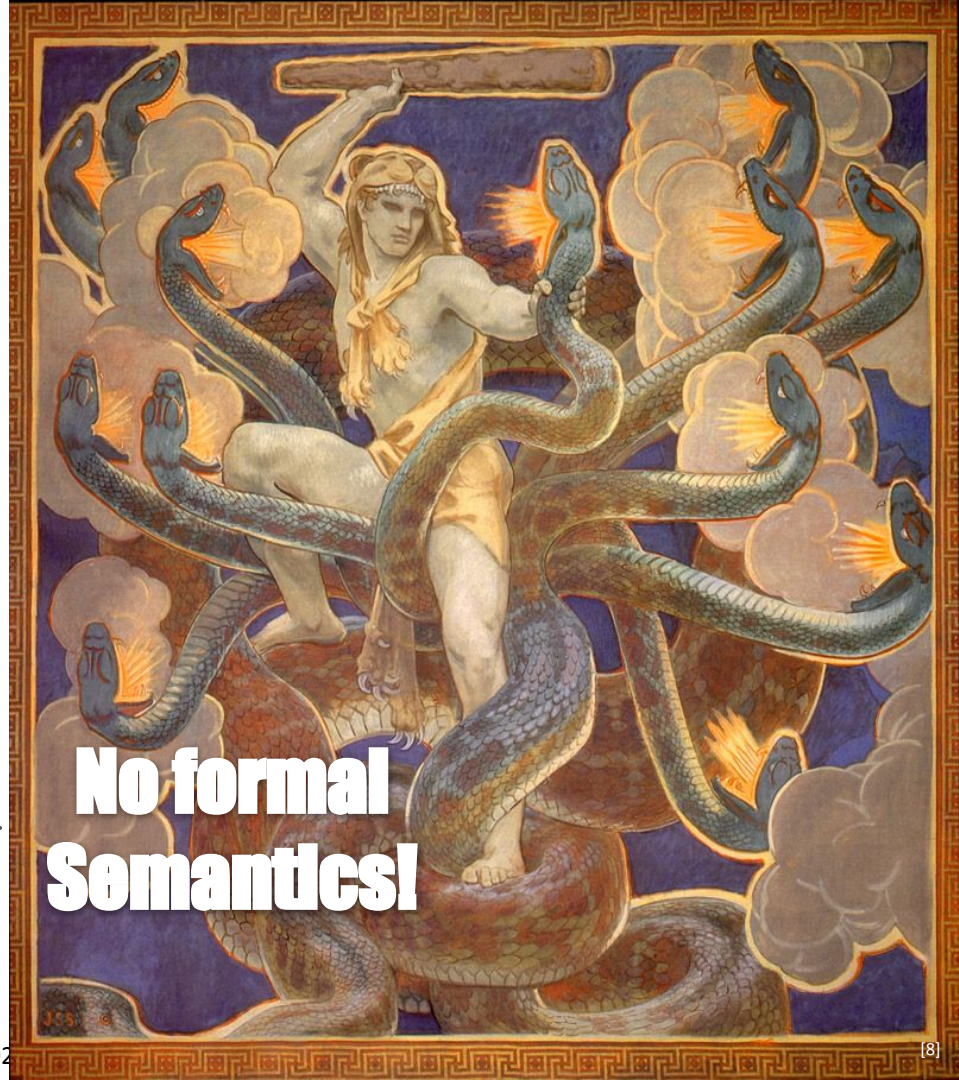
FIZ Karlsruhe
Leibniz-Institut für Informationsinfrastruktur

