

Strategies for the Integration of Metadata Catalogues, Data Publications and Archives

FWK Seminar, March, 2025

Oliver Knodel // contact: o.knodel@hzdr.de



Development of the HZDR

Open Data Ecosystem

Scientific Computing Department at HZDR — 2017

Establishment of the Scientific Computing division within the IT as a bridge to the science

Data Policy — 2018

Legal framework for the institutional data management at HZDR.

Data Repository RODARE — 2018

Provision of the Rosendorf Data Repository for data and software publications.



Scientific Data Management (and HPC) — since 2019

Analysis of data/service landscape at HZDR and development of an uniform data lifecycle.



HZDR Metadata Catalogue — 2024

Metadata catalogue for additional experiment-specific metadata as extension of RODARE.

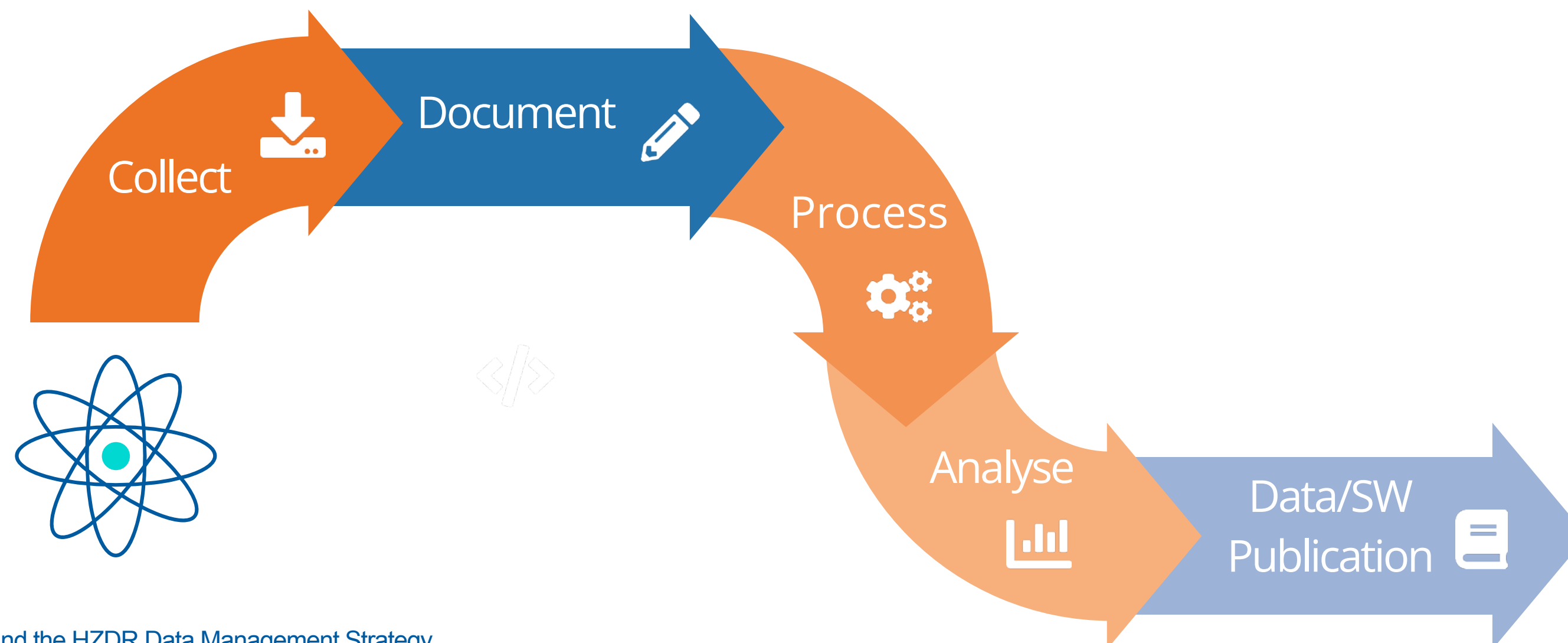


HZDR RDM Strategy — 2025+

Identification of the need to develop a data management strategy together with our scientists.

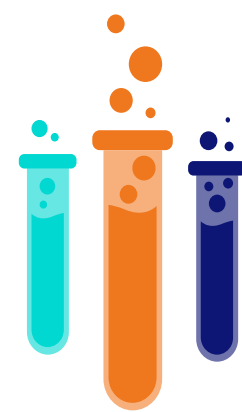
The Mission of the Scientific Data Management Group

- Supporting our scientists and research experiment with tools:
 - electronic lab books,
 - interactive analysis,
 - publication repositories for data and software,
 - scientific workflow management,
 - PID (handle) generation and management.
- Providing support in connecting experiments, detectors and diagnostics to our infrastructure.
- Establishment of analysis workflows and HPC applications.



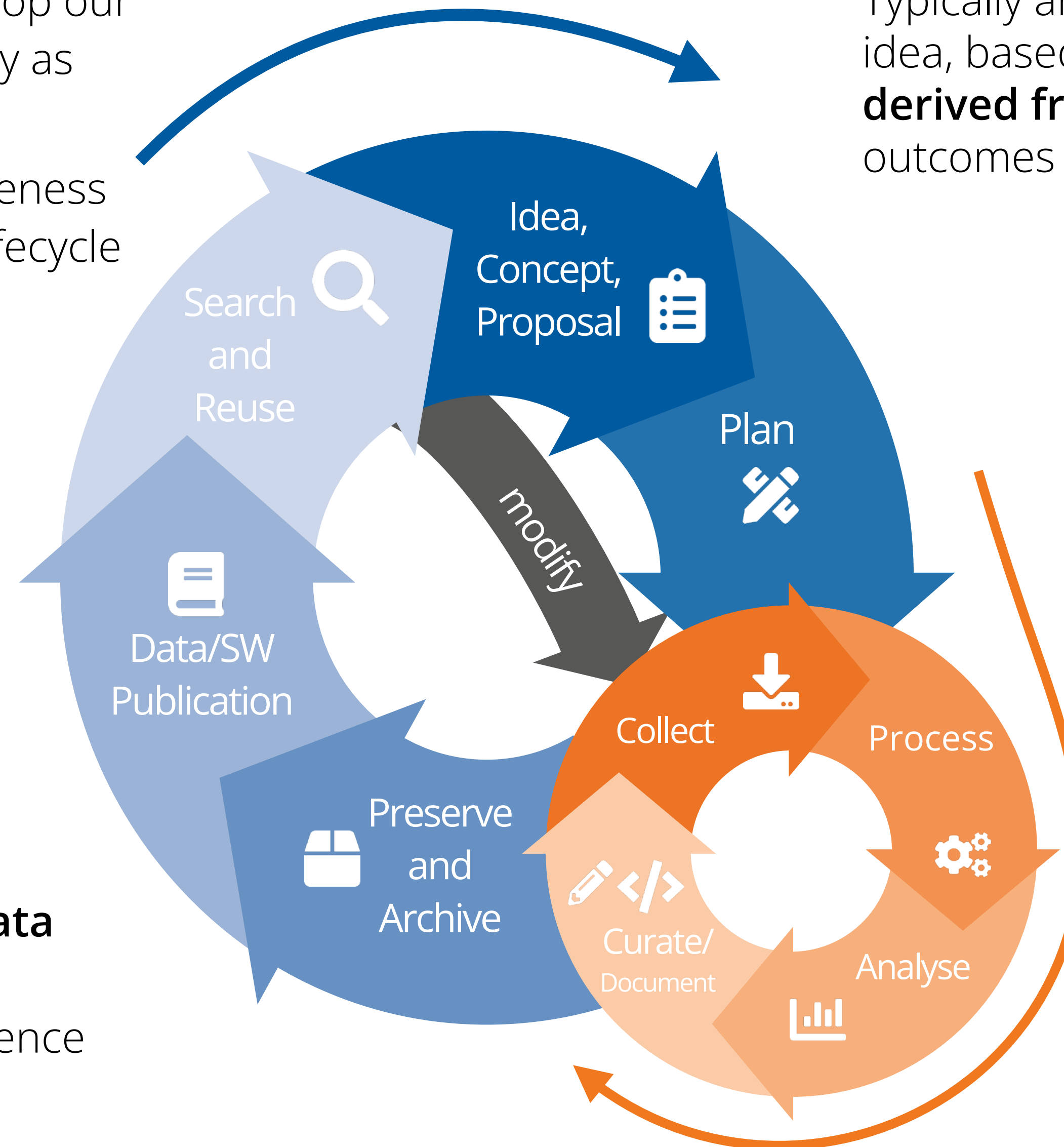
Data Management Lifecycle

- Necessary in order to develop our services as comprehensively as possible.
- Should also create an awareness for all components of the lifecycle



Publication of **data** and **metadata** is at the end of a typical research lifecycle and can be the starting point for new iteration cycles...

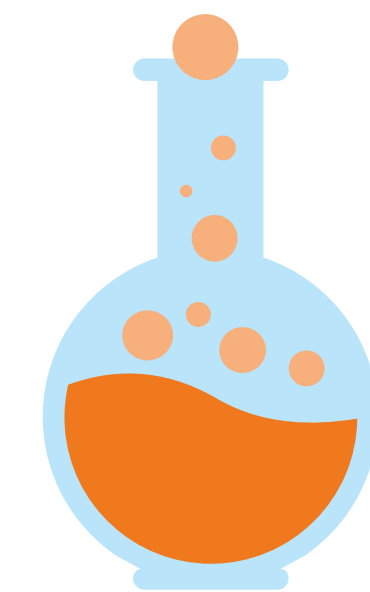
Archive or (additional) **data publication** to provide sustainable and open science




Typically an experiment starts with an idea, based on a research question, **derived from available research** outcomes

→ Valuable metadata in our proposal systems

During experiment **documentation** and **metadata** curation are essential for a **comprehensible** experiment



The Data and Software Publication Platform RODARE



Search

Upload

Communities

Log in


Recent uploads

September 24, 2024 (v1)

Dataset

Open Access

Data publication: Cavity-mediated thermal control of metal-to-insulator transition in 1T-TaS2

Jarc, Giacomo; Mathengattil, Shahla Yasmin; Montanaro, Angela; Giusti, Francesca; Rigoni, Enrico Maria; Sergo, Rudi; Fassioli, Francesca;  Winnerl, Stephan; Zilio, Simone Dal; Mihailovic, Dragan; Prelovšek, Peter; Eckstein, Martin; Fausti, Daniele

Original datasets corresponding to the publication.

Uploaded on September 24, 2024



View

September 19, 2024 (v1)

Dataset

Open Access

Heat flow data from the fungus Schizophyllum commune

 Fahmy, Karim; Günther, Alix; Bertheau, Rahel;  Pape, David

The data set contains three typical heat flow curves recorded from the fungus Schizophyllum commune and exemplifies the evaluation of such data by the software tool metabolator (<https://rodare.hzdr.de/record/3150>).

Uploaded on September 19, 2024







View

September 19, 2024 (v0.2.0)

Software

Open Access

METABOLATOR: Analysis of Microcalorimetric Metabolic Data Using Monod's Equation

 Pape, David;  Lokamani, Mani; Seal, Ayush;  Kelling, Jeffrey;  Knodel, Oliver;  Fahmy, Karim;  Juckeland, Guido

Curve fitting automation for metabolic load of bacteria in solutions.

Uploaded on September 19, 2024

2 more version(s) exist for this record

View

September 18, 2024 (v1)

Dataset

Open Access

Data publication: Formation of martensitic microstructure in epitaxial Ni-Mn-Ga films after fast cooling

View

RODARE Docs


Have a look at the restructured documentation of RODARE. We now can more easily notify about news and features, such as the search options in Rodare.

Visit us <https://rodare.hzdr.de/about>.

Featured HZDR Large-Scale Research Facilities

In order to make our large-scale HZDR facilities more visible, Rodare now offers specific communities and an overview on our front page!

Overview of published data sets of selected large-scale research facilities of the HZDR:

- **ATHENA** — Accelerator Technology Helmholtz iNfrAstructure
- **CARBOSOLA**
- **Dresden High Magnetic Field Laboratory**
- **DRESDYN** — DREsden Sodium facility for DYNamo and thermohydraulic studies
- **ELBE** — Electron Linac for beams with high Brilliance and low Emittance
 - ▶  ELBE Beamlines
- **Felsenkeller** — The underground ion accelerator lab
- **HECToR** — High-power ultrafast computed tomography
- **HIBEF** — Helmholtz International Beamline for Extreme Fields

Powered by:


Harvested via OAI-PMH by:



Registered in:

REGISTRY OF RESEARCH DATA REPOSITORIES

<http://doi.org/10.17616/R3BR40>



Metadata and Data

- In our data publication system the datasets are described via **DataCite** metadata to be FAIR.
- The DataCite metadata is attached to the DOI and harvested via portals, such as **B2Find**.
- In Theory: max 50 GB per file, max 100 GB per dataset (largest dataset: 7,2 TByte).

The screenshot shows the DataCite Fabrica interface for the DOI 10.14278/rodare.513. It displays the URL <https://rodare.hzdr.de/record/513> and provides options for metadata export (DataCite XML, DataCite JSON, Schema.org JSON-LD, BibTeX). The metadata is shown in a JSON format:

```
{
  "id": "https://doi.org/10.14278/rodare.513",
  "doi": "10.14278/RODARE.513",
  "url": "https://rodare.hzdr.de/record/513",
  "types": {
    "ris": "FIGURE",
    "bibtex": "misc",
    "citeproc": "graphic",
    "schemaOrg": "ImageObject",
    "resourceType": "Photo",
    "resourceTypeGeneral": "Image"
  },
  "creators": [
    {
      "name": "Podlipec, Roc",
      "givenName": "Roc",
      "familyName": "Podlipec",
      "affiliation": [
        {
          "name": "Helmholtz-Zentrum Dresden - Rossendorf"
        }
      ]
    }
  ]
}
```



The screenshot shows the RODARE dataset page for the DOI 10.14278/rodare.513. The page title is "Chronic Inflammation Prediction for Inhaled Particles, the Impact of Material Cycling and Quarantining in the Lung Epithelium". It includes a preview of a microscopy image, a list of files (e03_s02_t02_F10_LA4_CellMask_Tio2-Alexa_longTermExp_10X000.tif, e03_s02_t02_F10_LA4_CellMask_Tio2-Alexa_longTermExp_60X000.tif), and a list of versions (Version 1, Sep 8, 2020). The page also displays the publication date (September 8, 2020), the DOI, and the license (Creative Commons Attribution 4.0 International).

