

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



Our Research Facility and our Large Scale Research Infrastructures

The Helmholtz-Zentrum Dresden - Rossendorf

- Employees approx. 1,470. Thereof 670 scientists.

- **HELMHOLTZ**
RESEARCH FOR GRAND CHALLENGES

Research Fields

- Energy, Health and **Matter**.

ELBE – Center for High-Power Radiation Sources

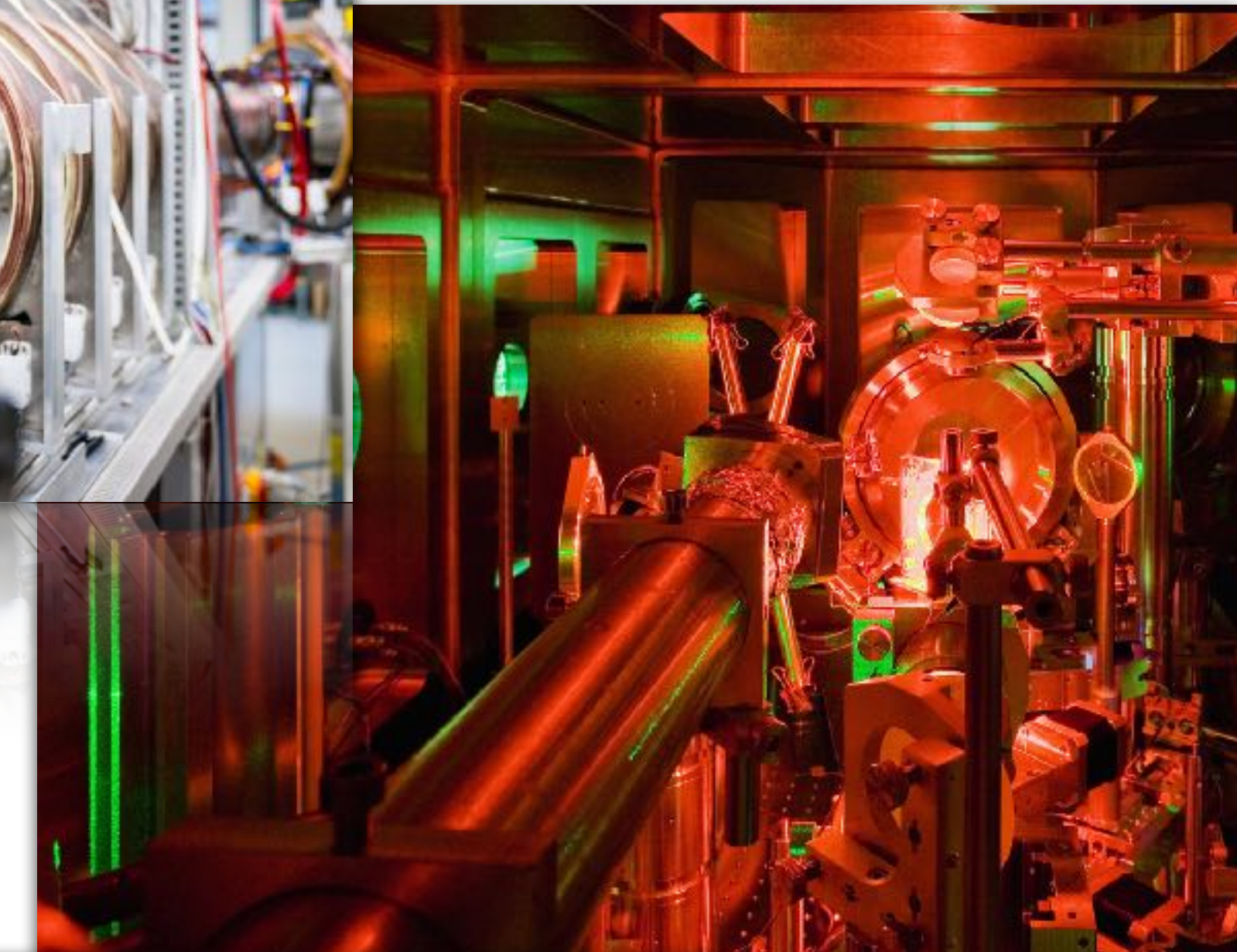
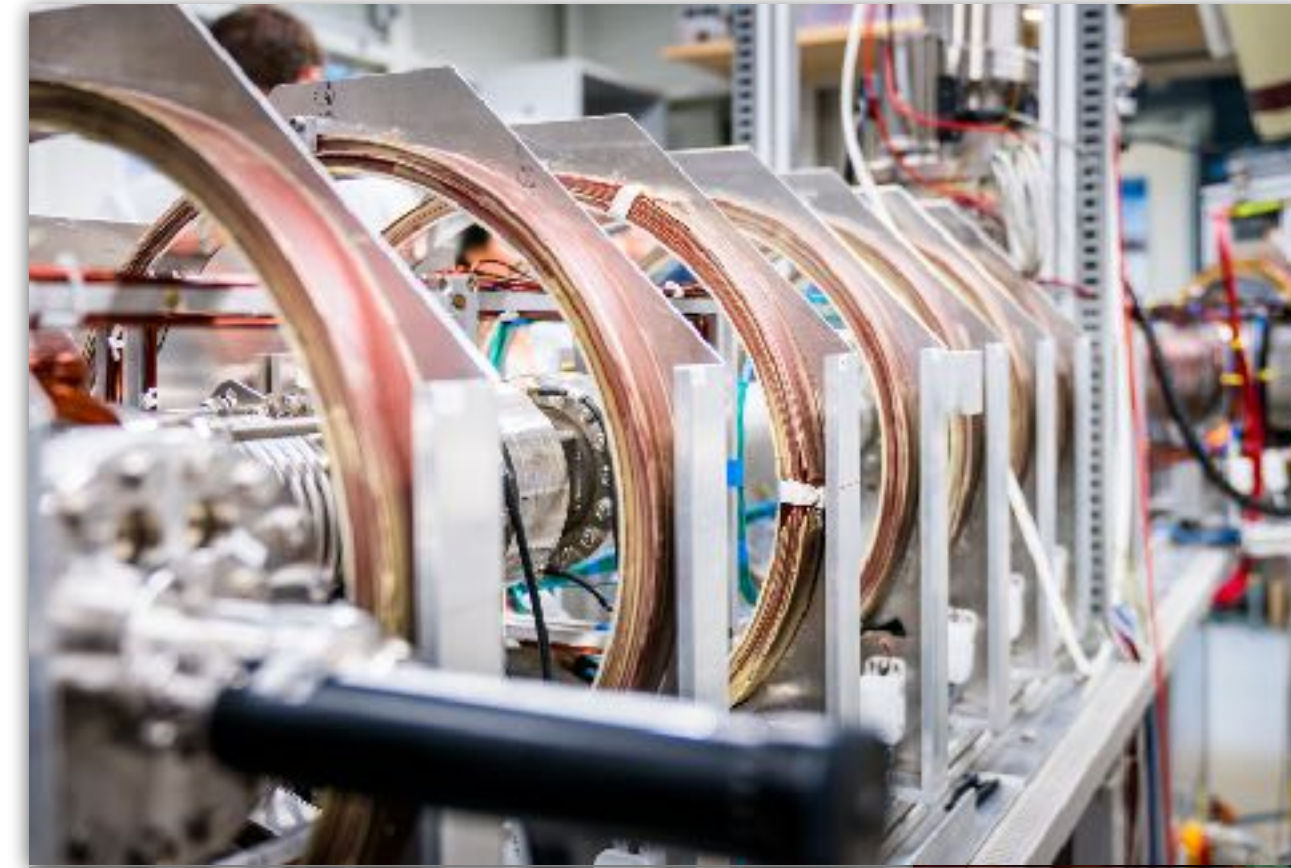
- Electron accelerator, free-electron lasers & THz source.
- Positrons, protons, neutrons as well as X-ray and gamma radiation.