WikiBase Architecture



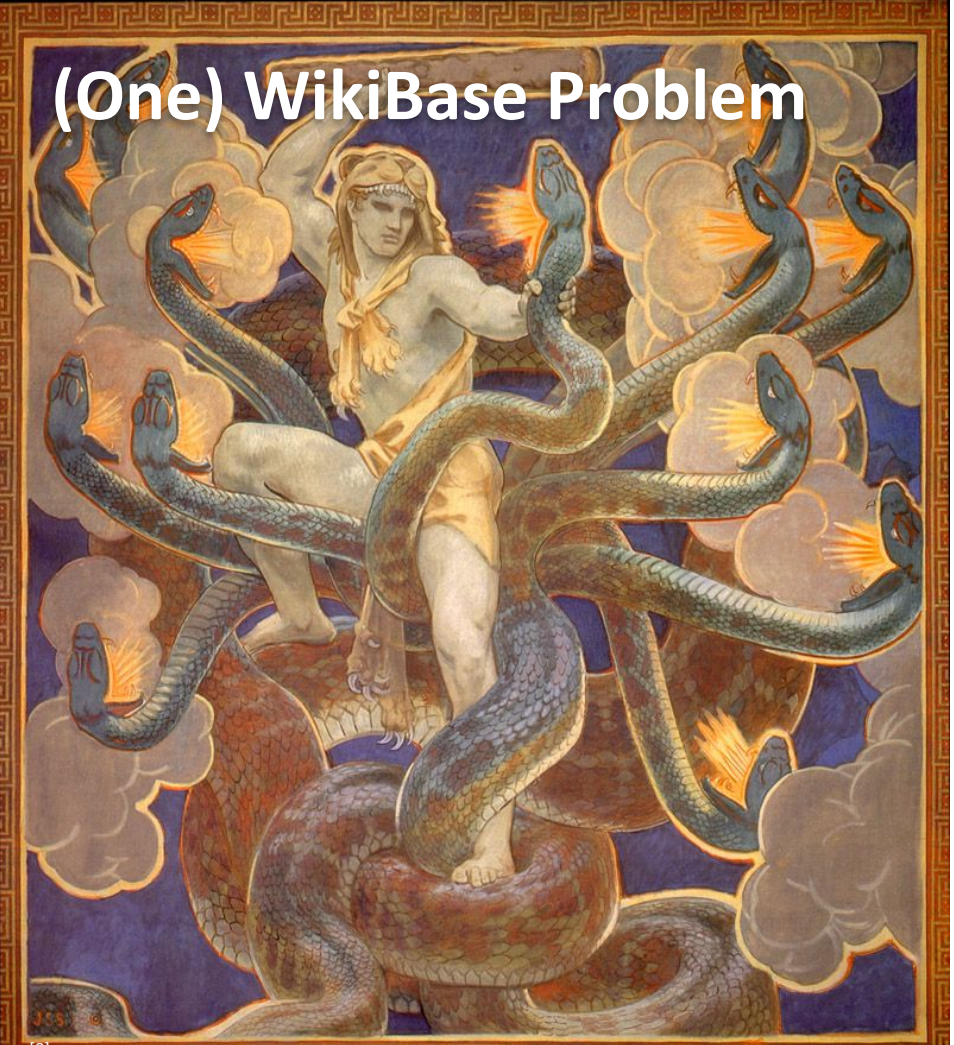
(One) WikiBase Problem

- Wikibase is originally a **Wiki only** and **no native RDF triple store**
- Data is only **relational data** (MariaDB/mySQL)
- Attached Triple Store only contains **“flat” triples** without explicit semantics
  - **No use of standard W3C Semantic Web vocabularies (RDF, RDFS, OWL)**



[8]

(One) WikiBase Problem



- Triple Store (blazegraph) is **only an addOn**
- **No transparent synchronous bidirectional interaction** between Wikibase and blazegraph
- All data changes have to be committed via
  - Wiki interface
  - Wiki API
  - Tools using Wiki API

[8]