File Access:
Rest API



Filesystem

Metadata API:
Rest API, OAI-PMH





OSCARS

Open Science Clusters' Action
for Research & Society

Funded Project

NAPMIX - Nuclear, Astroparticle, and Particle Physics Metadata Integration for eXperiments

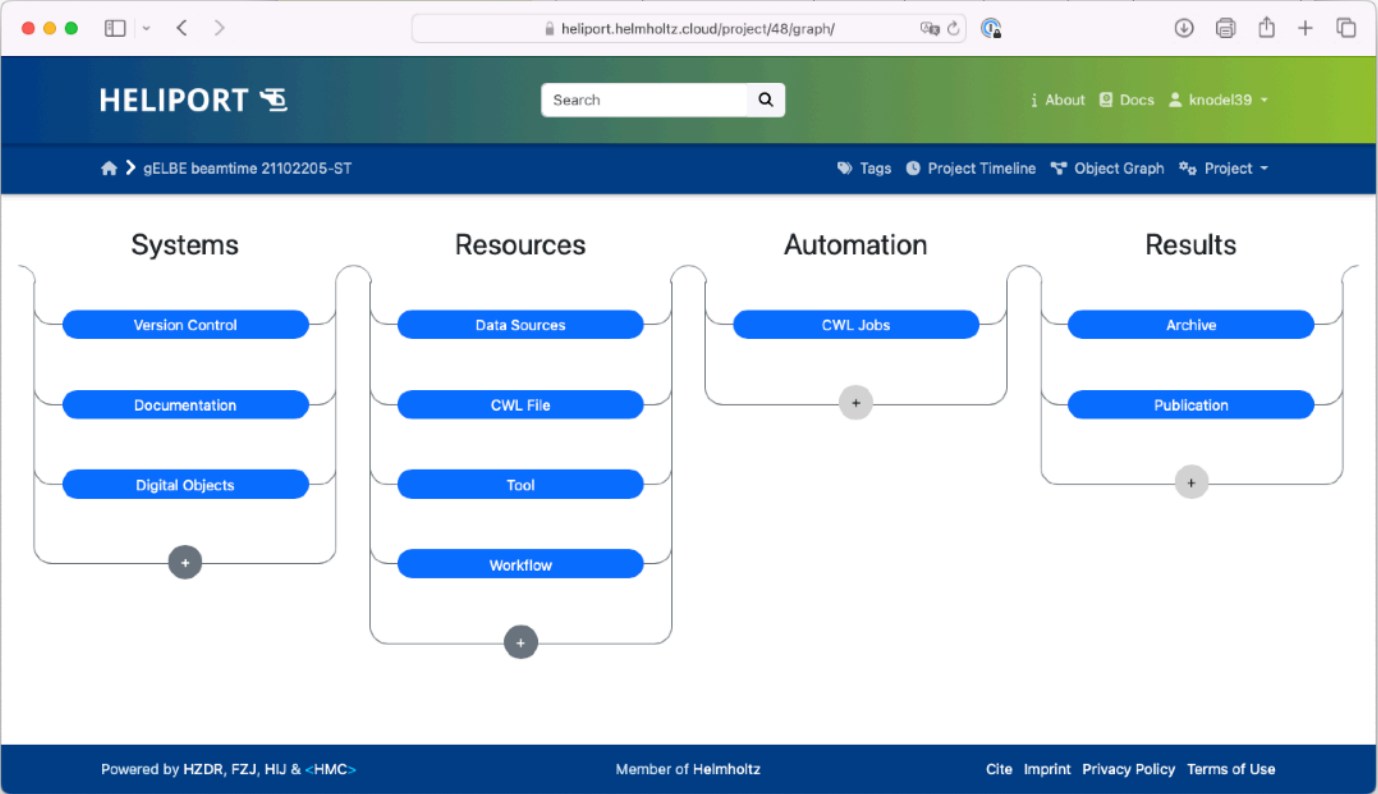
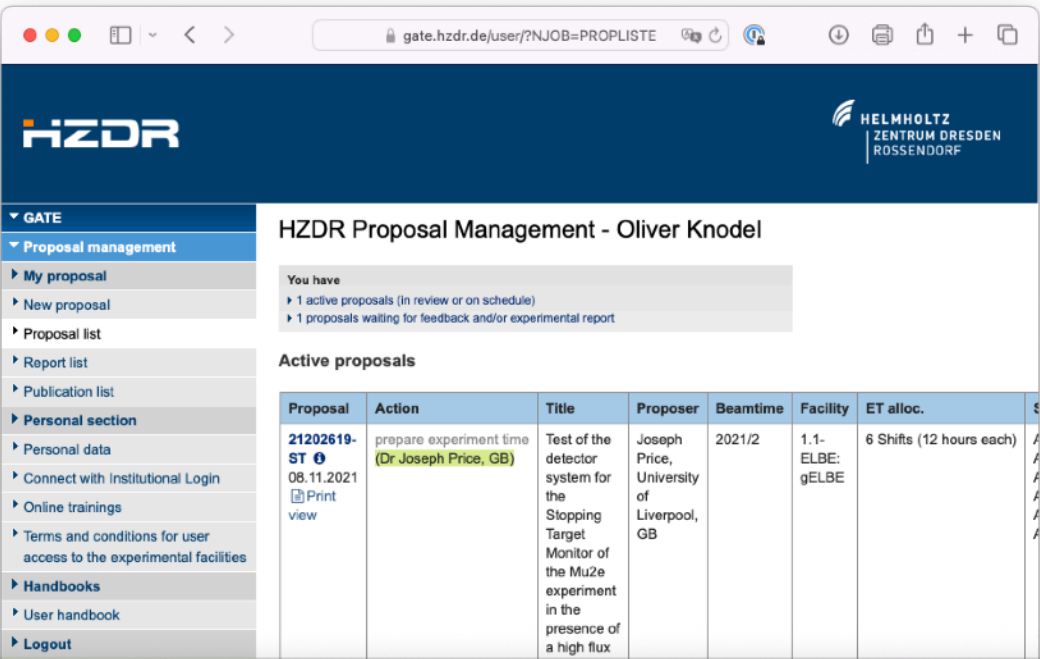
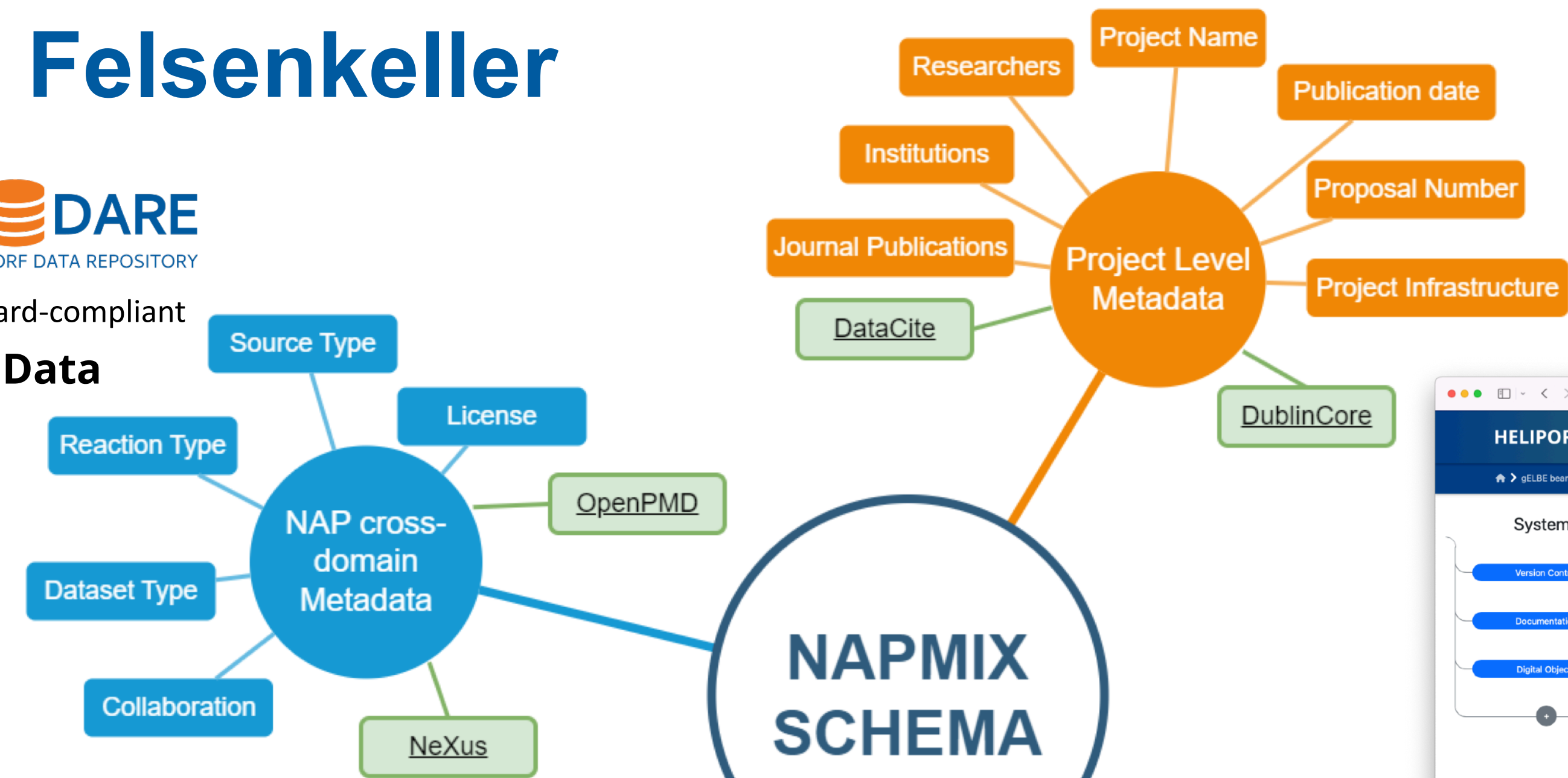
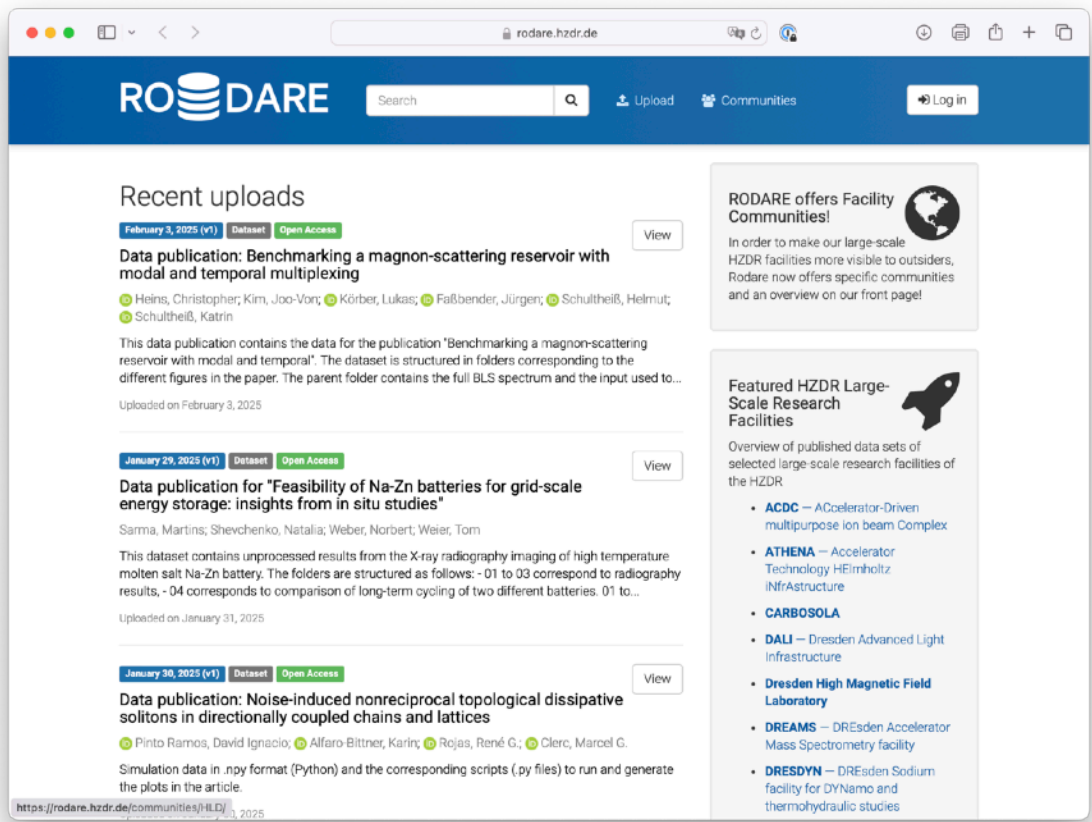


Implemented by:

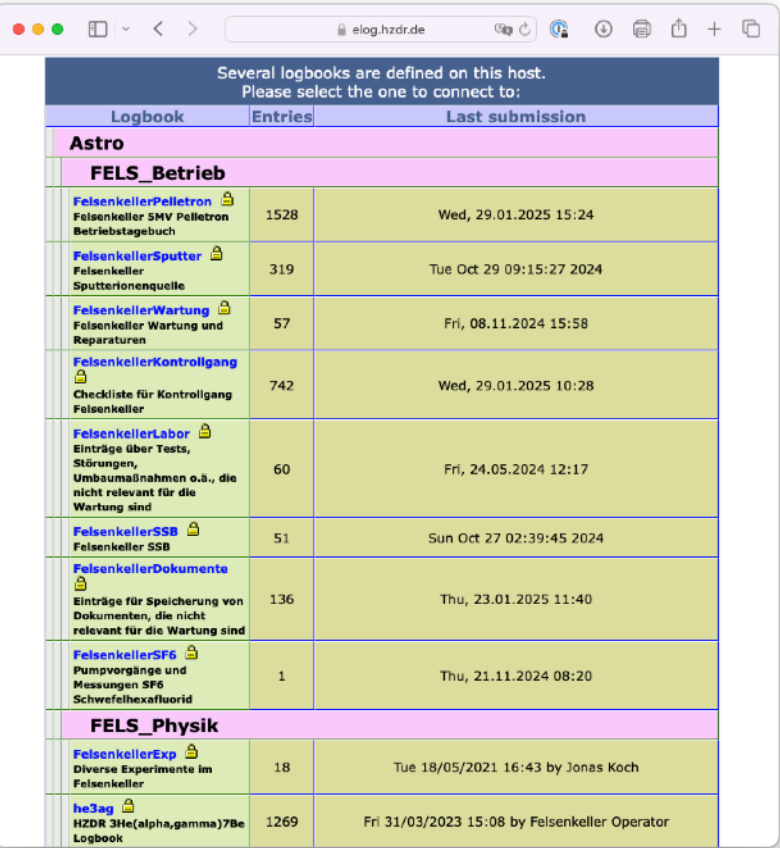
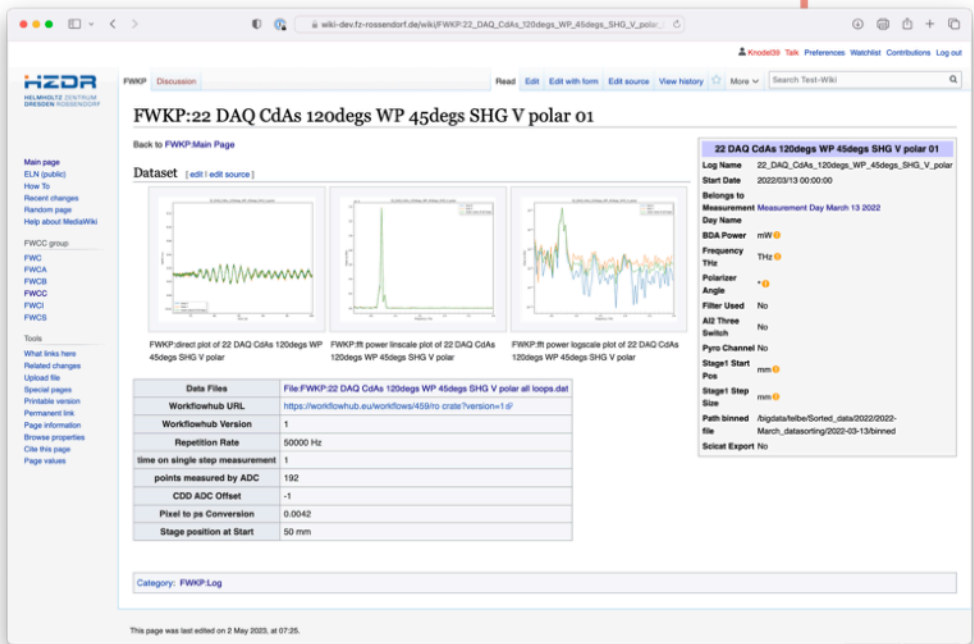
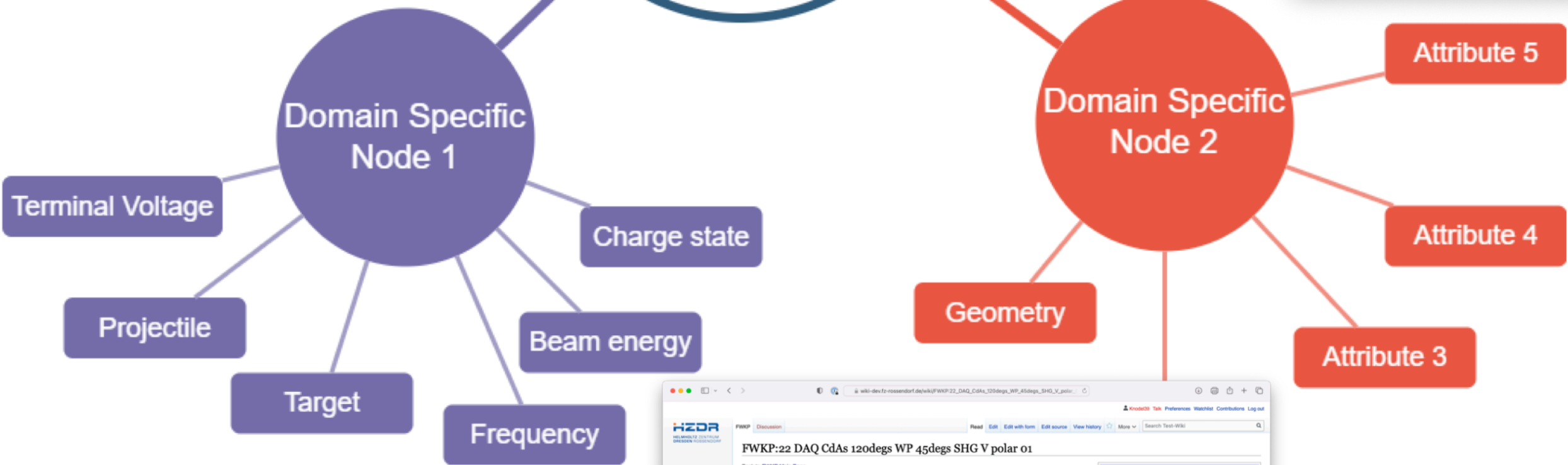


Funded by
the European Union

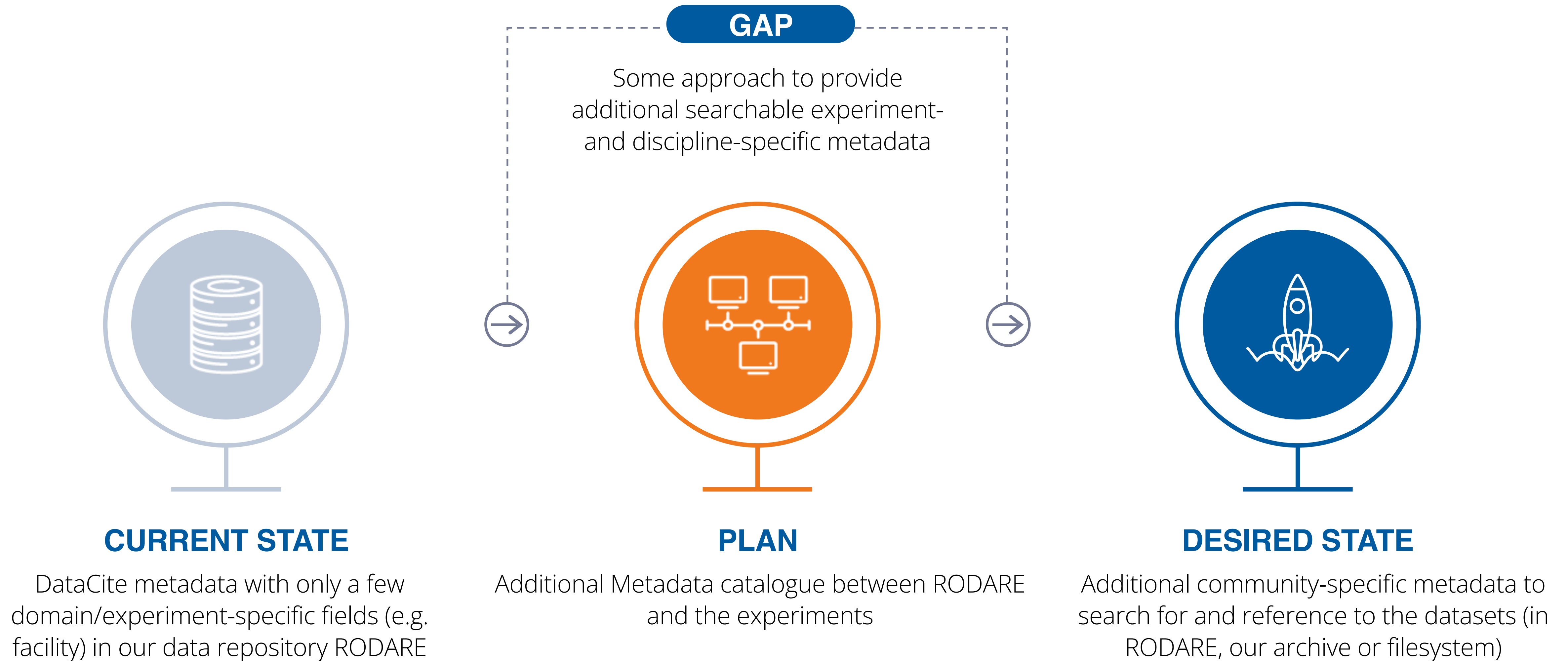
NAPMIX at the Felsenkeller



?



Gap Analysis of our Metadata Ecosystem

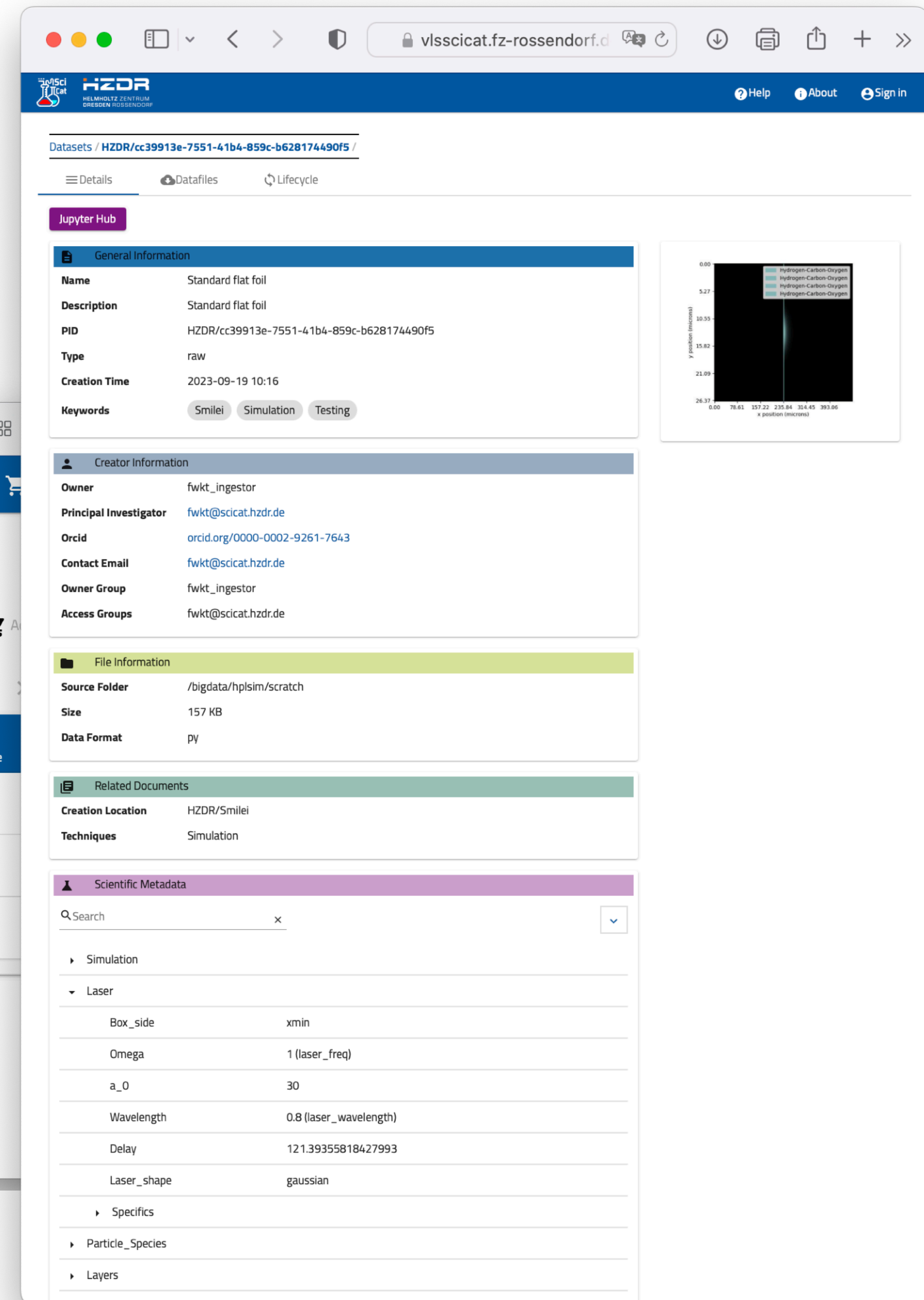
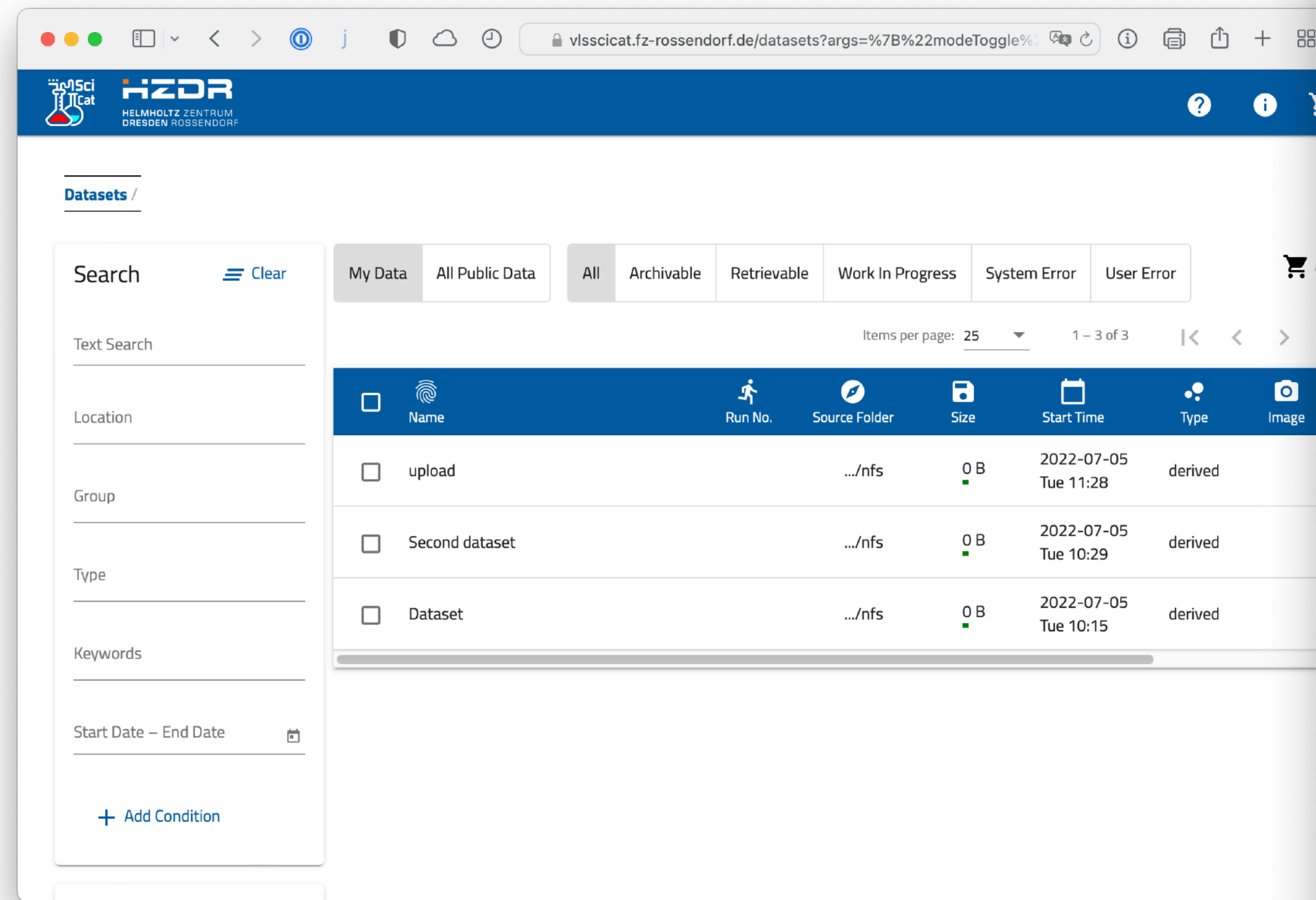
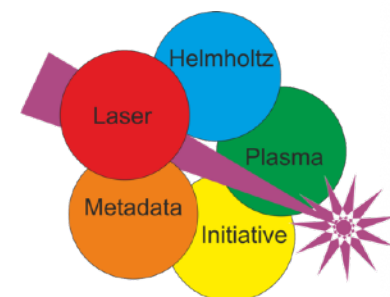