Dresden High Magnetic Field Laboratory (HLD)

- Europe's highest pulsed magnetic fields.

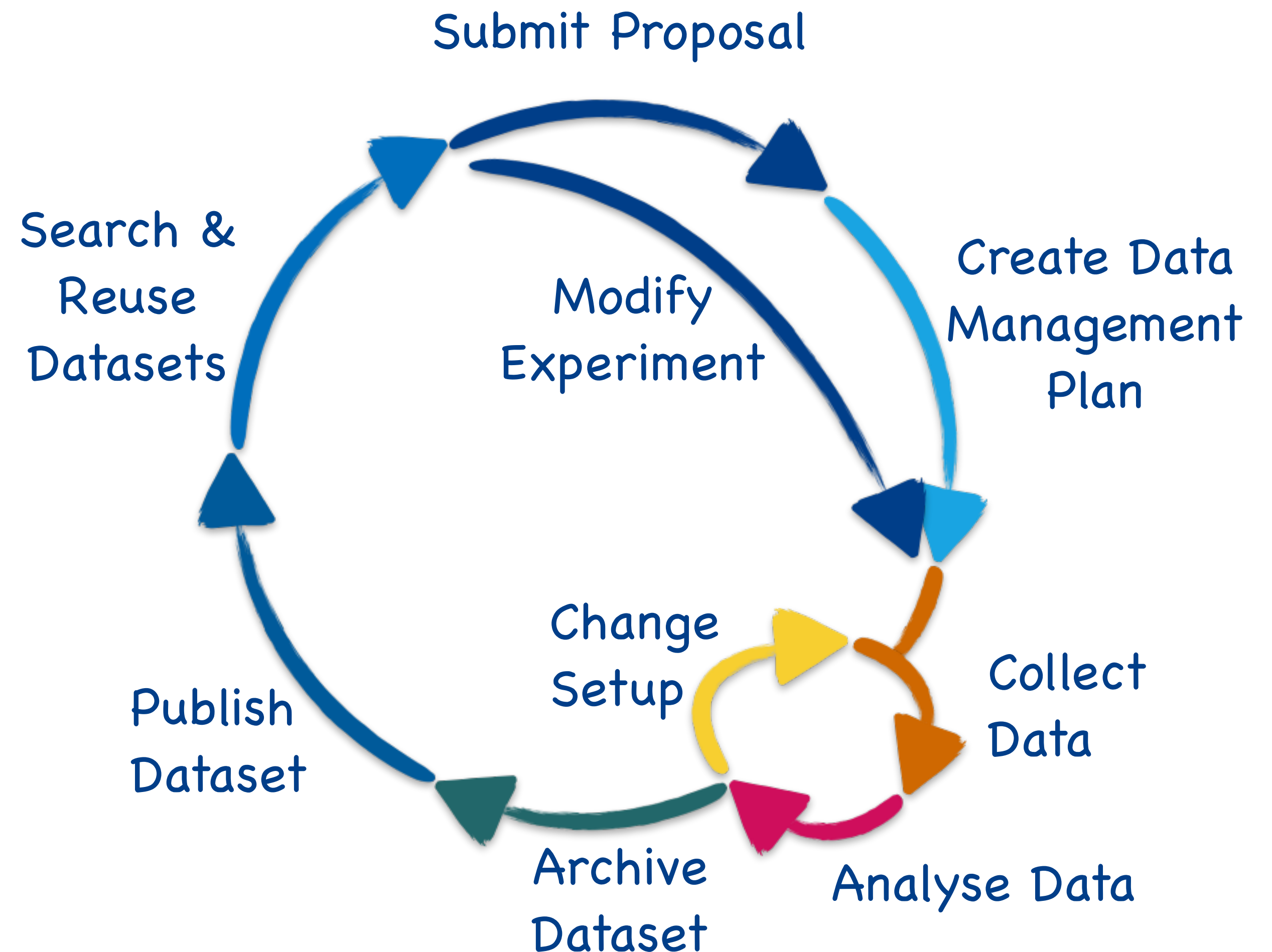
Ion Beam Center (IBC)

- Nanoscale surface analysis and modification.