FIZ Karlsruhe
Leibniz-Institut für Informationsinfrastruktur

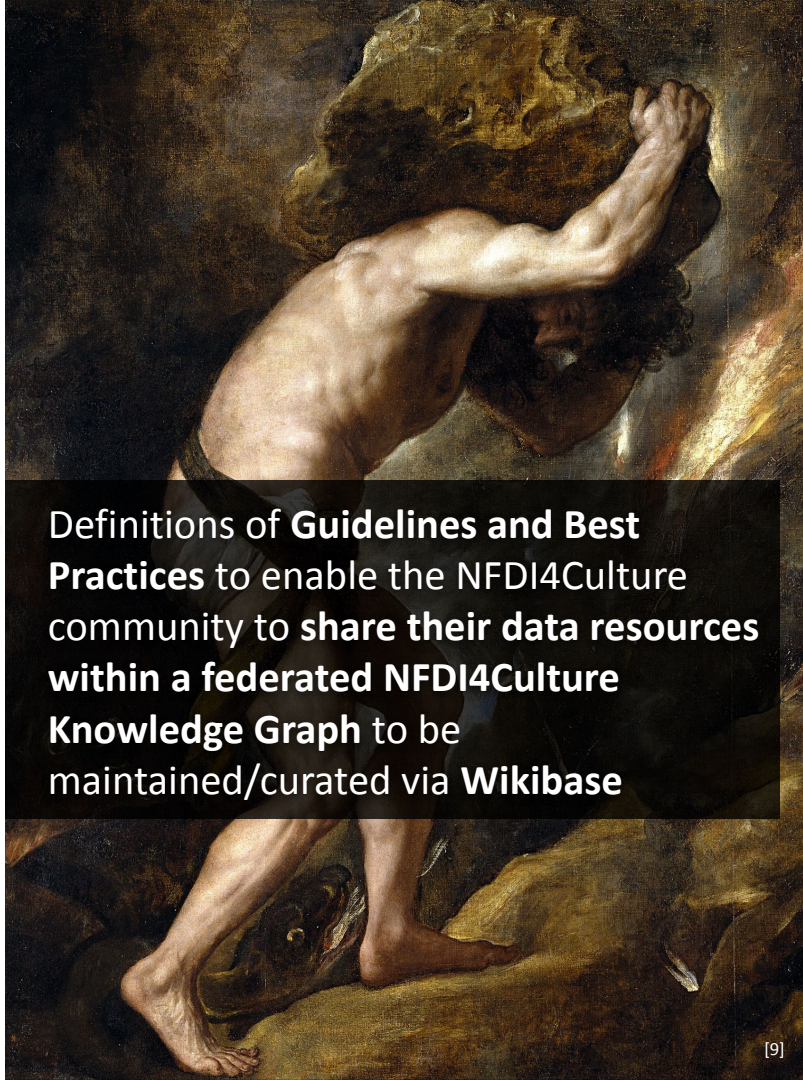
NFDI4Culture - Ongoing Work (Part 1)

- Evaluation of **viable workarounds and solutions** to enable real data semantics with existing Wikibase
  - Declarative semantic mapping
  - Data import / export
  - Semantic Extension

[9]

NFDI4Culture - Ongoing Work (Part 1)

