



# The Helmholtz KG – Status Quo

## Helmholtz Metadata backbone

Widespread integration into scientific, infrastructure and administrative workflows in Helmholtz.

### Increase coverage

- Onboard new data sources
- Extend and refine HKG datamodel
- Feedback metadata analytics

### Prototyping

- Methods testing
- Requirement Analysis
- Semantic Analysis

### Refine & extend scope

- Increase semantic expressiveness
- Increase graph data quality actively
- Explore new, additional modes of use

### Infrastructure Foundation

- Robust, adaptable semantic framework
- Scalable pipeline and storage
- Sustainable and robust deployment



HELMHOLTZKG

# The Helmholtz KG – Status Quo



[sparq.unhide.helmholtz-metadaten.de](https://sparq.unhide.helmholtz-metadaten.de)

[Find all events since 2020 with the topic "metadata"](#)

[List all Helmholtz publications published in journal "nature"](#)

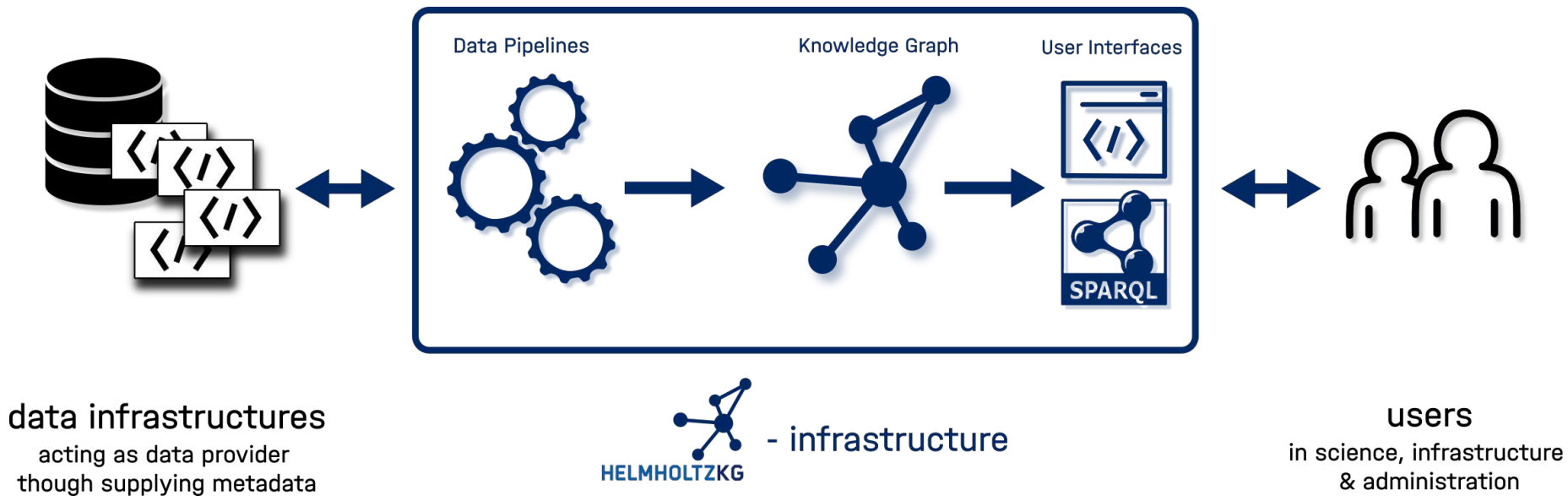
# The Helmholtz KG - Status Quo

---

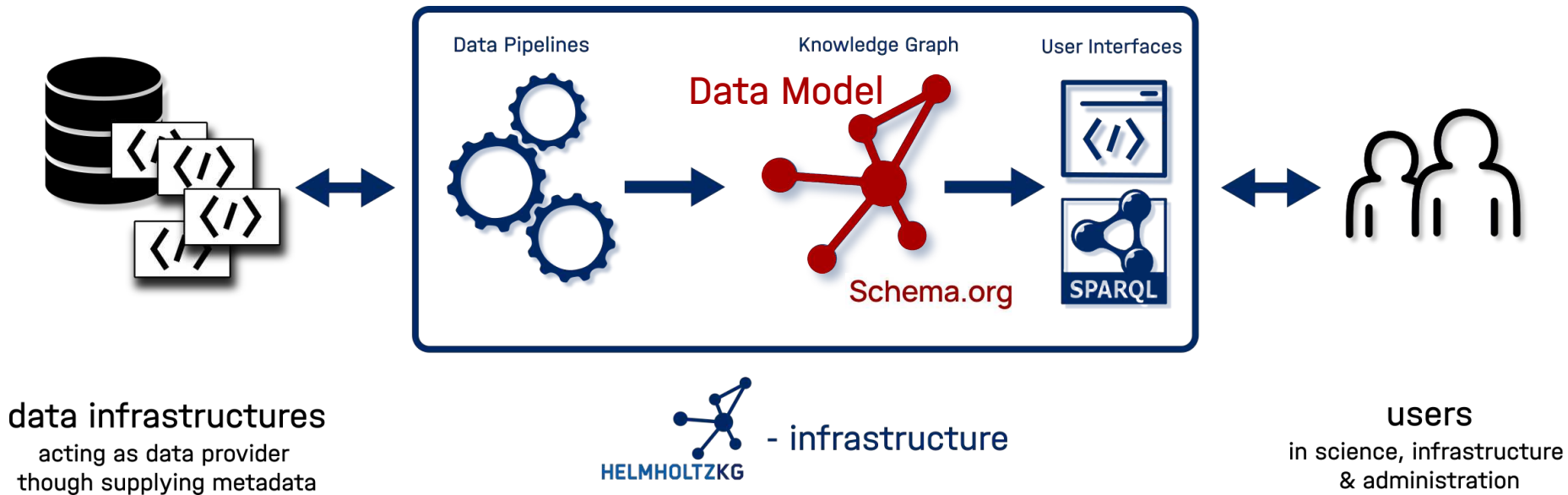


[search.unhide.helmholtz-metadaten.de](https://search.unhide.helmholtz-metadaten.de)