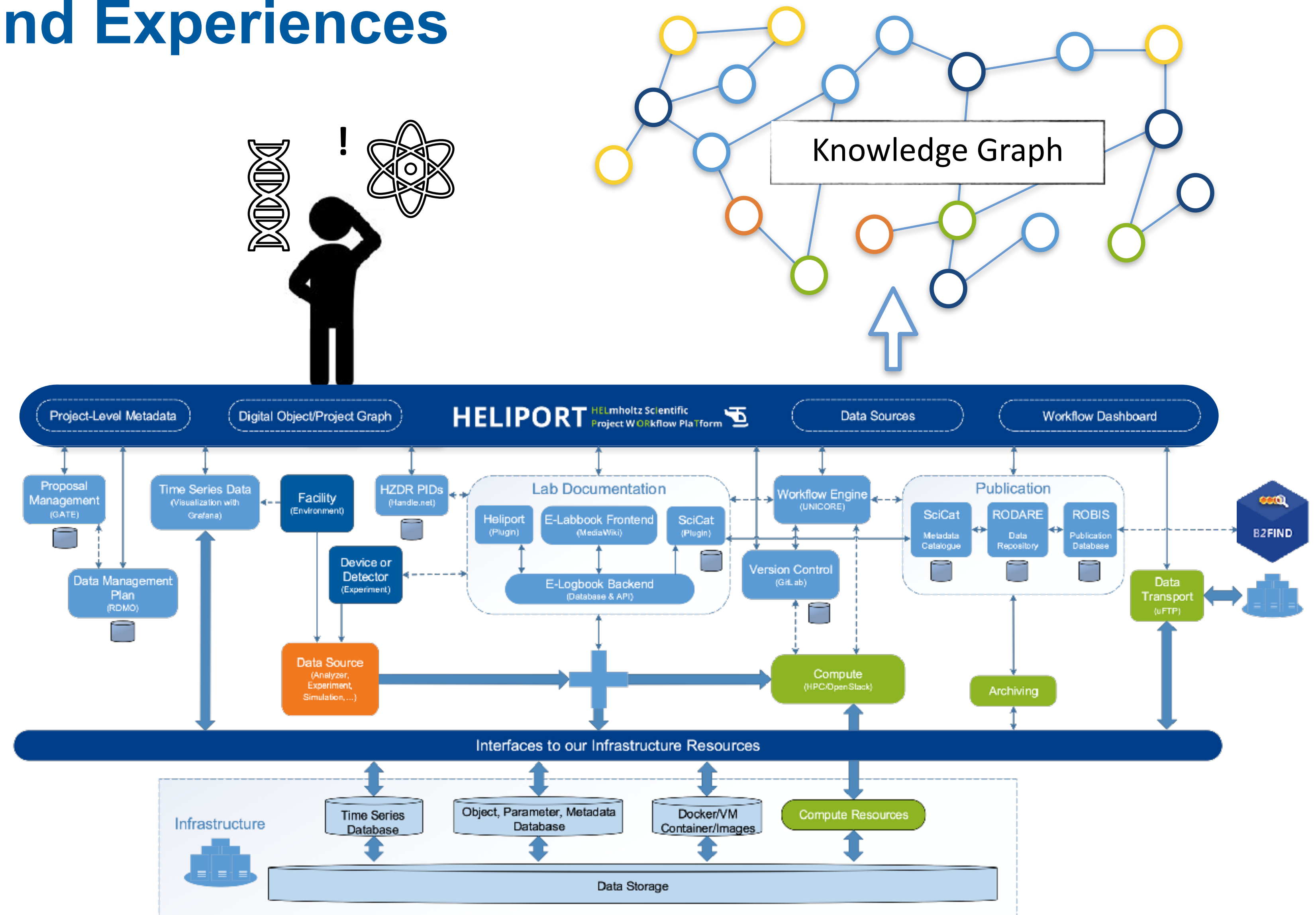
Our Challenge: An End-to-End Digital Data Lifecycle

- We support many steps of our different research experiment (matter, energy and health) with tools:
 - electronic lab books,
 - interactive analysis,
 - publication of datasets,
 - scientific workflow management,
 - Handle generation and management.
- A uniform and smooth access to and between all services and systems in our ecosystem is necessary.
- The documentation of all these linked resources is essential to create a comprehensible and FAIR data lifecycle.