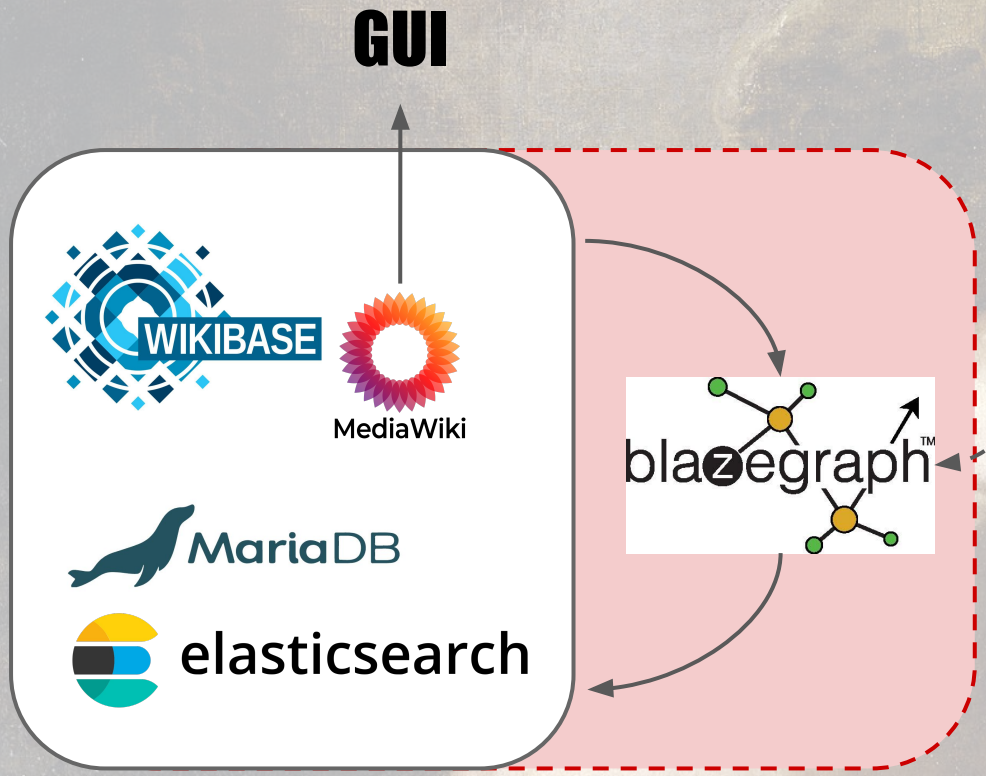
- Evaluation of **viable workarounds and solutions** to enable real data semantics with existing Wikibase
  - Declarative semantic mapping
  - Data import / export
  - Semantic Extension



Definitions of **Guidelines and Best Practices** to enable the NFDI4Culture community to **share their data resources within a federated NFDI4Culture Knowledge Graph** to be maintained/curated via **Wikibase**

[9]

Potential NFDI4Culture KG Architecture(s) - Variant 1



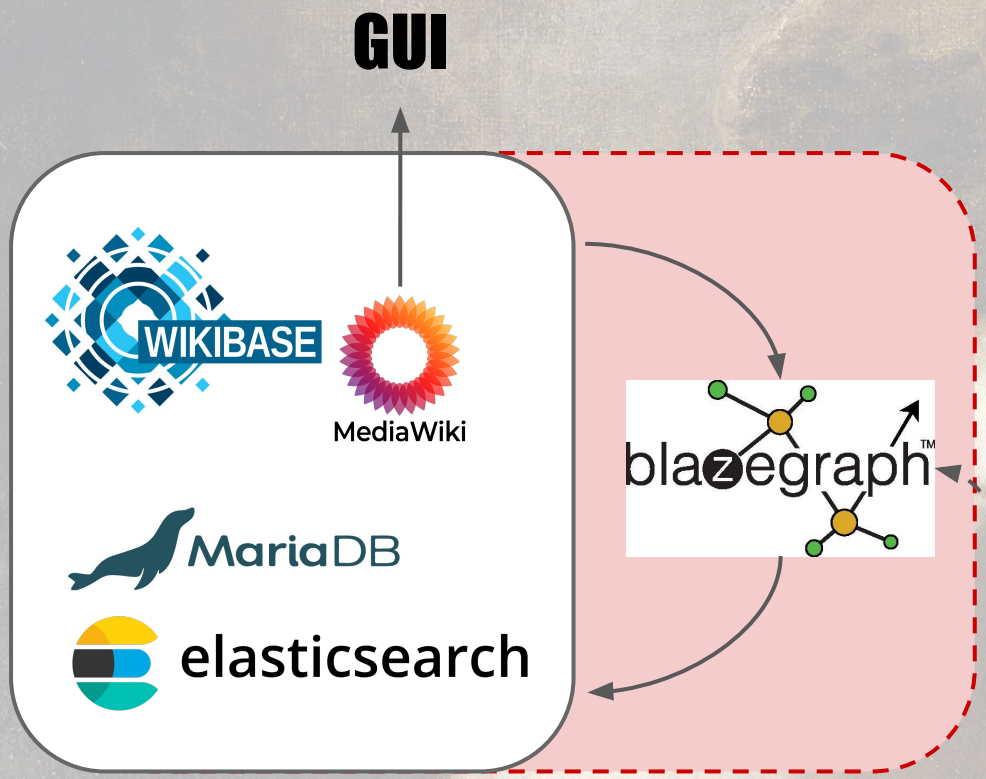
to be done only once...

