FAIR Data Spaces Final Event

Architecture Overview

Sebastian Beyvers, *Technical Lead*

Bioinformatics & Systemsbiology, JLU Giessen

GEFÖRDERT VOM

Bundesministerium
für Bildung
und Forschung

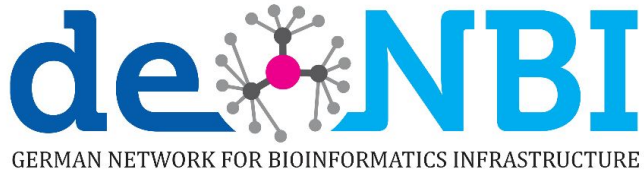
FAIR Data Spaces

Motivation

Development of a shared **cloud-based** dataspace for industry and science

Cloud foundations

- **Community Cloud**: de.NBI
    - One of the largest scientific Community Clouds in Germany
    - Provides free storage & compute resources for researchers in life science
    - Operational since 2016
- **Public Cloud**: Open Telekom Cloud



Technologies

- Free and Open Source
- Large communities
  - Science & Industry
  - Widespread adoption
- Vendor agnostic
  - Sovereignty

Operations

Standards

Object storage

Infrastructure

Interfaces

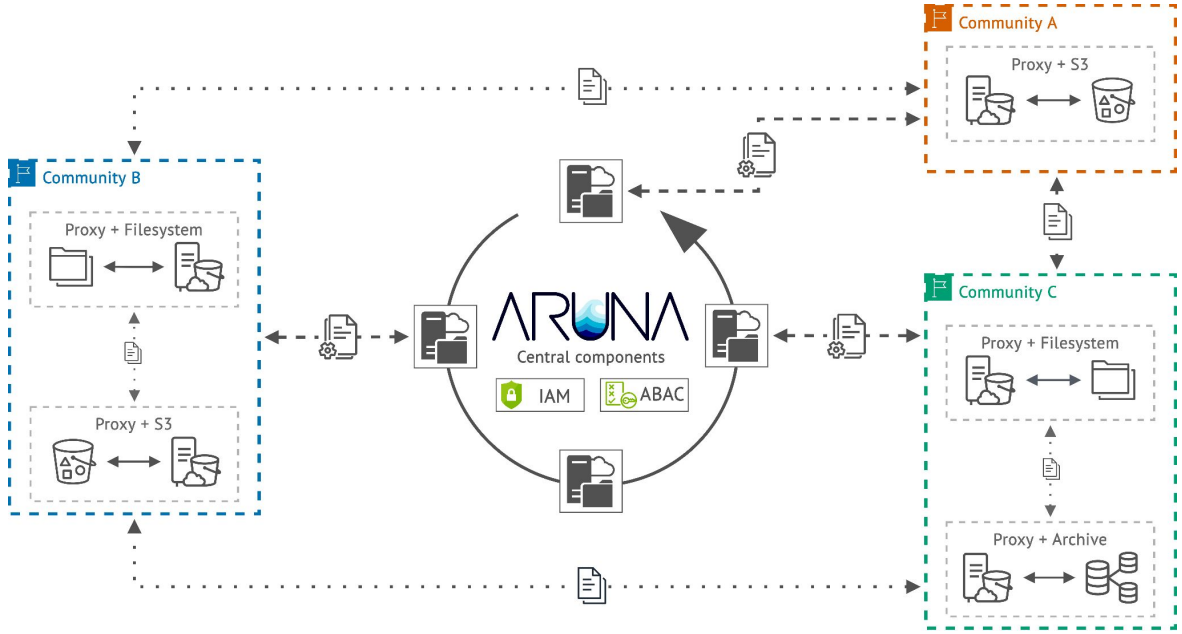
- **Data exchange (external): Data Space Protocol**
  - Contract negotiation via Open Digital Rights Language (ODRL)
  - Data catalog (DCAT)