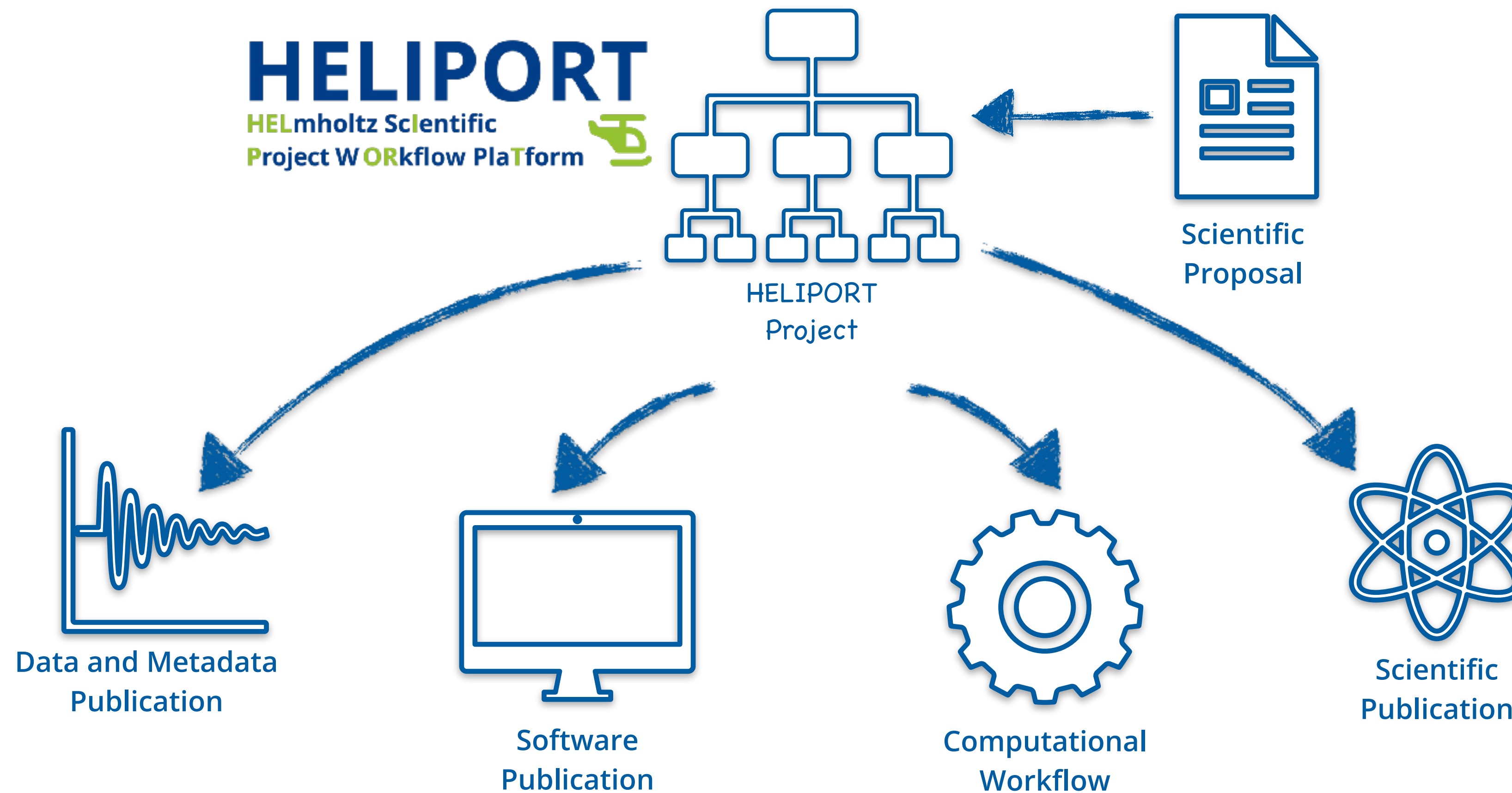


Our Observations and Experiences

- Our IT infrastructures can support various experiments, but they are complex...
- Scientists often don't know which services are available and how to use them.
- An overarching system guiding our scientists (and visitors) through the lifecycle of their research project is essential.
- In the future we can provide an overall Helmholtz-wide knowledge graph!



The Idea for a HMC Project: Common entry point for experiments



The HMC Proposal HELIPORT

HELMholtz Scientific
Project WORKflow Platform

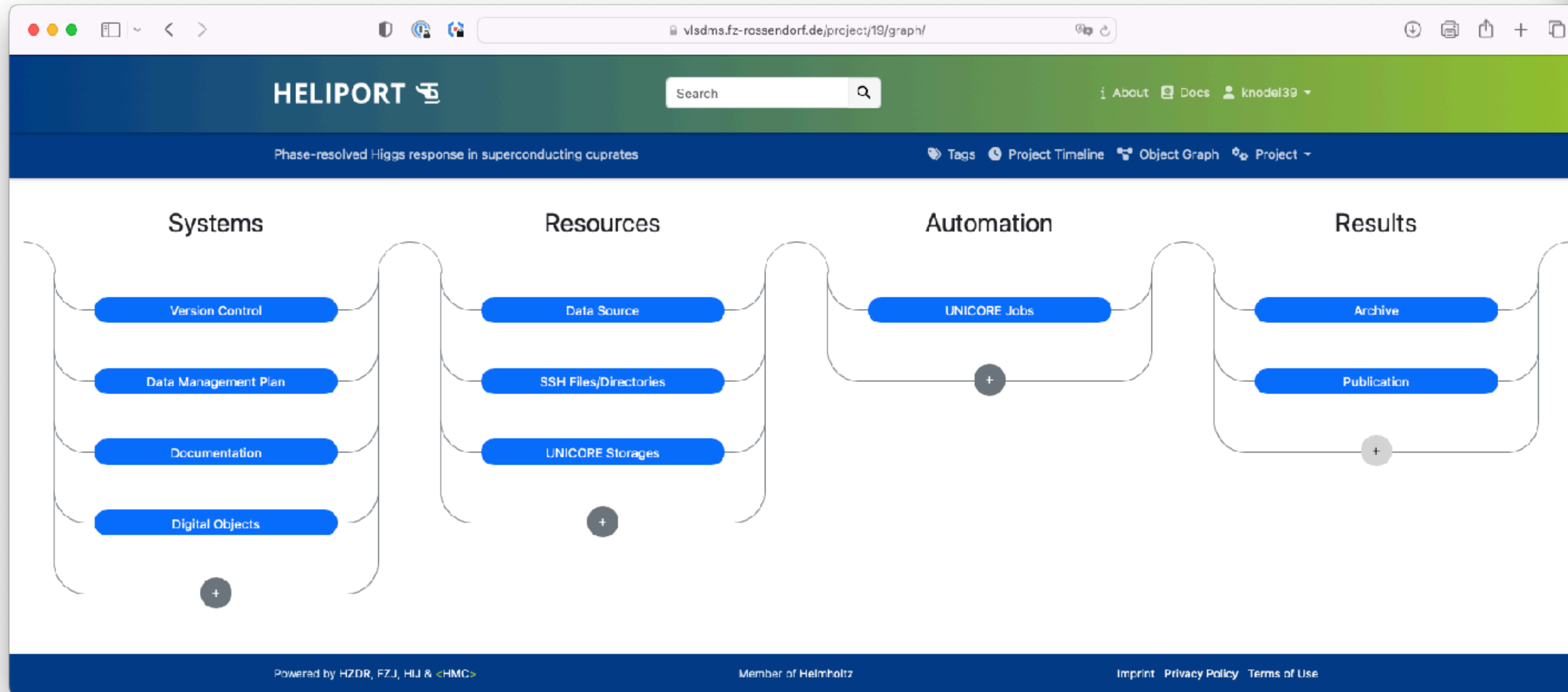


“ The HELIPORT project aims at developing a platform which accommodates the **complete life cycle** of a scientific project and links all corresponding programs, systems and workflows to create a more **FAIR and comprehensible** project description.

Project Members:



Funded by:

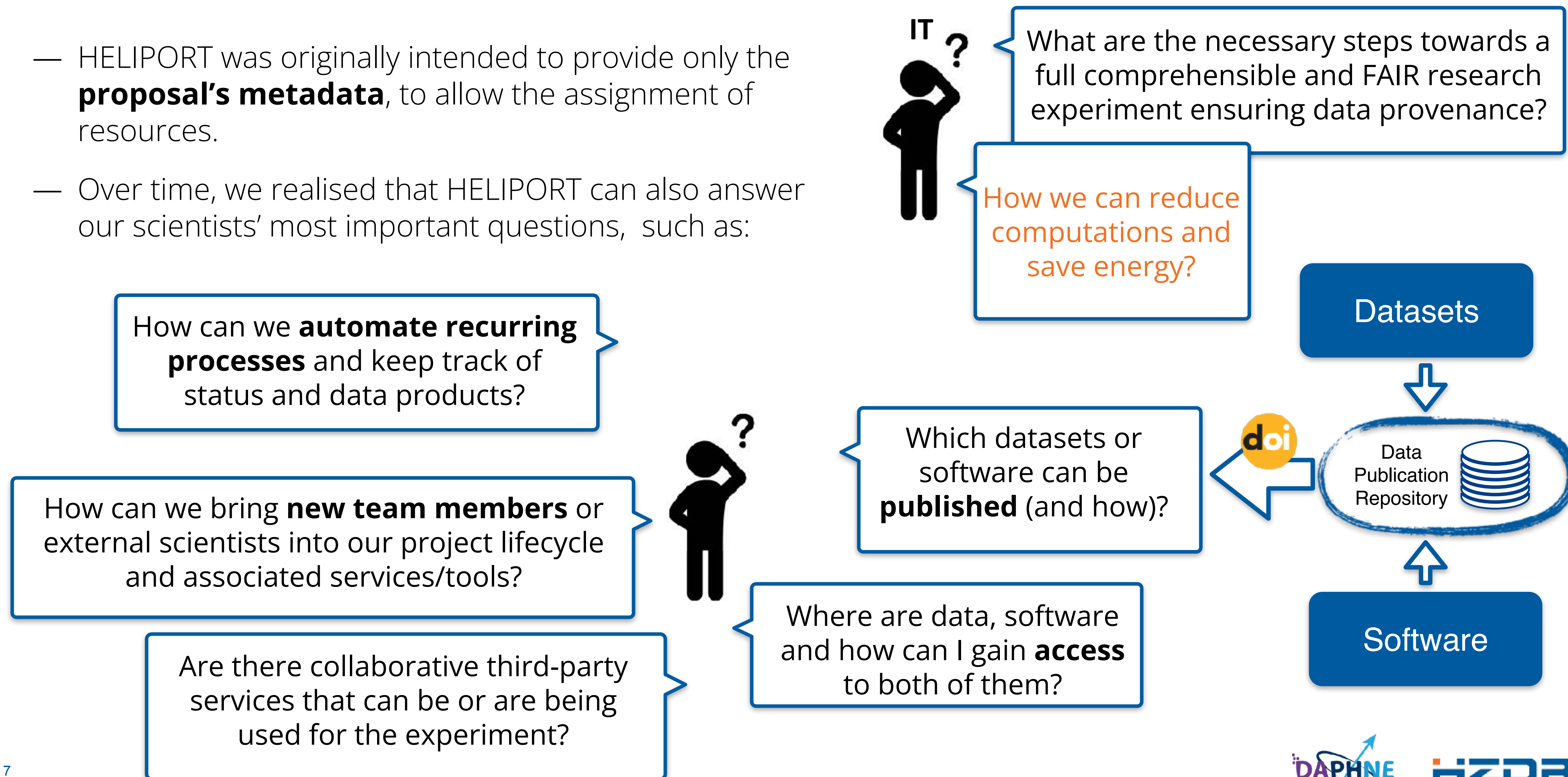


```
"namespaces": {
  "datacite": "http://purl.org/spar/datacite/",
  "rdfs": "http://www.w3.org/2000/01/rdf-schema#",
  "heliport": "https://heliport/schema/",
  "time": "http://www.w3.org/2006/time#",
  "dc": "http://purl.org/dc/terms/"
},
"heliport:project_id": 28,
"datacite:hasIdentifier": "HZDR.FWCC.2021.84769",
"heliport:uuid": "09779261-200c-48c4-be9c-f298369d6a1c",
"datacite:handle": "https://hdl.handle.net/None",
"heliport:project_name": "PaN Research Project",
"time:hasBeginning": "2021-04-01 09:14:34.295524+00:00",
"datacite:hasDescription": "",
"heliport:group": "FWCC",
"heliport:owner": {
  "datacite:hasIdentifier": "132739",
  "datacite:orcid": null,
  "rdfs:label": "Knodel, Dr. Oliver (FWCC) - 132739"
},
"heliport:has_VersionControl": [
  {
    "heliport:version_control_id": 15,
    "datacite:uri": "https://ddd",
    "rdfs:label": "ddd"
  }
],
"heliport:has_Archive": [
  {
    "heliport:archive_id": 7,
    "datacite:uri": "https://ddd",
    "heliport:documentation_system": "MediaWiki1",
    "datacite:hasDescription": "ddd"
  }
],
"heliport:has_DataSource": [
  {
    "heliport:data_source_id": 11,
    "datacite:uri": "http://ddd",
    "heliport:use_computer": null,
    "rdfs:label": "ddd",
    "datacite:hasDescription": ""
  }
]
```