- **W3C vocabularies** mapping to Wikibase entities and properties

```
wdt:P31 a rdf:Property .
wdt:P31 owl:equivalentProperty rdf:type .

wdt:P279 a rdf:Property .
wdt:P279 owl:equivalentProperty rdfs:subClassOf .

wdt:P1647 a rdf:Property .
wdt:P1647 owl:equivalentProperty rdfs:subPropertyOf .

...
```


FIZ Karlsruhe
Leibniz-Institut für Informationsinfrastruktur

Potential NFDI4Culture KG Architecture(s) - Variant 1



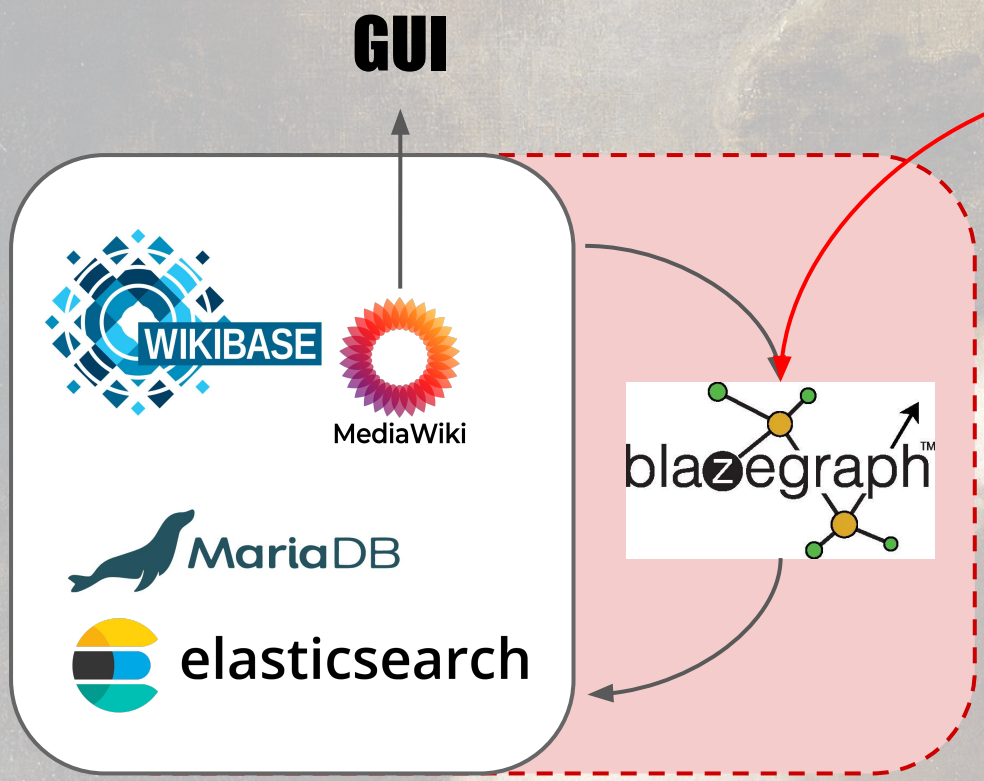
- W3C vocabula
  entities and prop
- Mapping of further Wikibase entities and properties to **existing semantic vocabularies/ontologies**

to be synchronized with every new entry or fact...

```
wd:Q42 owl:sameAs dbp:Douglas_Adams .
wd:Q43 owl:sameAs dbp:Turkey .
wd:Q44 owl:sameAs dbp:Beer .
wd:Q5  owl:equivalentClass dbo:Person .

...
```