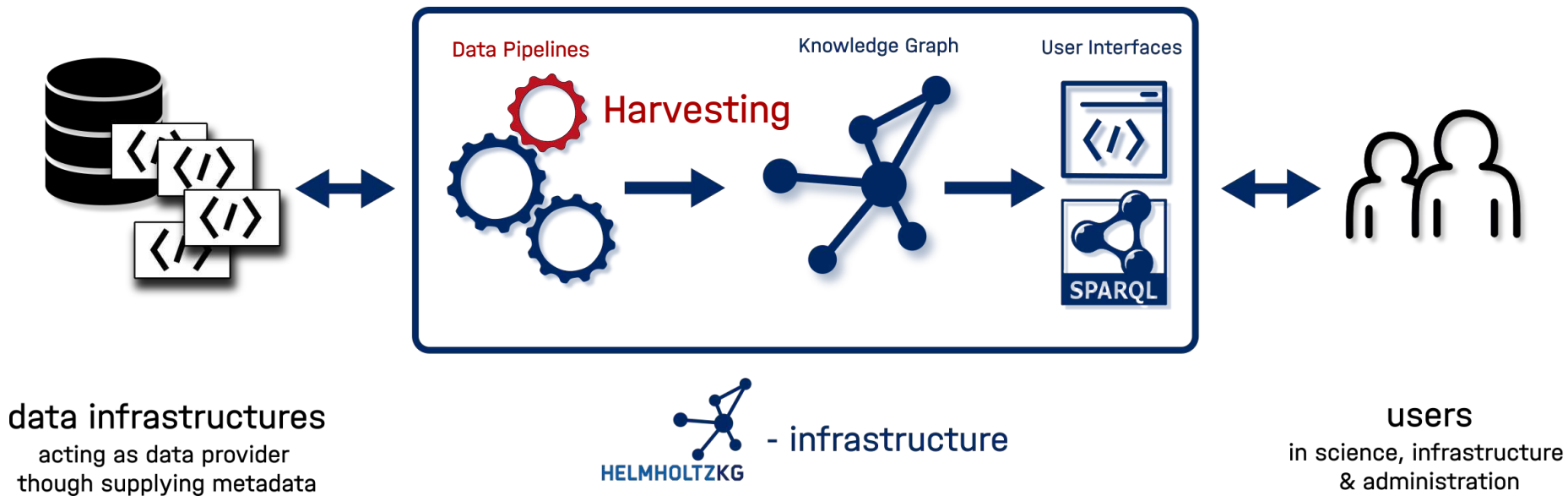
# Architecture and System Overview



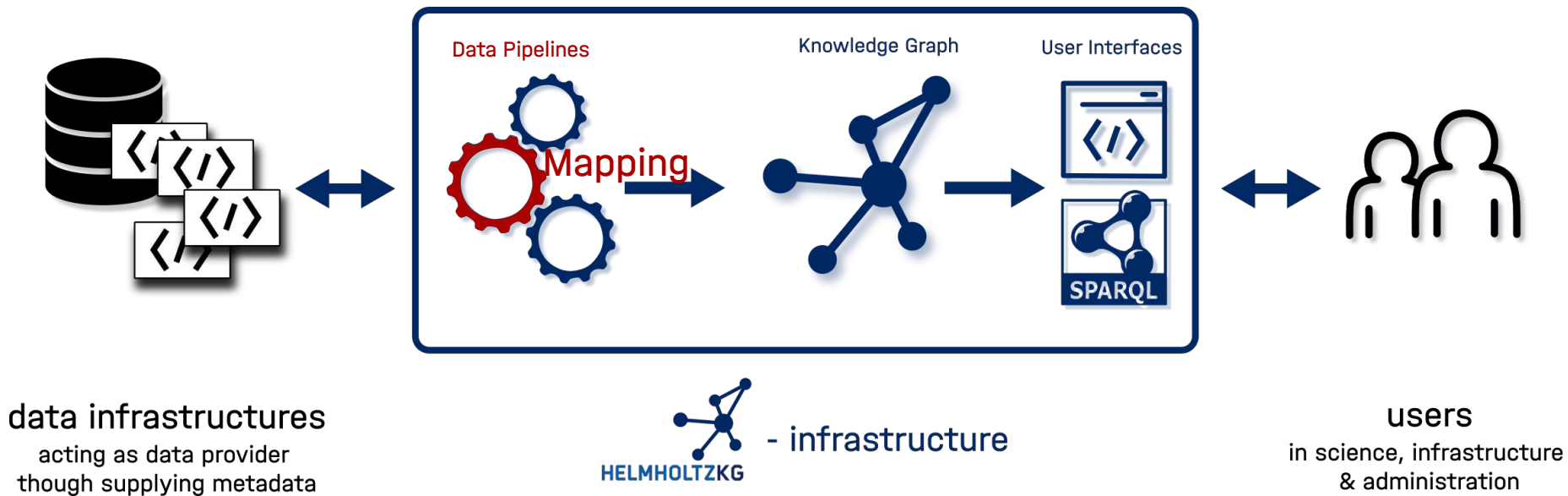
# Where is the Data Model?