Metadata crosswalk to
schema.org
ResearchProject

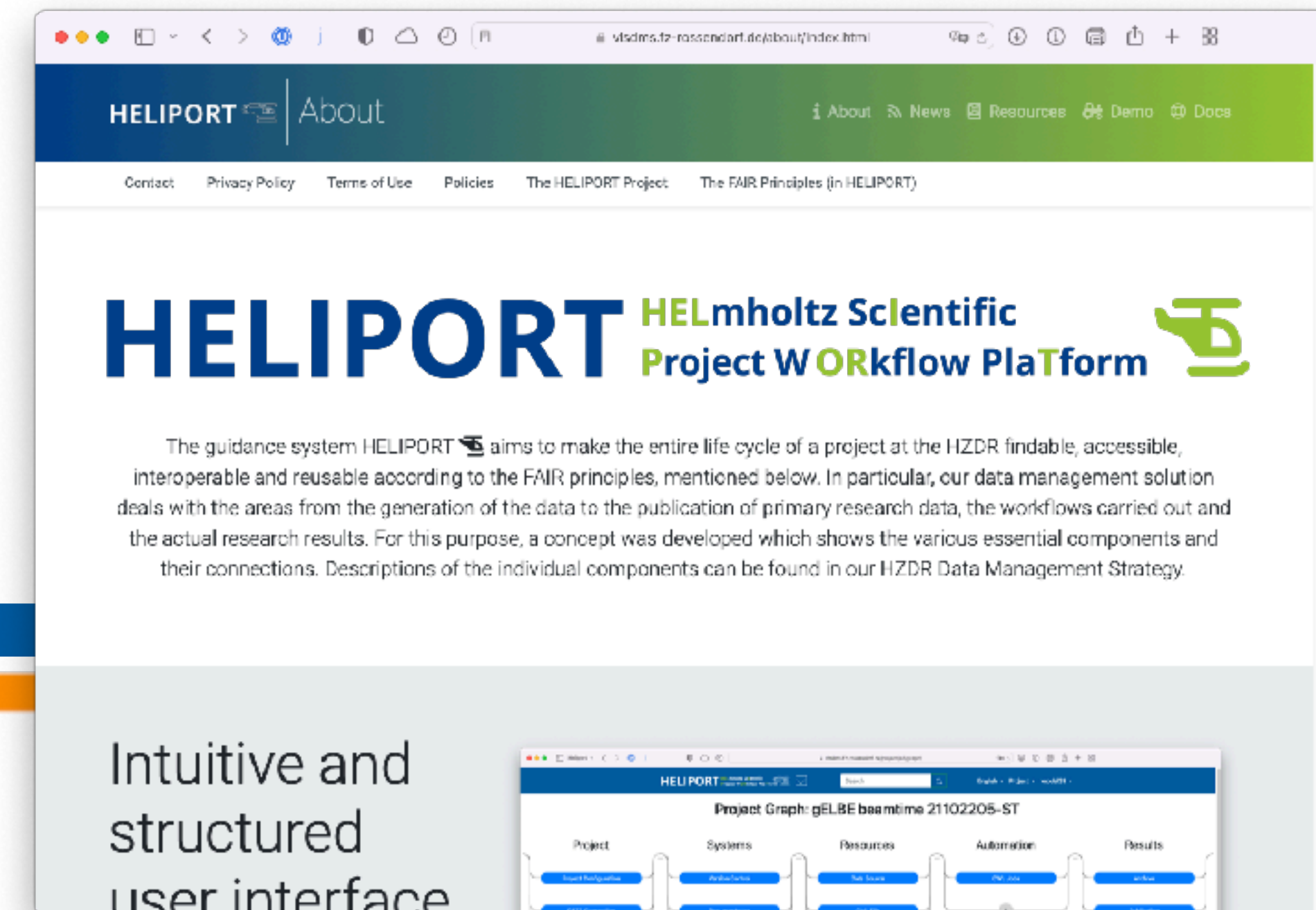
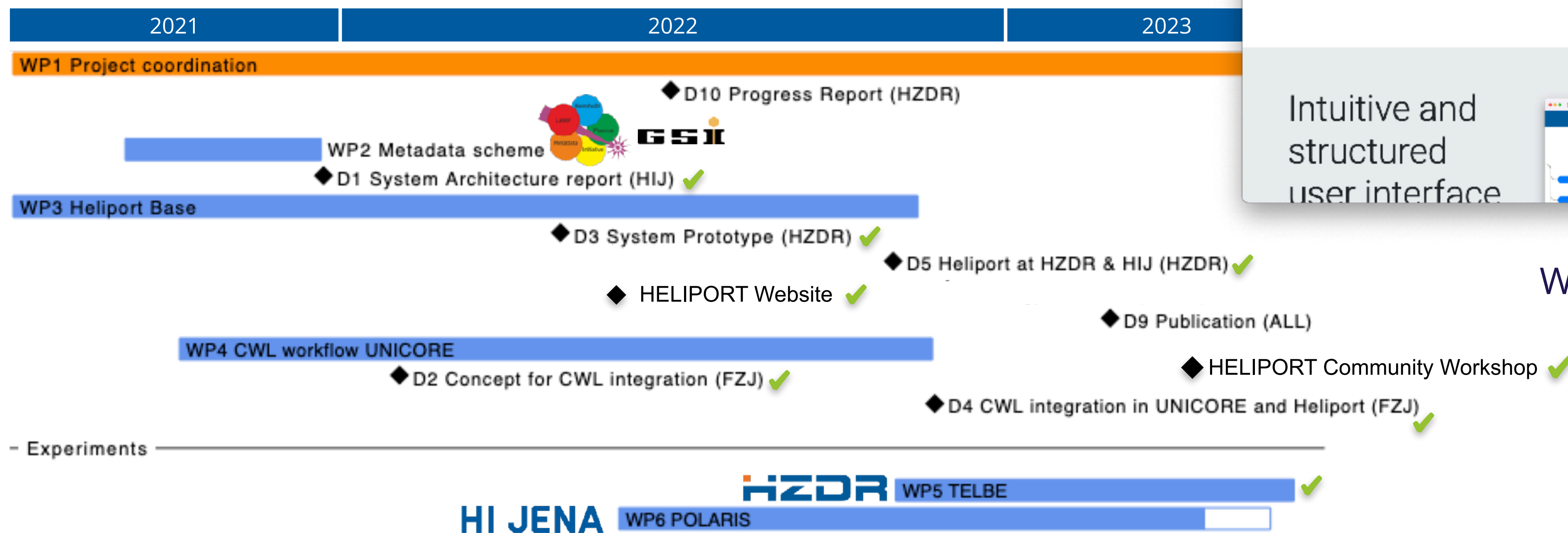
Motivation for HELIPORT — User Perspective

- HELIPORT was originally intended to provide only the **proposal's metadata**, to allow the assignment of resources.
- Over time, we realised that HELIPORT can also answer our scientists' most important questions, such as:



Timescale for the HMC-Project HELIPORT

- The deliverables and our prototype are available on our website.
- We are in contact with different Helmholtz centers, universities and European partners and build a HELIPORT community.
- Overview of work packages and milestones:



Website: heliport.hzdr.de



Heliport (Project) Timeline

First Draft: Project Plan (August 2020)

- Project and user management
- Configurable stages
- **REST API** for proposal information
- CWL visualization prototype

Modular Structure (July 2021)

- Official start of the HMC founded Heliport project:



- Redesign to provide modular and highly configurable system

HELIPORT Community Workshop (July 2023)



0.1.0

Initial Version (June 2020)

- DMS Projects and proposal information from the **HZDR GATE proposal** database
- Webinterface with user authentication (LDAP)

0.2.0

0.3.X

Improved Project Plan (December 2020)

- Configurable stages and modules
- Infrastructure and database updates
- Daily proposal database update
- Advanced logging and monitoring

0.4.0

0.5.X

Integration of various Apps and Features (Mid 2022)

- Export for (different) metadata schemas
- Computational/scientific workflow execution
 - UNICORE support
 - Computing job management and monitoring
- Handle management with public landing pages
- **HELIPORT Website**: heliport.hzdr.de