FIZ Karlsruhe
Leibniz-Institut für Informationsinfrastruktur

Potential NFDI4Culture KG Architecture(s) - Variant 1

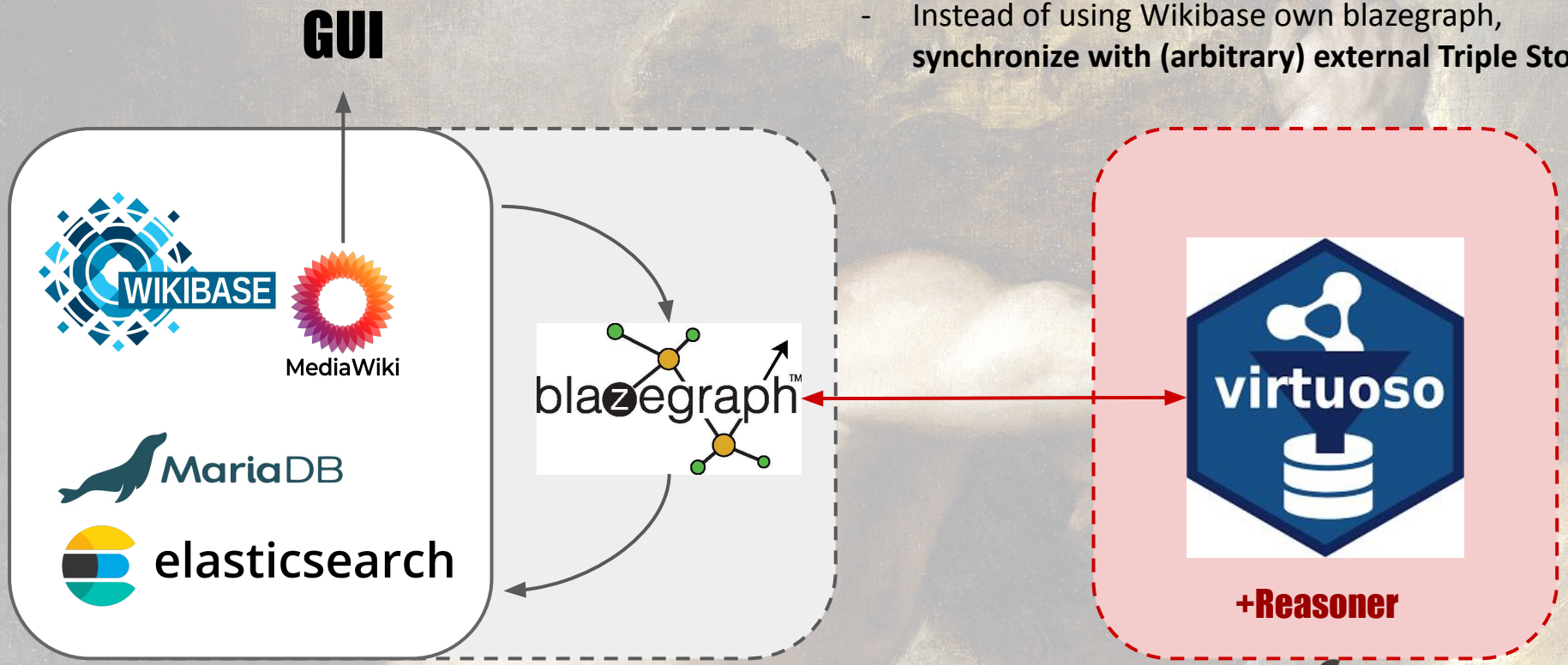


Reasoner

- Apply **reasoner**
  - to **materialize missing semantic triples** from mapping files via inferences
  - to make implicit knowledge explicit
  - and to **detect potential inconsistencies**
- One-way **synchronization with Wikibase** ***(way back is still missing)***