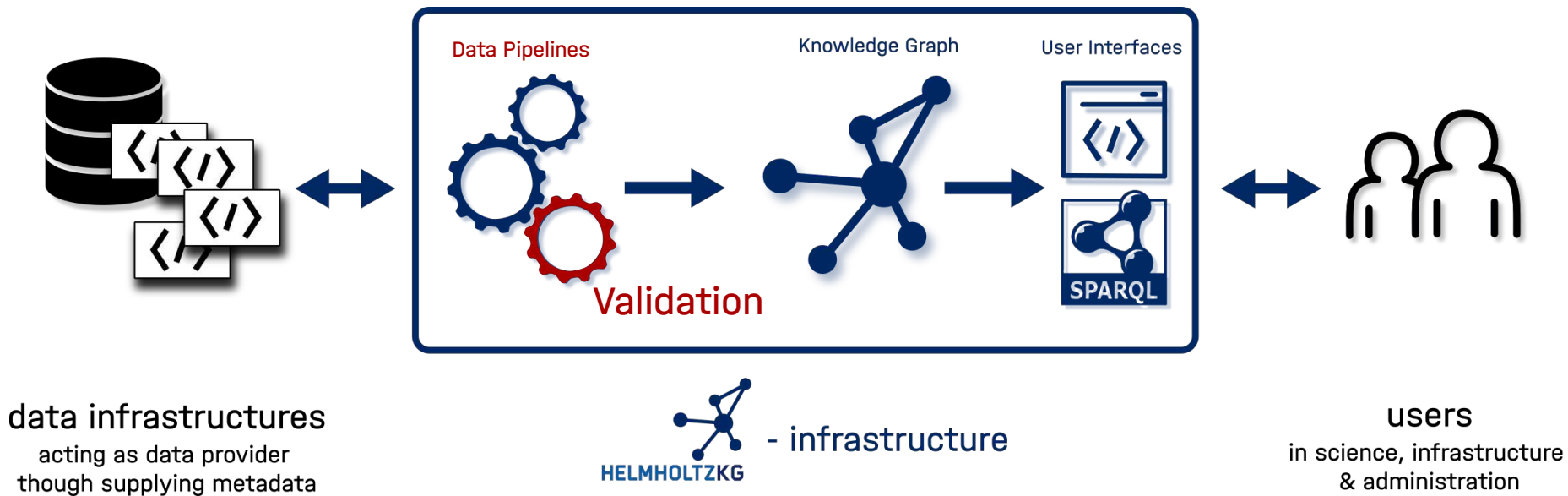
# Data Pipelines



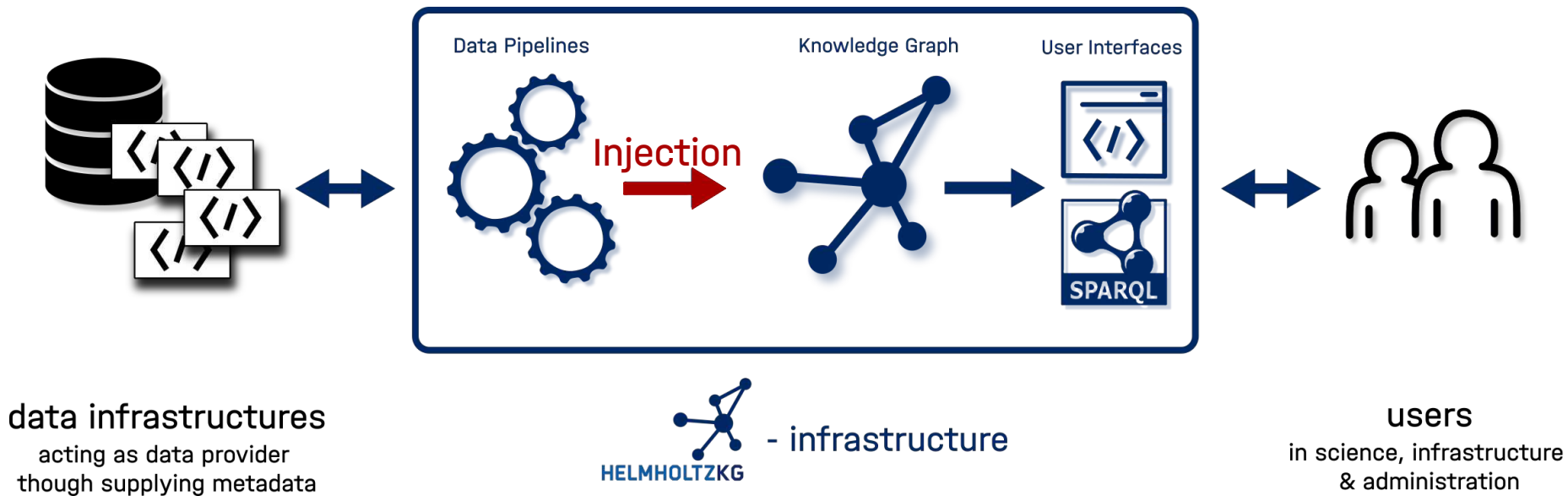
# Data Pipelines



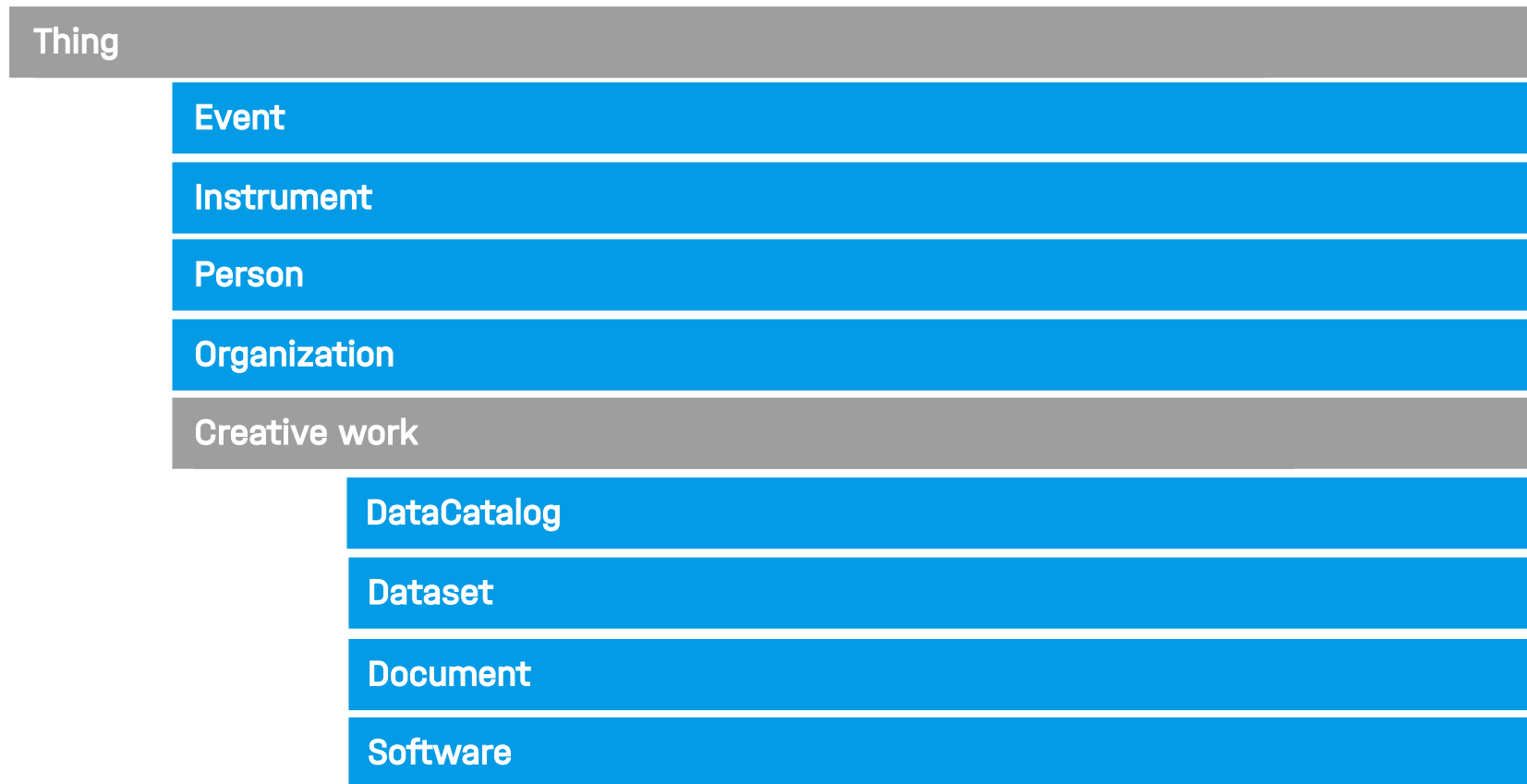
# Data Pipelines



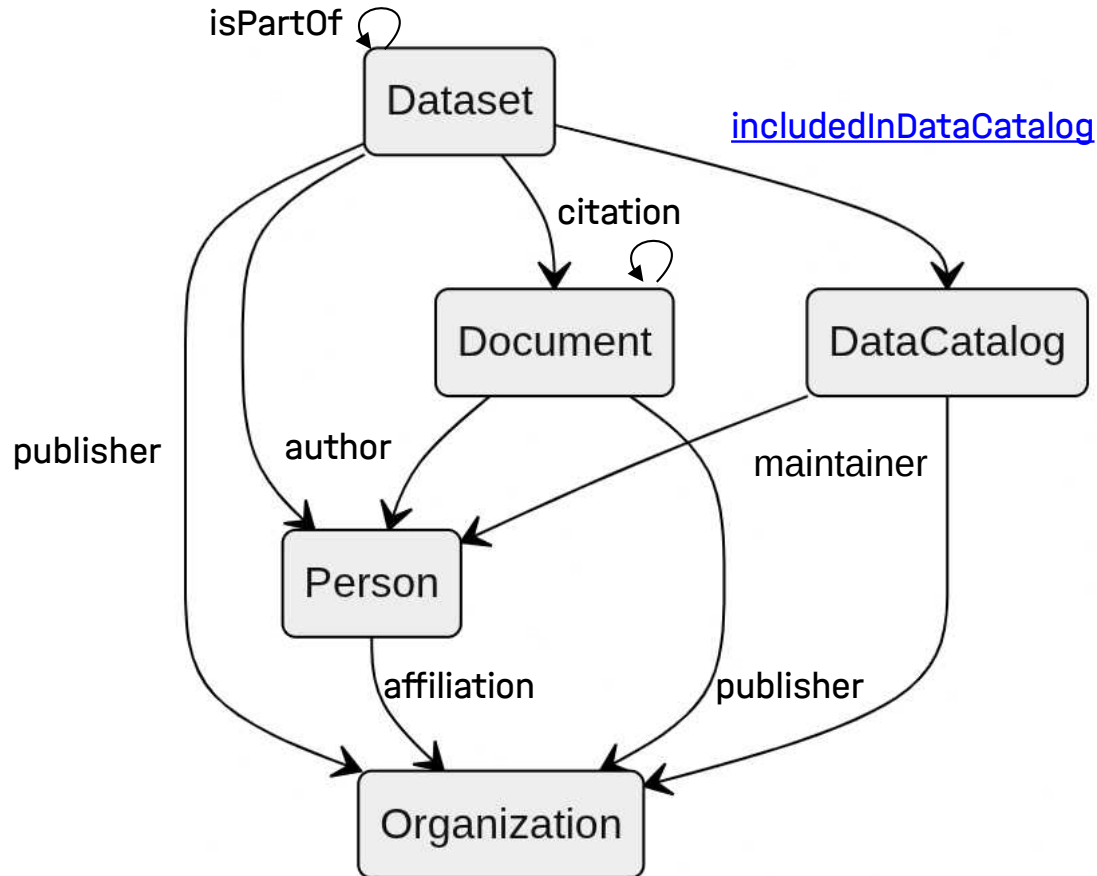
# Data Pipelines



# Data Model – Main components and their Hierarchy



# Data Model Relationships (example)



# Helmholtz KG as Umbrella Infrastructure



- The RDI's metadata will always be more expressive than the HKG representation
- We want to increase the semantic overlap between HKG and RDIs to allow for
  - a) better findability of the resources,
  - b) increase harmonization at "deeper levels", and
  - c) allow federated queries.