0.6.0

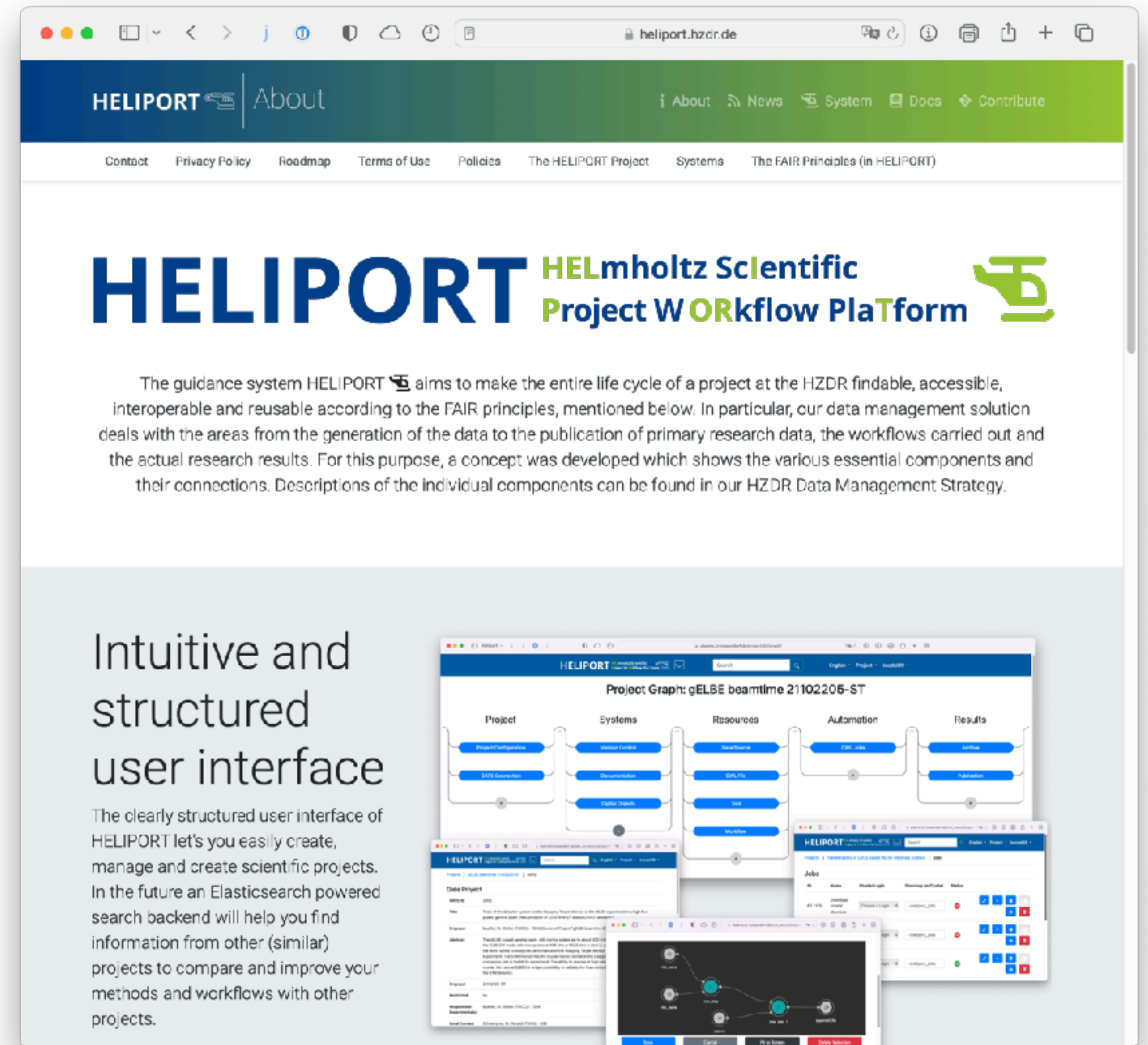
Productively operating HELIPORT for different RIs

- Extended support for a proposal system (GATE)
- Authentication with OpenID Connect (Helmholtz ID)
- Public available HELIPORT instance for remote/visiting scientists at HZDR



Community Website: heliport.hzdr.de

- Central access point to our resources at **heliport.hzdr.de** with:
 - Overall information on HELIPORT,
 - Documentation,
 - Ressources (Poster, Presentation, Paper),
 - News and Events section and
- Our HELIPORT instances is a first step towards an overall community portal (maybe heliport.helmholtz.cloud?).