Metadata Catalogue SciCat

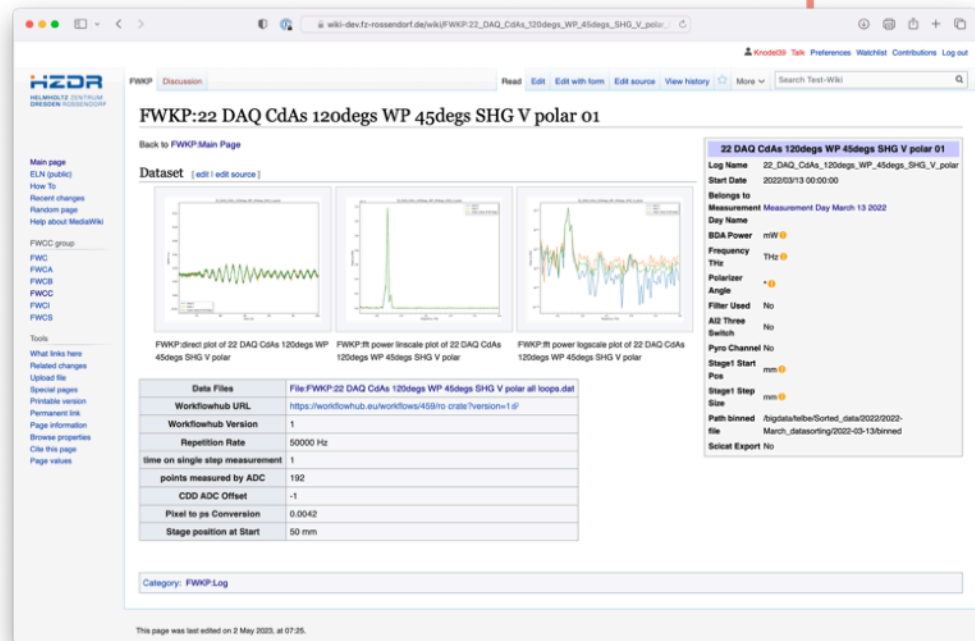
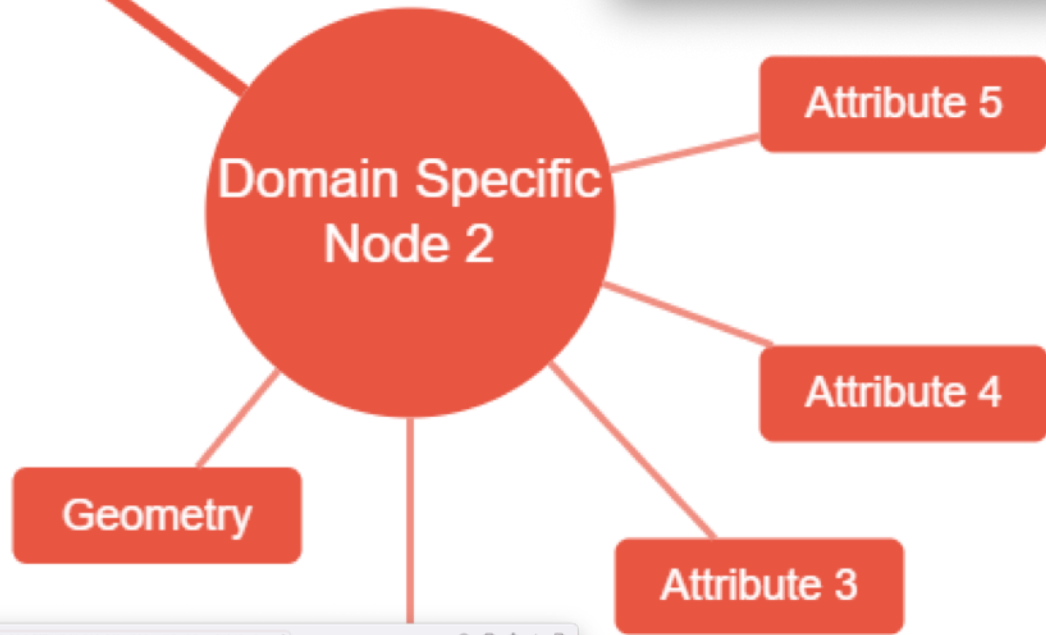
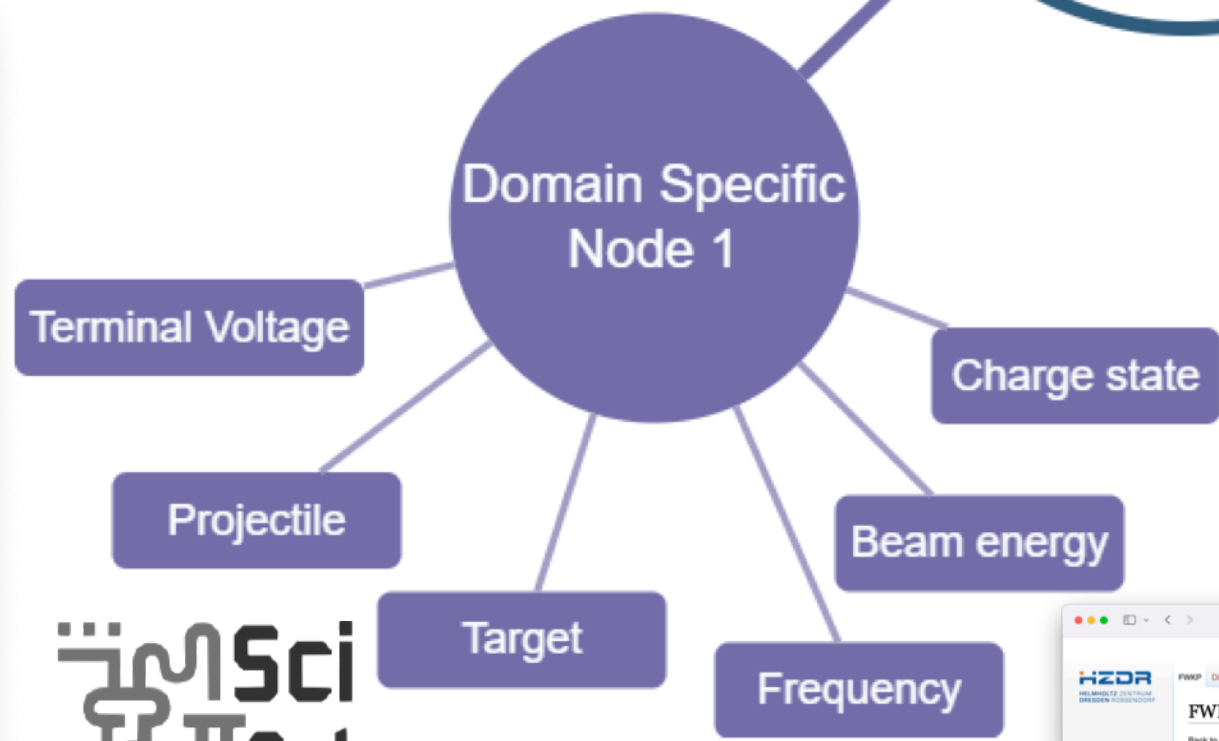
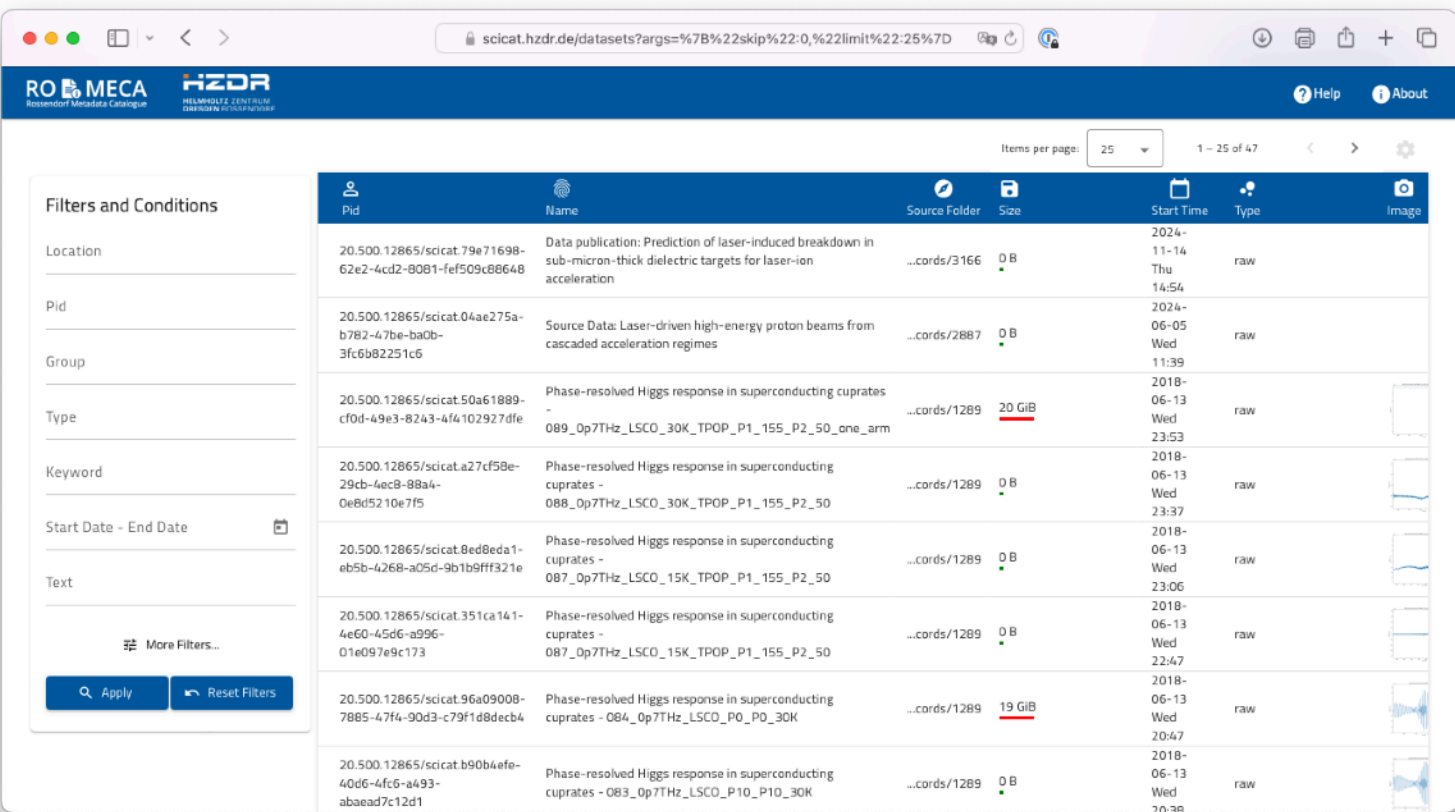
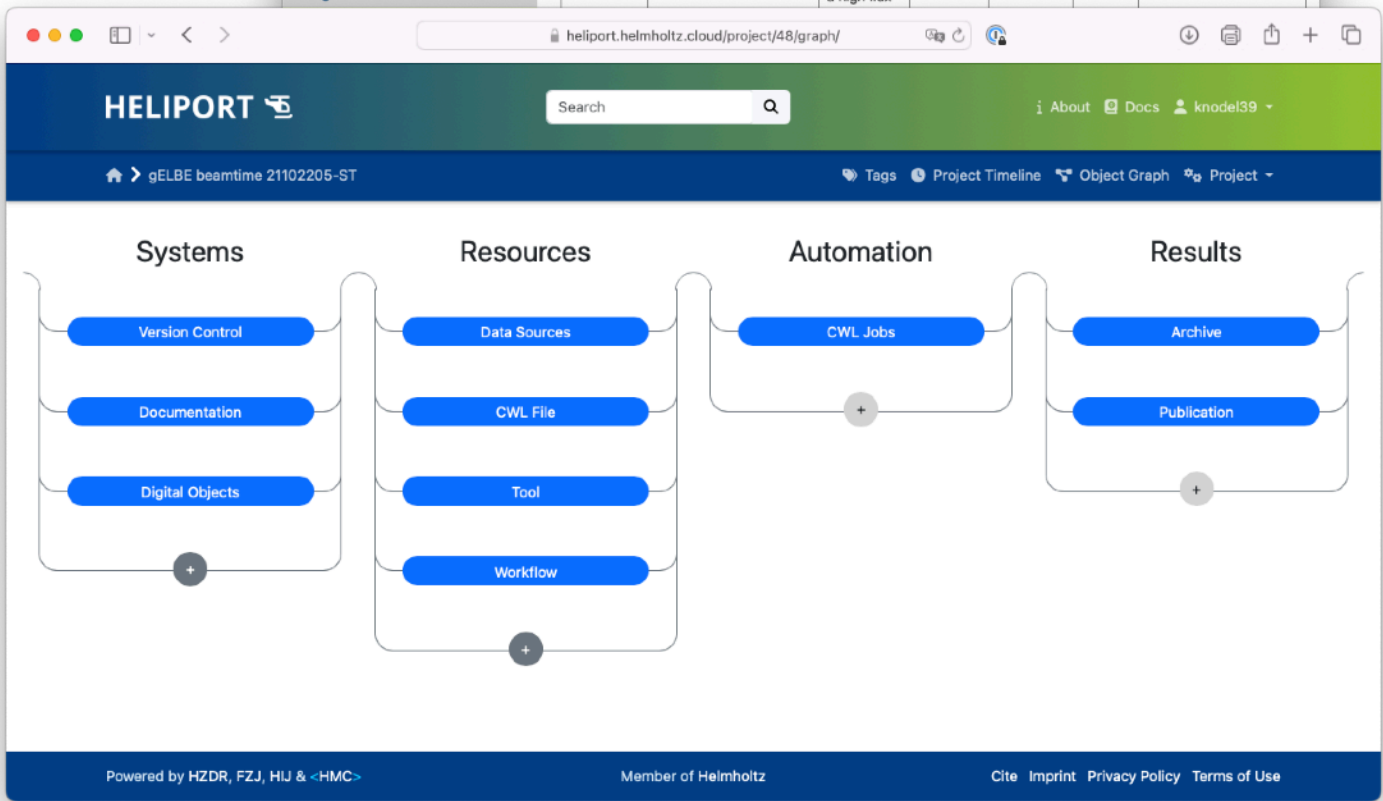
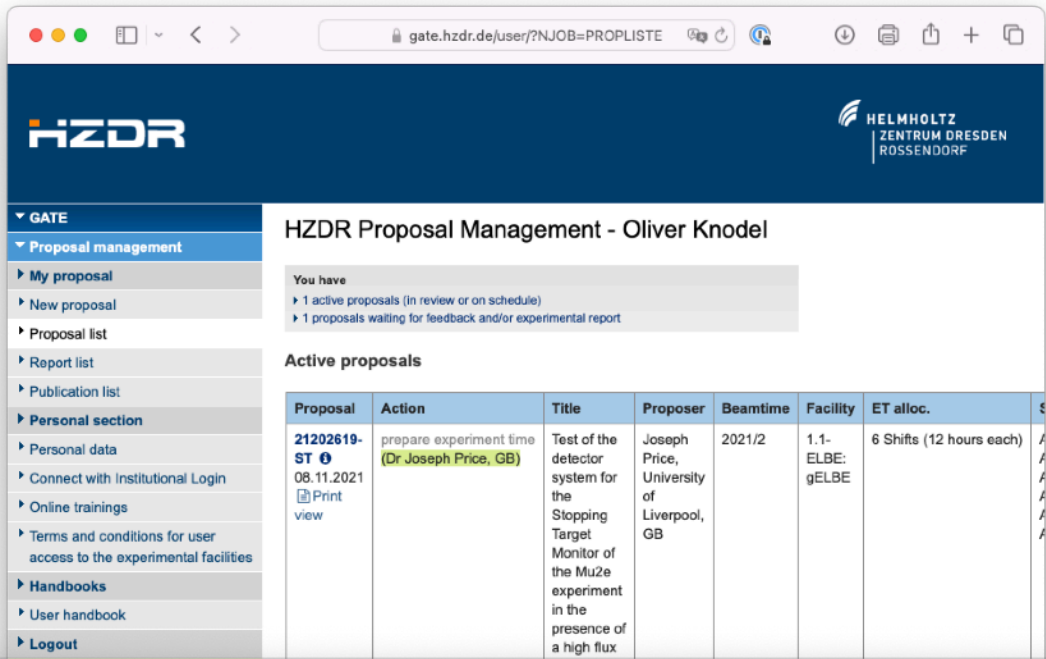
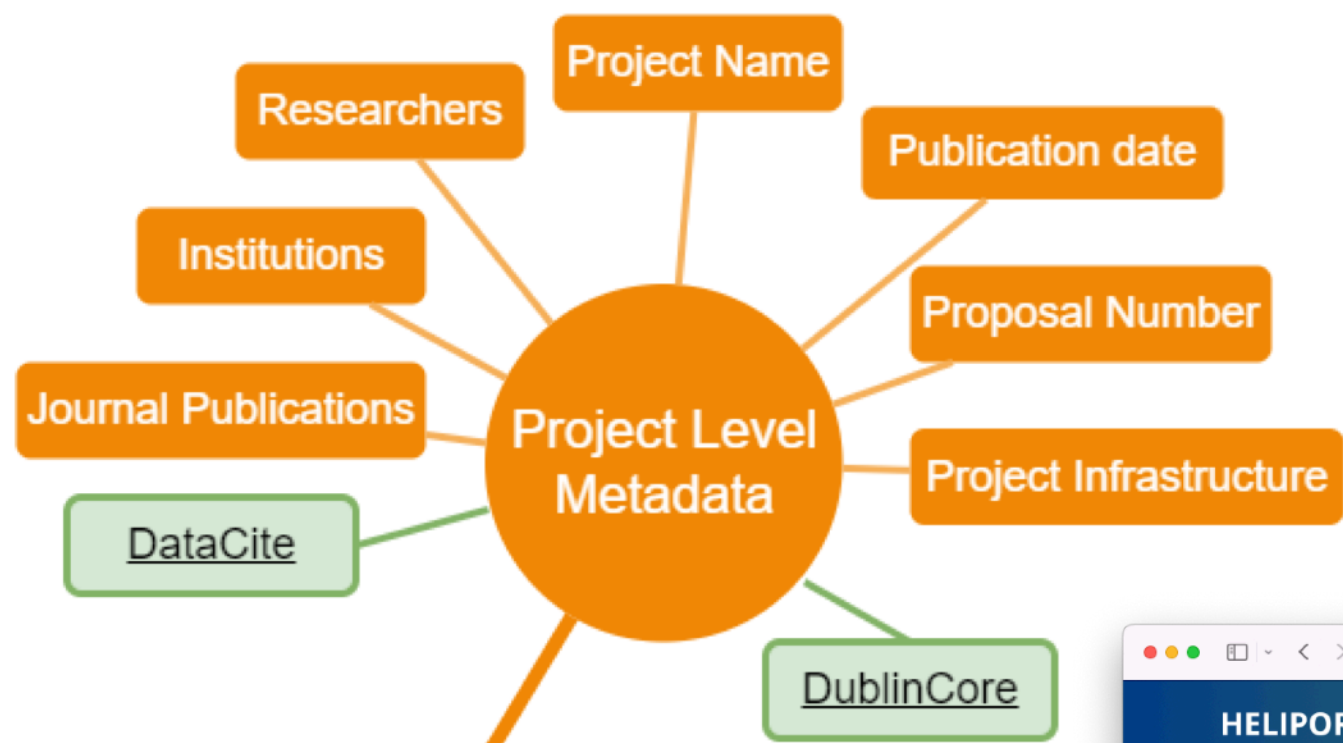
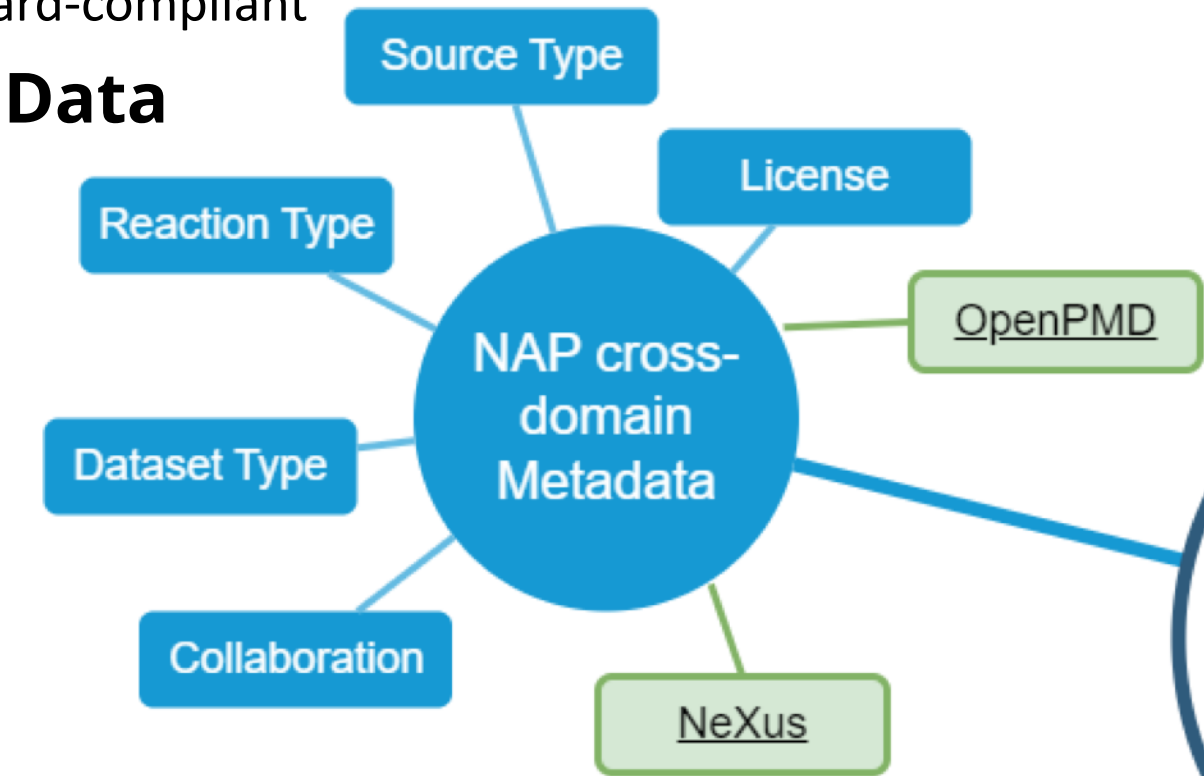
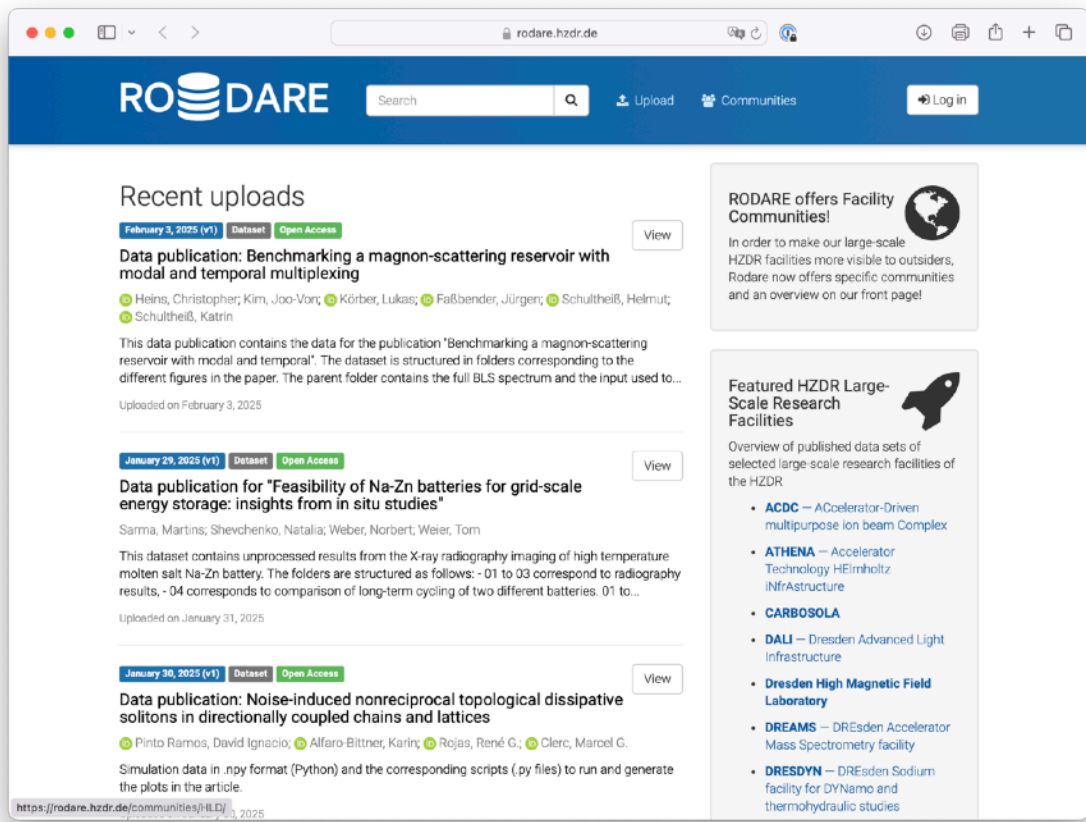
- SciCat catalogue provides experiment-specific metadata.
- Developed in in the projects ExPANDS, PANOSC and DAPHNE.