Potential NFDI4Culture KG Architecture(s) - Variant 2

- Instead of using Wikibase own blazegraph, **synchronize with (arbitrary) external Triple Store**



FIZ Karlsruhe
Leibniz-Institut für Informationsinfrastruktur

Further Wikibase Requirements from the Research Community

- More **flexible Wiki templates**
- Removal of existing **Bottlenecks**
- Free (Re-)Use of **existing vocabularies/ontologies**
- Free (Re-)Use of **own/existing URIs**
- Fully **transparent communication** between Wiki System and Triple/Graph Store
- **Fully fledged Access Control** on the entity/fact level
- Possibility to (collaboratively) **edit/design ontologies**

[1]

...y much for your attention!

FIZ Karlsruhe
Leibniz-Institut für Informationsinfrastruktur

Picture References

[1] Harald Sack, *The Creation of Linked Open Data* (2014), remixing Michelangelo, *The Creation of Adam* (1511) https://commons.wikimedia.org/wiki/File:Michelangelo_-_Creation_of_Adam_(cropped).jpg, and *Linking Open Data cloud diagram*, by Richard Cyganiak and Anja Jentzsch. http://lod-cloud.net/, [CC-BY-SA 3.0]

[2] Baldassare Peruzzi (1481-1536) (after) - *Apollo and the Muses Dancing* - 609020 - National Trust (1890-1899), Art UK, https://commons.wikimedia.org/wiki/File:Baldassare_Peruzzi_(1481-1536)_(after)_-_Apollo_and_the_Muses_Dancing_-_609020_-_National_Trust.jpg?uselang=de [public domain]

[3] Carole Raddato, *Oedipus and the Sphinx of Thebes*, Red Figure Kylix, c. 470 BC, from Vulci, attributed to the Oedipus Painter, Vatican Museums, https://commons.wikimedia.org/wiki/File:Oedipus_and_the_Sphinx_of_Thebes,_Red_Figure_Kylix,_c._470_BC,_from_Vulci,_attributed_to_the_Oedipus_Painter,_Vatican_Museums_(9665213064).jpg, [CC-BY-SA 2.0]

[4] John Lenniel, *Alice's Mad Tea Party*, Alice in Wonderland (1865), https://commons.wikimedia.org/wiki/File:John_Tenniel-_Alice%27s_mad_tea_party,_colour.jpg, [public domain]

[5] Alchemist-hp, *Molybdenum crystaline fragment and 1cm3 cube* (2010), https://commons.wikimedia.org/wiki/File:Molybdenum_crystaline_fragment_and_1cm3_cube.jpg, [CC-BY-NC-ND 3.0]

[6] NASA, ESA, M. Robberto (Space Telescope Science Institute/ESA) and the Hubble Space Telescope Orion Treasury Project Team, *Orion Nebula - Hubble 2006*, https://commons.wikimedia.org/wiki/File:Orion_Nebula_-_Hubble_2006_mosaic_18000.jpg, [public domain]

[7] Pixabay, https://pixabay.com/photos/data-computer-internet-online-www-2899901/, [pixabay license, https://pixabay.com/service/license/]

[8] John Singer-Sargent, *Hercules* (1921), https://commons.wikimedia.org/wiki/Category:Paintings_of_Hercules_(Heracles)_and_Lernaean_Hydra#/media/File:Singer_Sargent,_John_-_Hercules_-_1921.jpg , [public domain]

[9] Titian, *Sysiphus* (cropped) (1548/49), https://commons.wikimedia.org/wiki/File:Punishment_sisyph.jpg, [public domain]

[10] Jacques-Louis David, *The Oath of the Horatii* (1784), https://commons.wikimedia.org/wiki/File:David-Oath_of_the_Horatii-1784.jpg, [public domain]

FIZ Karlsruhe
Leibniz-Institut für Informationsinfrastruktur