- **Data exchange (internal): S3 compatible**

- **Service Catalog**
  - Discovery of services via Self Descriptions

Storage Architecture

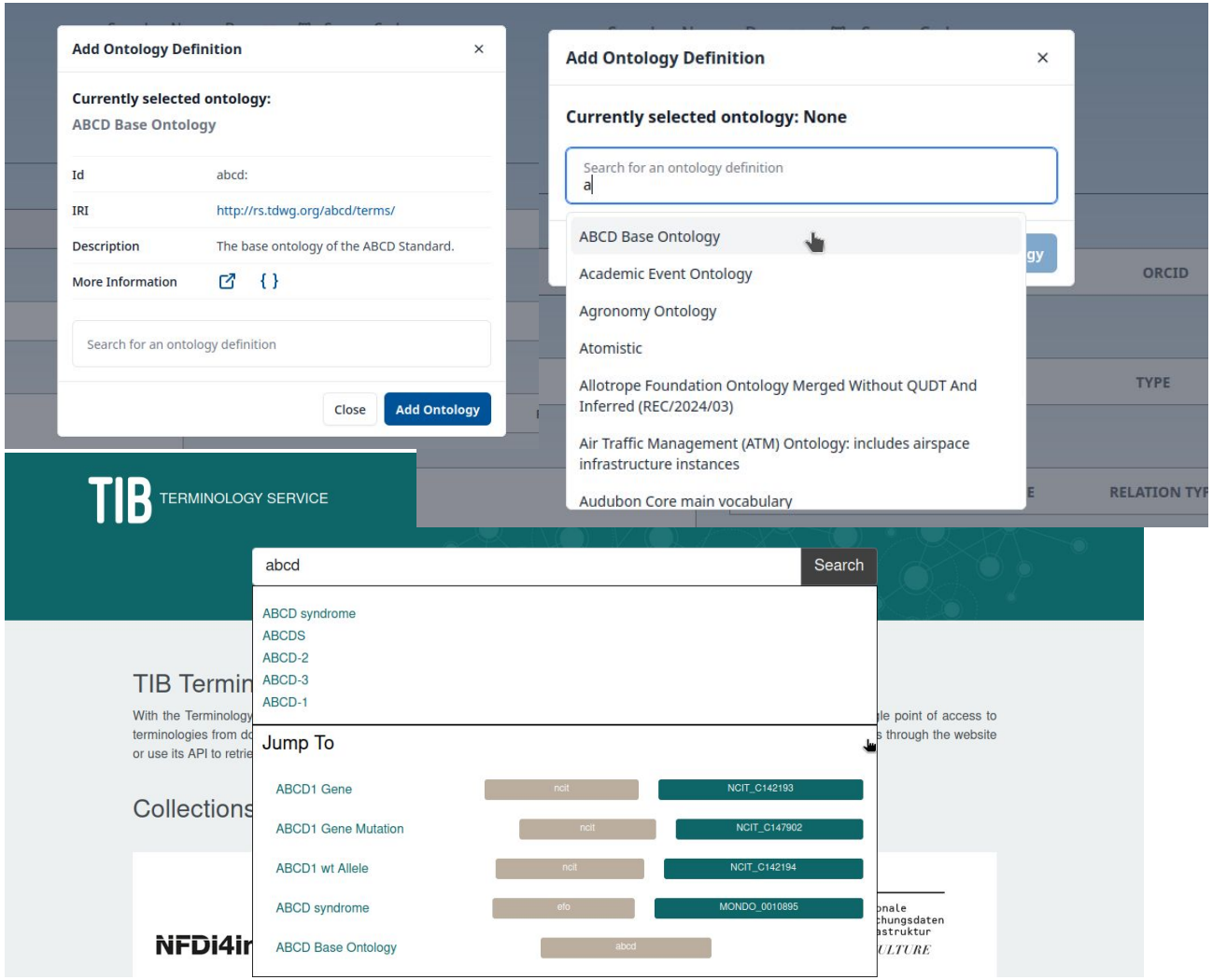
- **Aruna:** Storage engine developed by NFDI4Biodiversity and NFDI4Microbiota
- Integrated Data Catalog (Custom + DCAT)
- Standardized Metadata with PIDs and full-text search (schema.org)
- Unified (S3) interface for existing storage
- Event driven architecture
  - automated trigger of internal & external actions workflows
  - **Compliance Monitoring**
- Custom **EDC** compatible **DSP** integration