ExPANDS



NAPMIX at the Felsenkeller

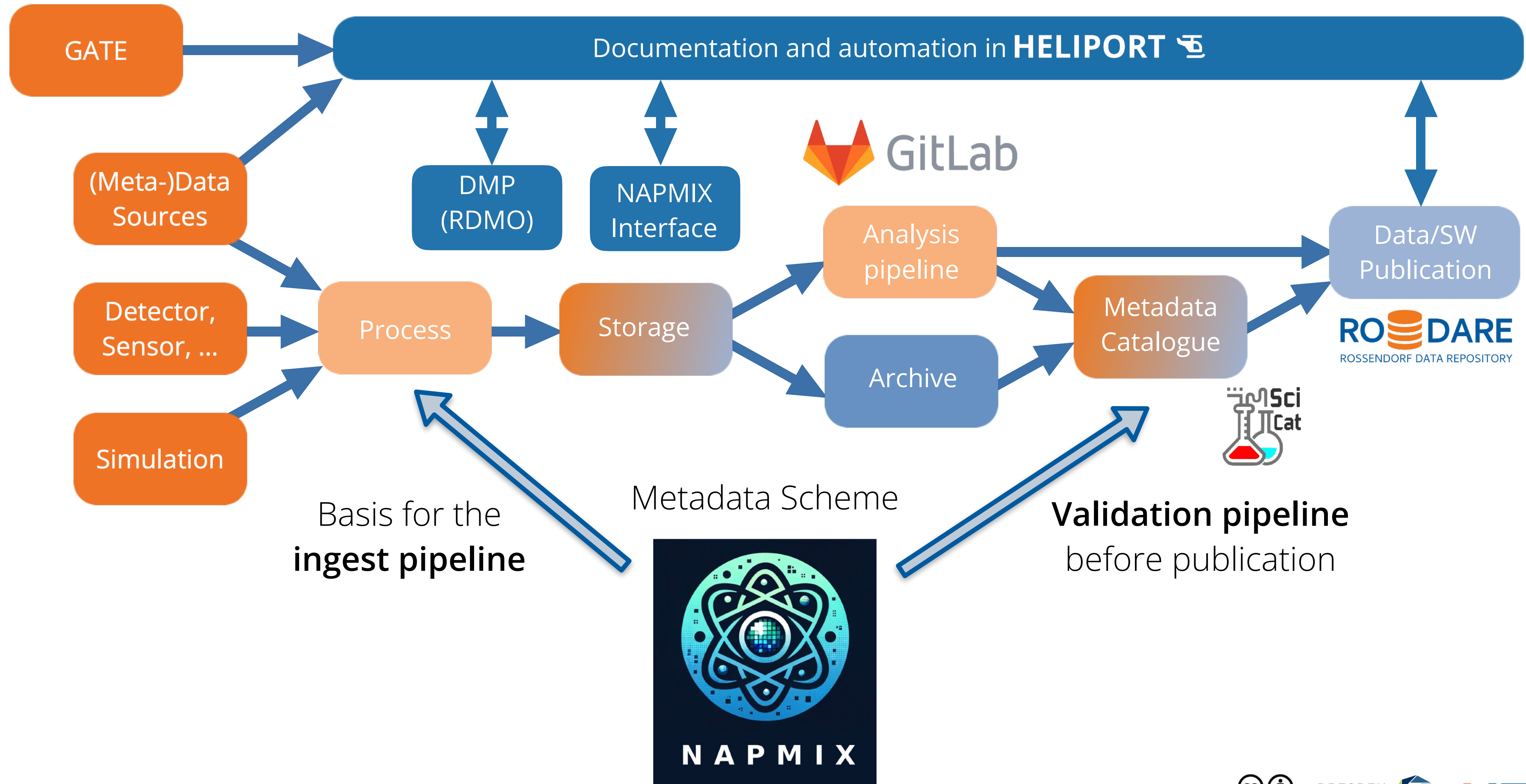


Several logbooks are defined on this host. Please select the one to connect to:		
Logbook	Entries	Last submission
Astro		
FELS_Betrieb		
FelsenkellerProduktion	1528	Wed, 29.01.2025 15:24
Felsenkeller SHV Felsenkeller Betriebstagebuch		
FelsenkellerSputter	319	Tue Oct 29 09:15:27 2024
FelsenkellerSputterentonnequelle		
FelsenkellerWartung	57	Fri, 08.11.2024 15:58
FelsenkellerWartung und Reparaturen		
Checkliste für Kontrollgang Felsenkeller	742	Wed, 29.01.2025 10:28
FelsenkellerLabor	60	Fri, 24.05.2024 12:17
Einträge über Tests, Störungen, Umbaumaßnahmen o.ä., die nicht relevant für die Wartung sind		
FelsenkellerSSB	51	Sun Oct 27 02:39:45 2024
FelsenkellerSSB		
FelsenkellerDokumente	136	Thu, 23.01.2025 11:40
Einträge für Speicherung von Dokumenten, die nicht relevant für die Wartung sind		
FelsenkellerS56	1	Thu, 21.11.2024 08:20
Pumpvorgänge und Messungen S56 Schwefelhexafluorid		
FELS_Physik		
FelsenkellerExp	18	Tue 18/05/2021 16:43 by Jonas Koch
Diverse Experimente im Felsenkeller		
heTag	1269	Fri 31/03/2023 15:08 by Felsenkeller Operator
HZDR 3He(alpha,gamma)7Be Logbook		

standard-compliant Metadata



NAPMIX in the Open Science Ecosystem of HZDR



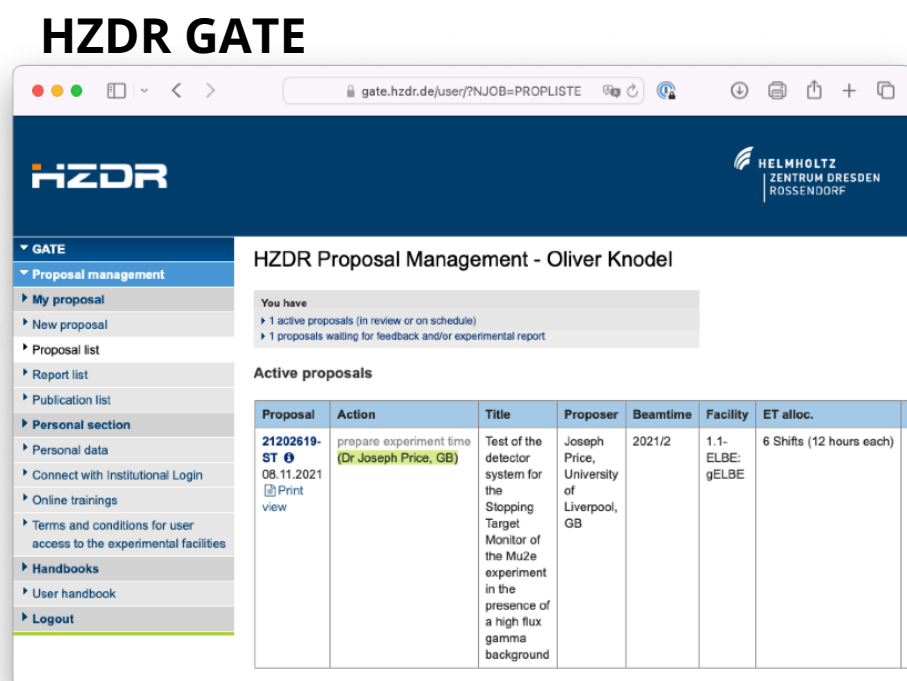
Embedding NAPMIX in the HZDR Data Management Chain

(Curated) Metadata Sources

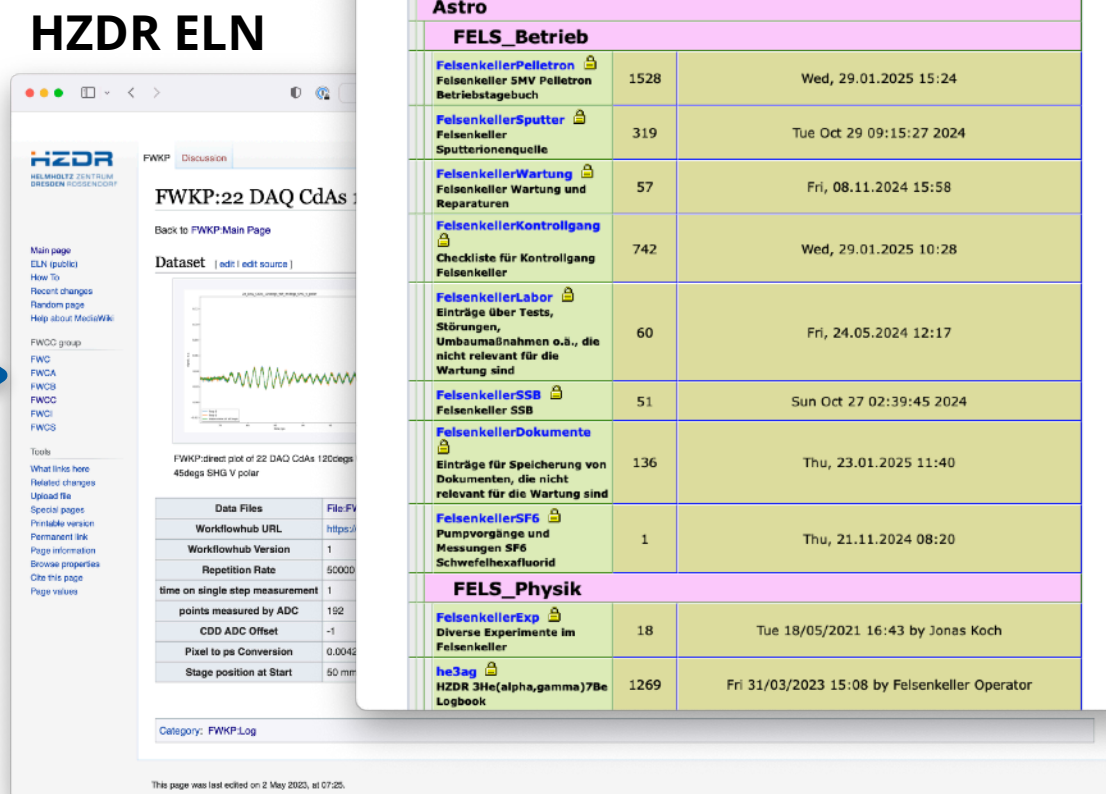
Public Metadata

Subsequent Access to Data

Proposal in GATE

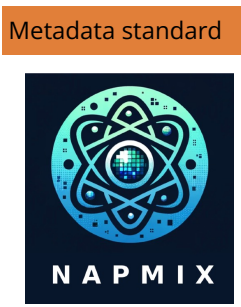
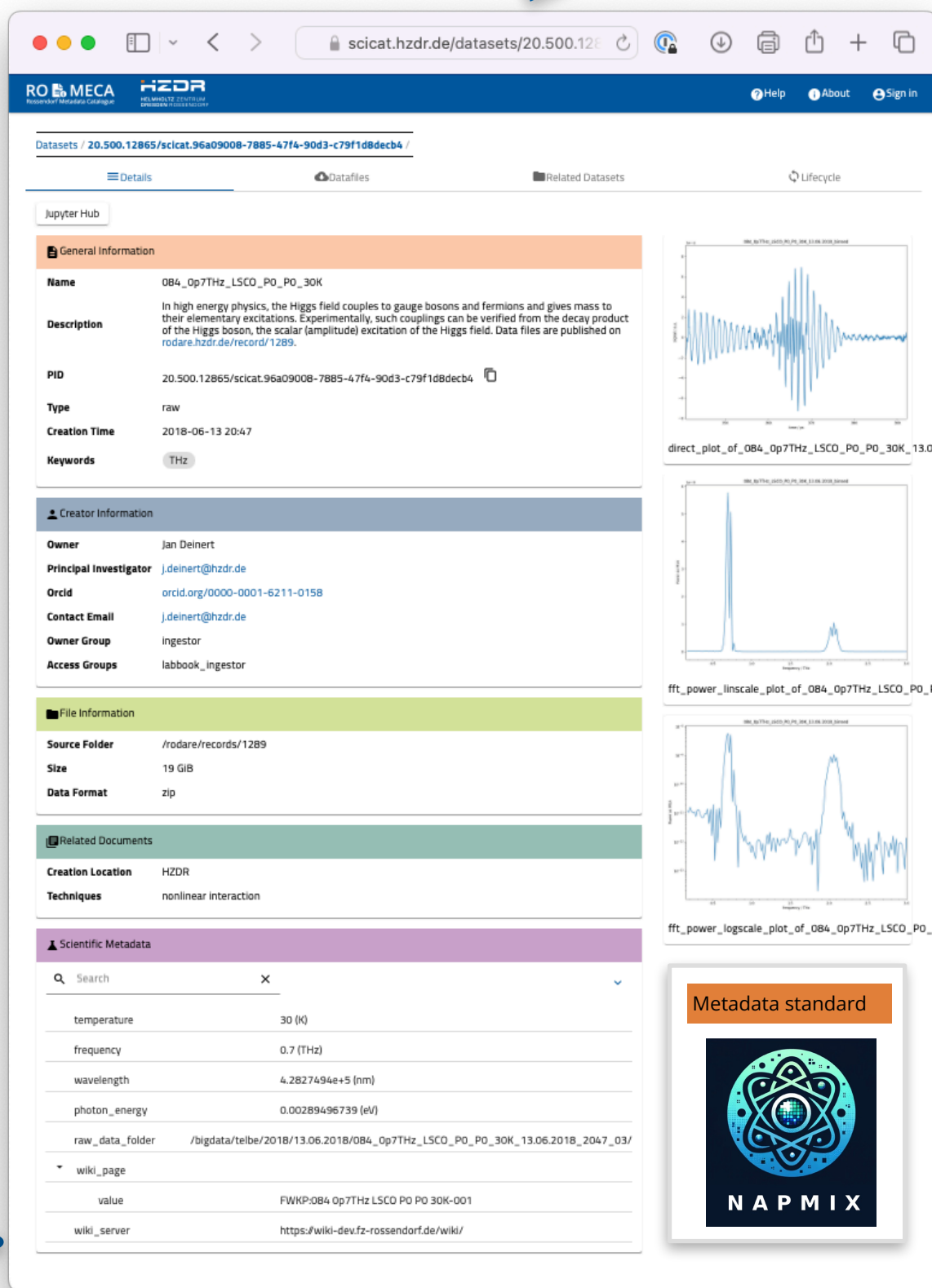


Metadata from Experiment



Automated Ingest Pipelines

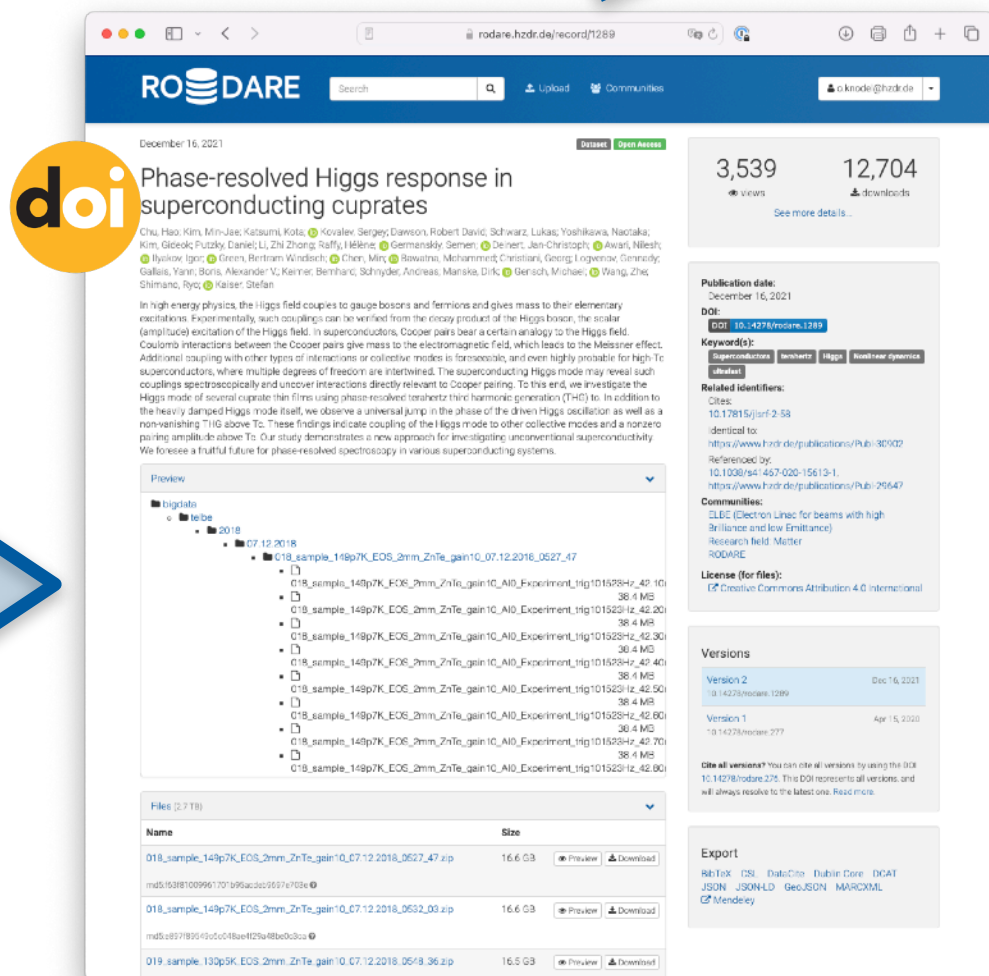
Catalogue



Access: Public



RODARE (Invenio)