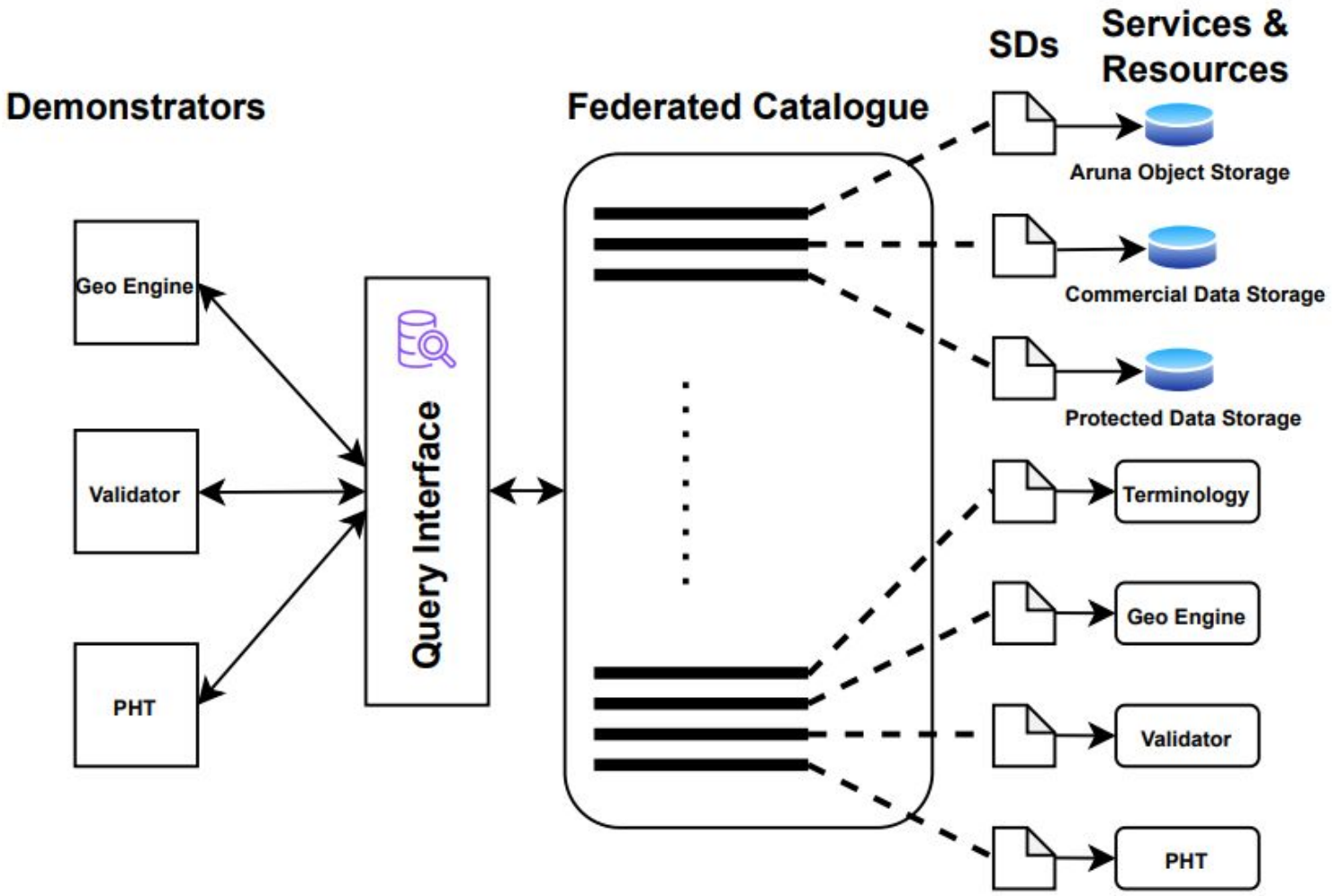
https://aruna-engine.org/

Terminology Service

- Service to query terminologies from various domains
- Integration into validator demonstrator and storage engine (aruna)
- Continuous developments in TS4NFDI



Demonstrator integration



Thank you for your interest!

FAIR Data Spaces

GEFÖRDERT VOM

Bundesministerium
für Bildung
und Forschung