Dataset



Tape Archive



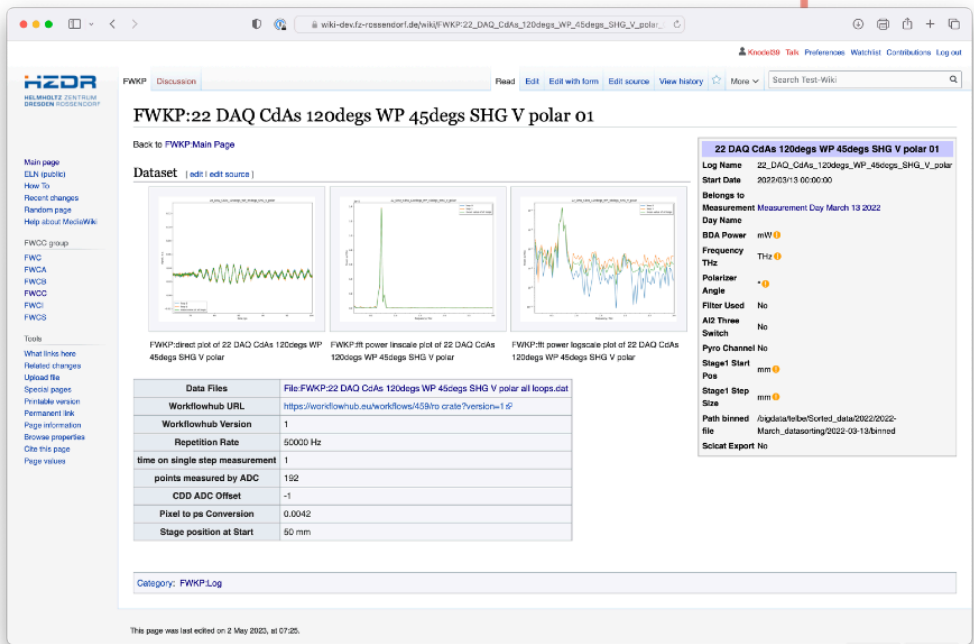
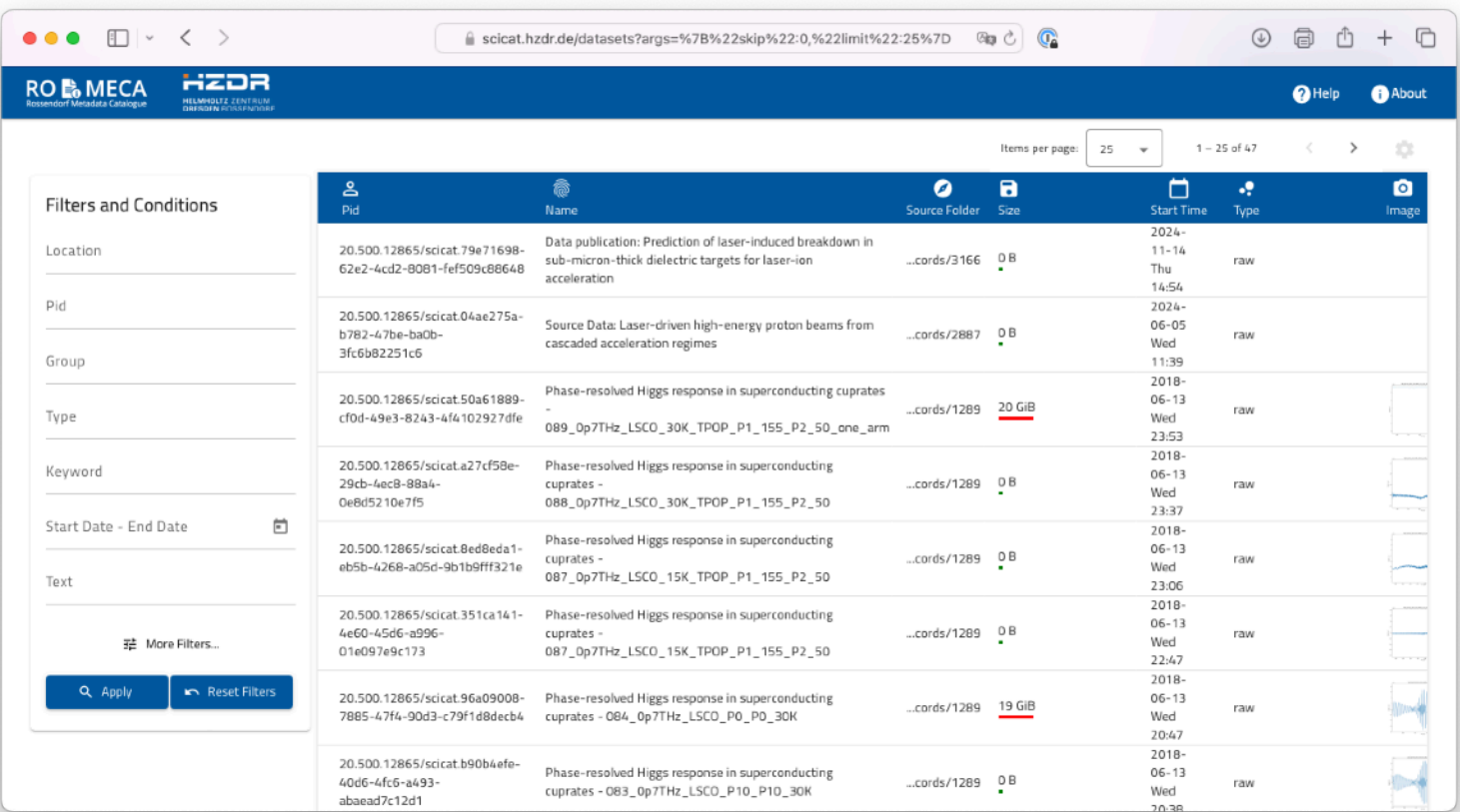
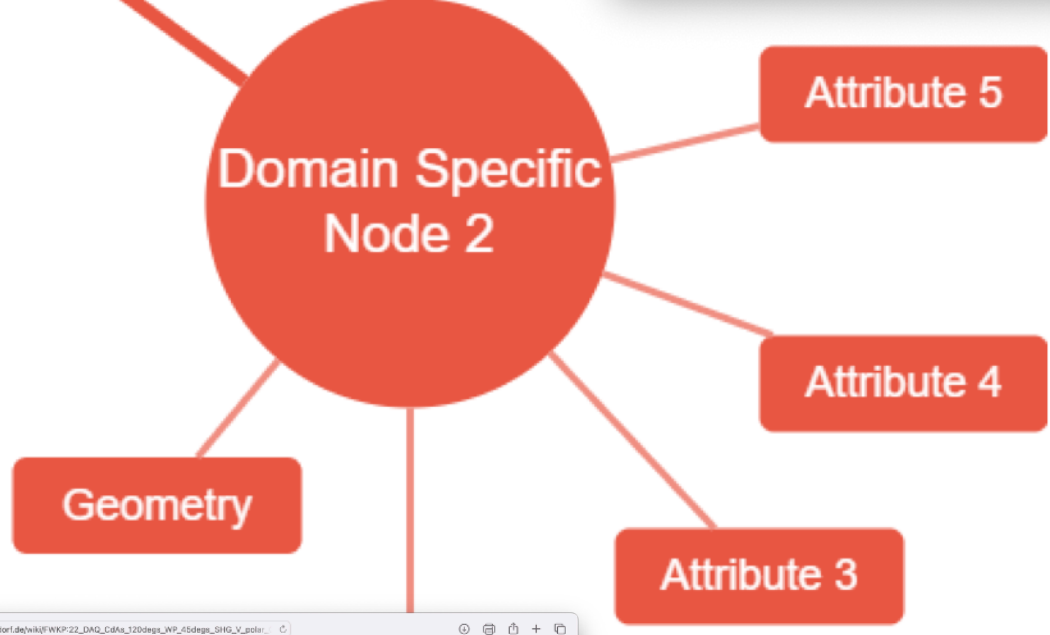
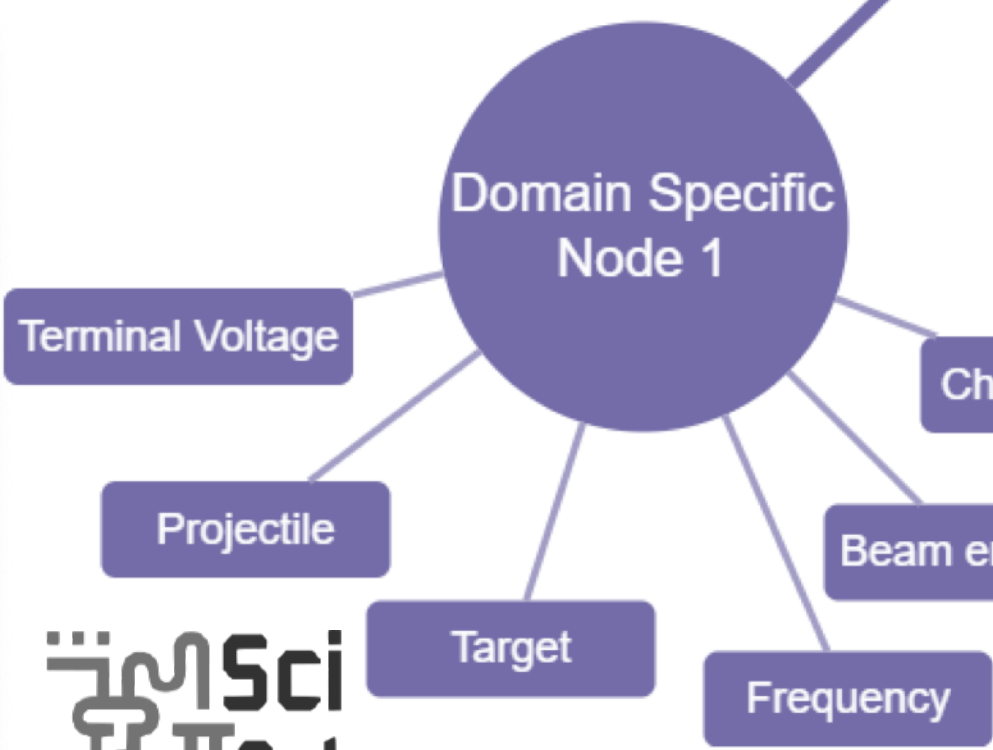
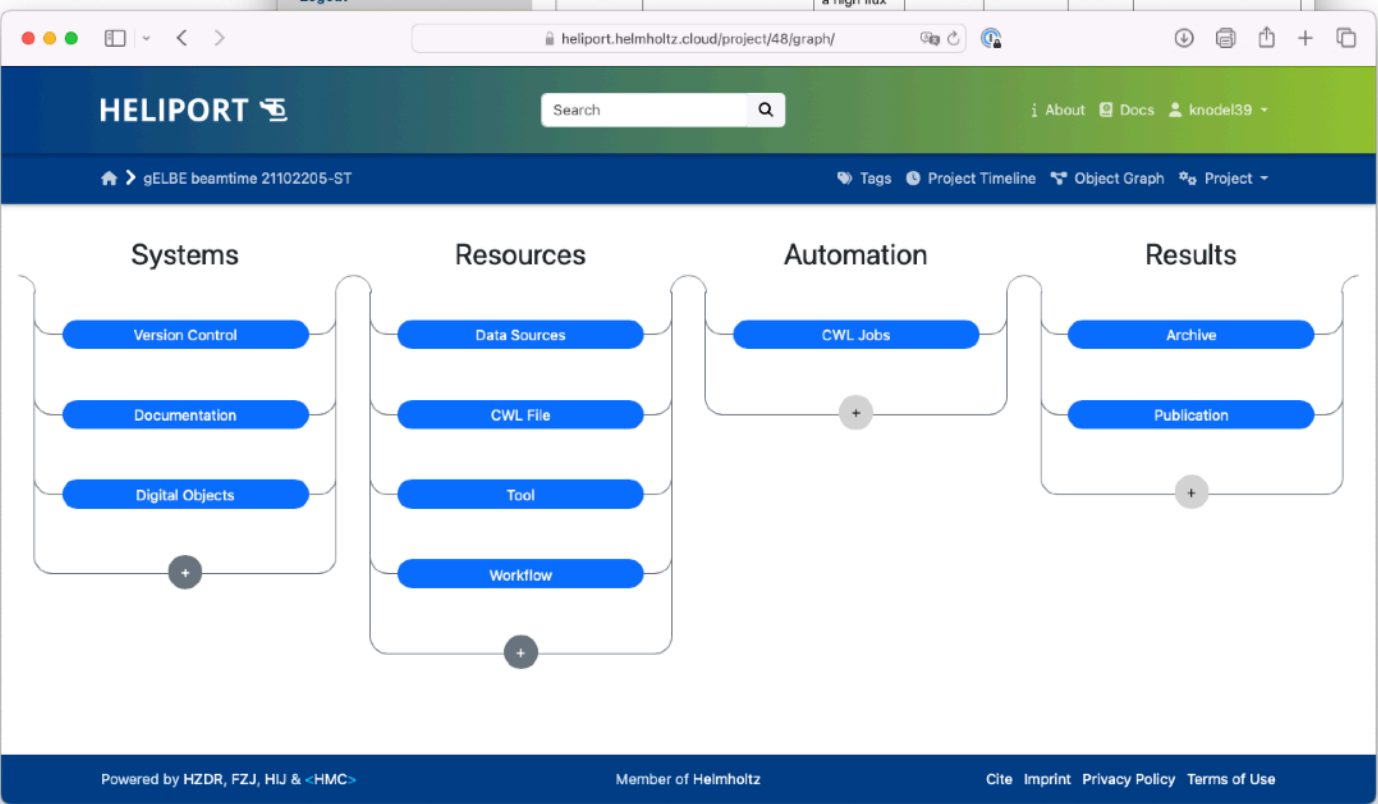
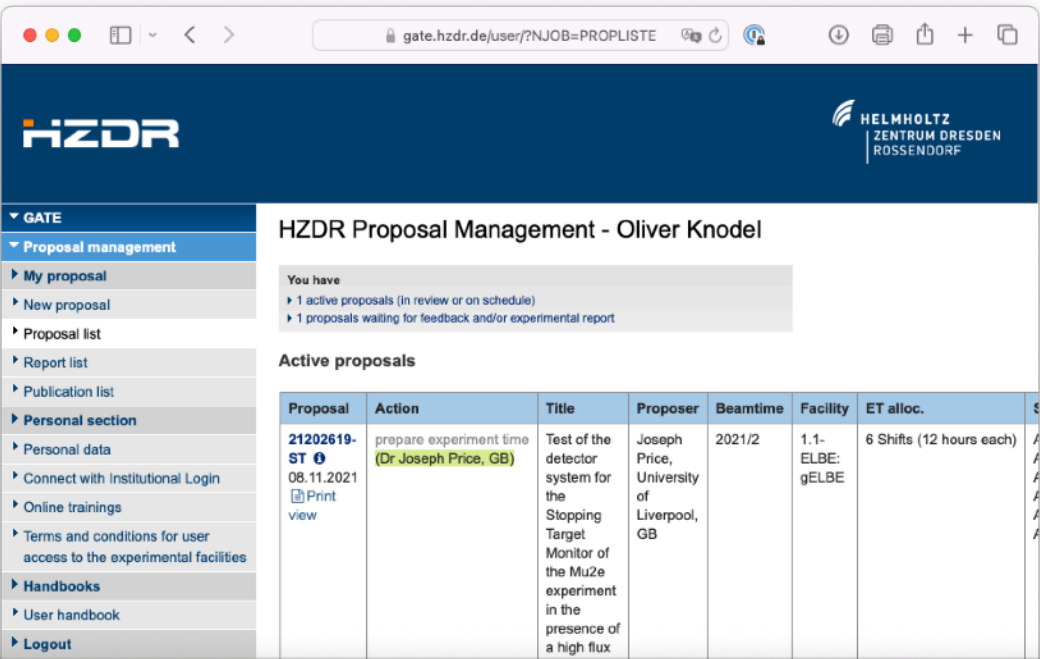
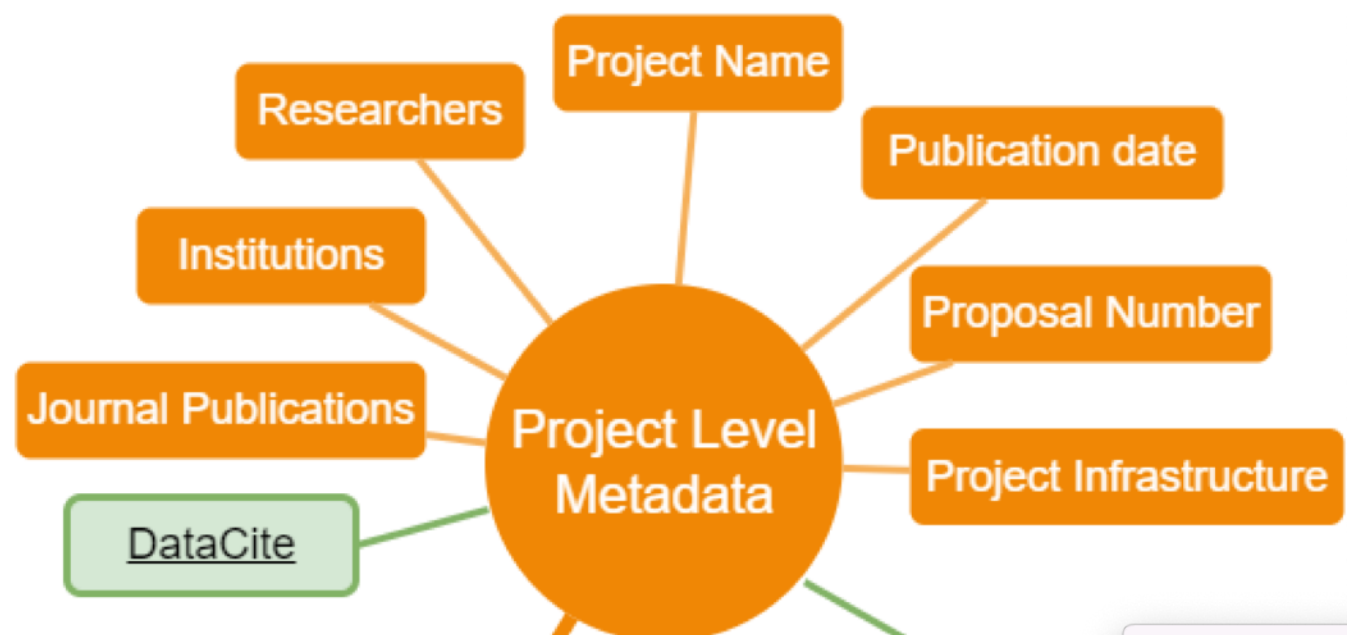
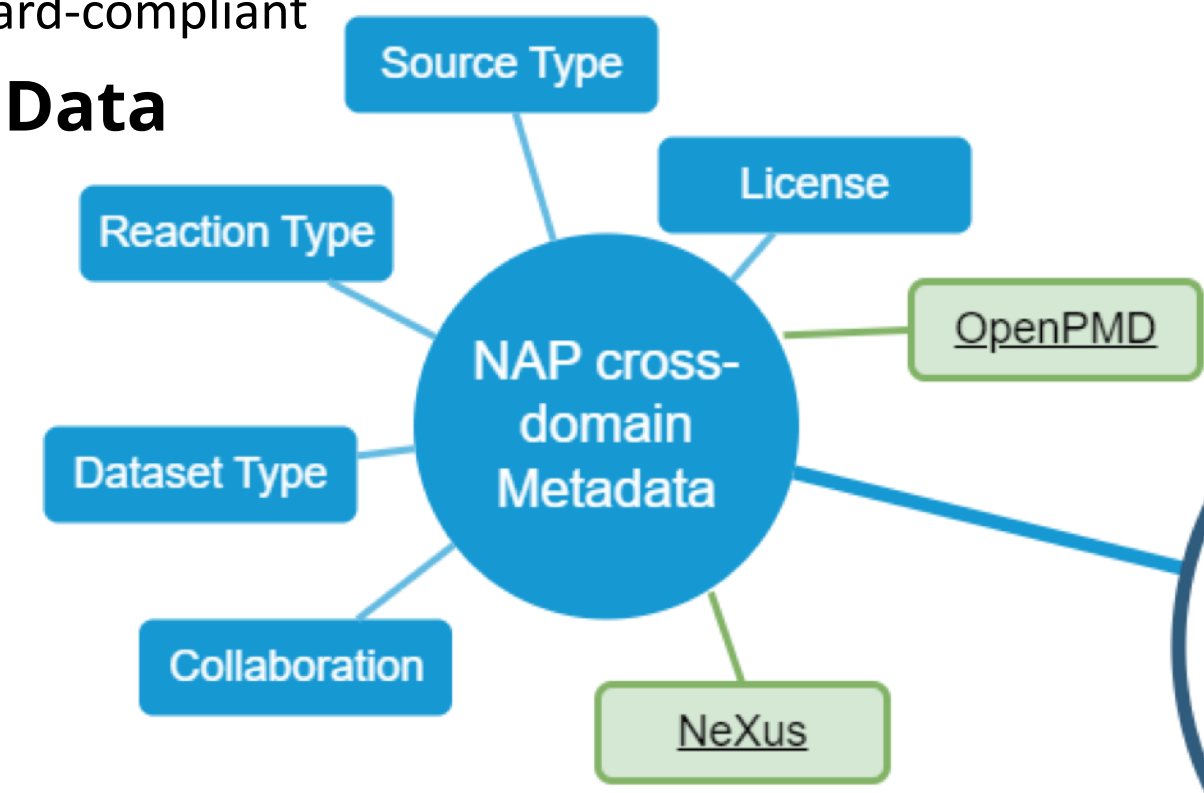
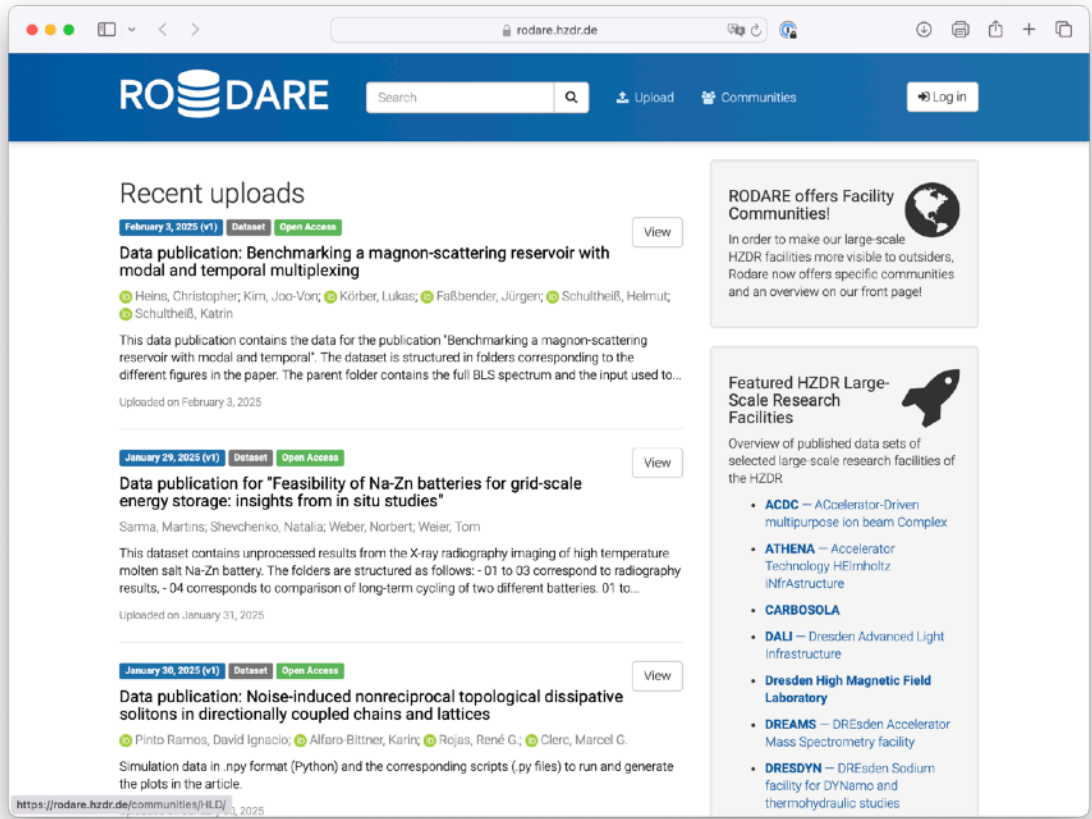
Filesystem



Access: Private/Internal



NAPMIX at the Felsenkeller



Several logbooks are defined on this host. Please select the one to connect to:		
Logbook	Entries	Last submission
Astro		
FELS_Betrieb		
FelsenkellerBetriebsrat	1528	Wed, 29.01.2025 15:24
FelsenkellerSHV FelsenkellerBetriebsratgebuch		
FelsenkellerSputter	319	Tue Oct 29 09:15:27 2024
FelsenkellerSputterentomnequelle		
Felsenkellerwartung	57	Fri, 08.11.2024 15:58
Felsenkellerwartung und Reparaturen		
FelsenkellerKontrollgang	742	Wed, 29.01.2025 10:28
Checkliste für Kontrollgang Felsenkeller		
FelsenkellerLabor	60	Fri, 24.05.2024 12:17
Einträge über Tests, Störungen, Umbaumaßnahmen o.ä., die nicht relevant für die Wartung sind		
FelsenkellerSSB	51	Sun Oct 27 02:39:45 2024
FelsenkellerSSB		
FelsenkellerDokumente	136	Thu, 23.01.2025 11:40
Einträge für Speicherung von Dokumenten, die nicht relevant für die Wartung sind		
FelsenkellerSFS	1	Thu, 21.11.2024 08:20
Pumpvorgänge und Messungen SFS Schwefelhexafluorid		
FELS_Physik		
Diverse Experimente im Felsenkeller	18	Tue 18/05/2021 16:43 by Jonas Koch
heTag		
HZDR Site(alpha,gamma)79e Logbook	1269	Fri 31/03/2023 15:08 by Felsenkeller Operator

standard-compliant Metadata