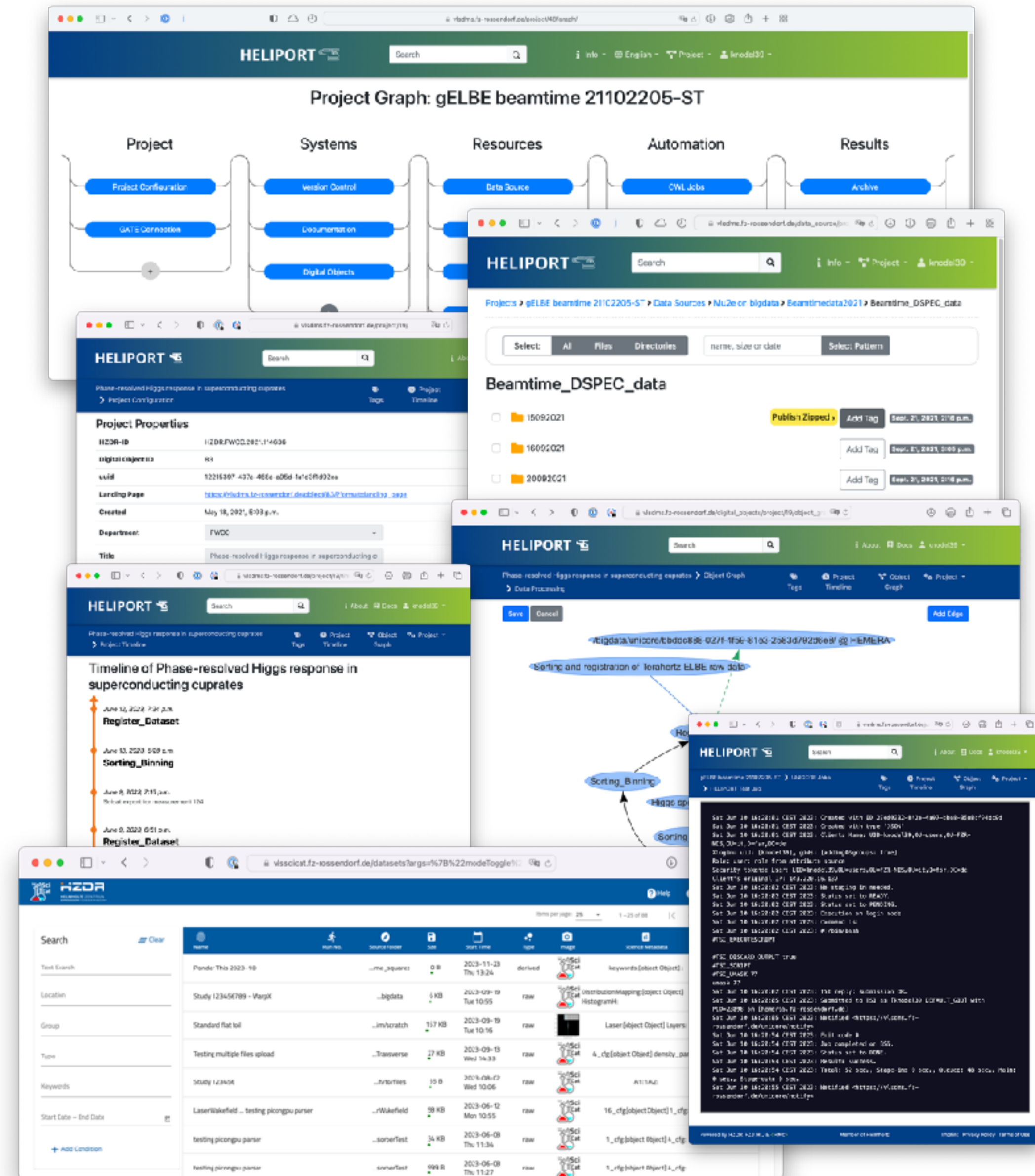


~~HZDR
Version
(Internal)~~



HELIPORT Features

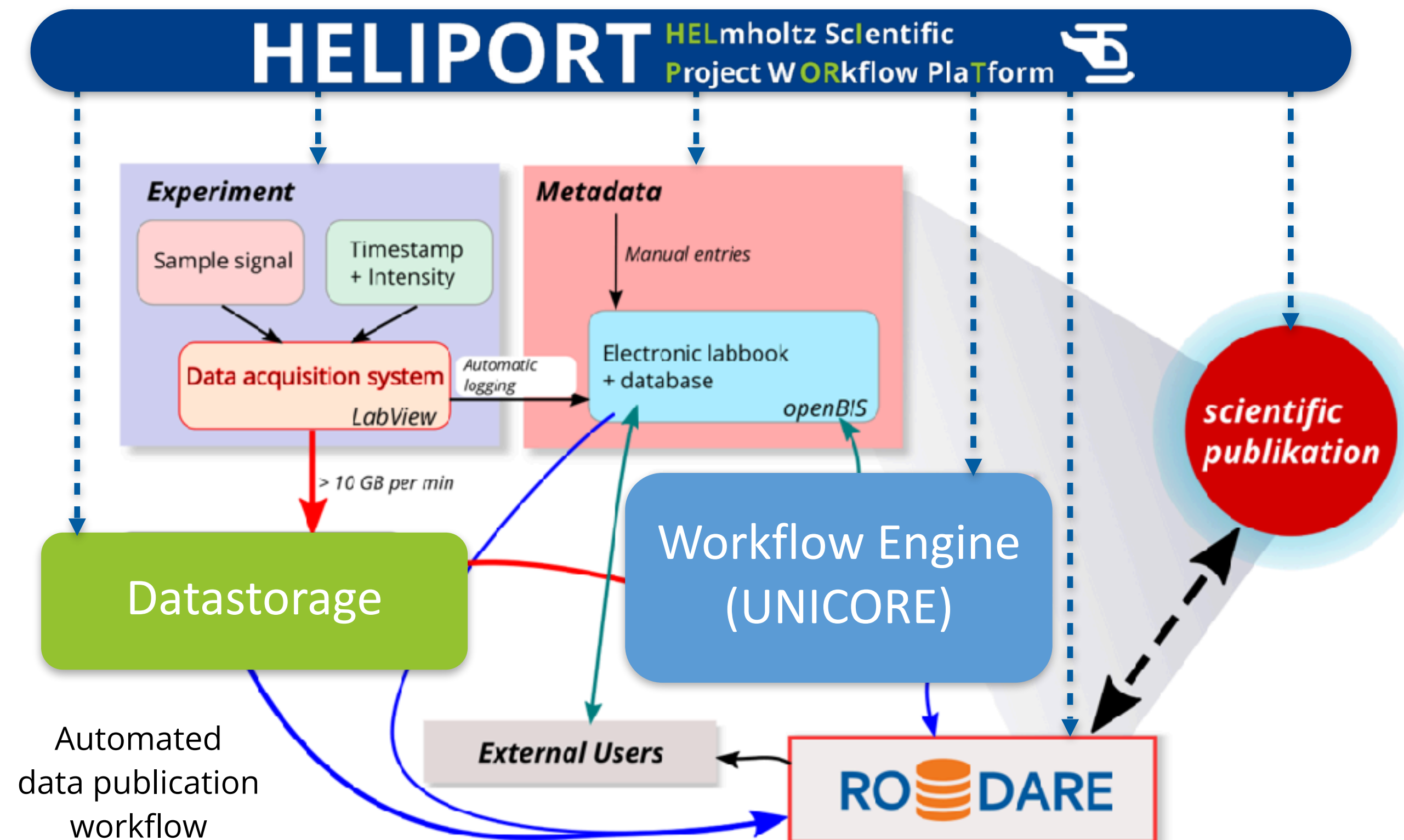
- Entry point for experiments and scientific projects
- User and group authorisation/management
- Overview of systems and services involved in an experiment
- Provision of metadata from proposal systems (e.g. GATE)
- Registration of and access to internal file systems
- Automated transfer of metadata between involved systems/services
- Background data publication of datasets (e.g. Zenodo, Rodare)
- Digital object and handle management with graph visualisation
- Timeline representing changes
- HELIPORT REST API
- Authentication via Helmholtz ID
- Todos:
 - Integration of reproducible computational workflows
 - HPC cluster access (slurm, UNICORE)



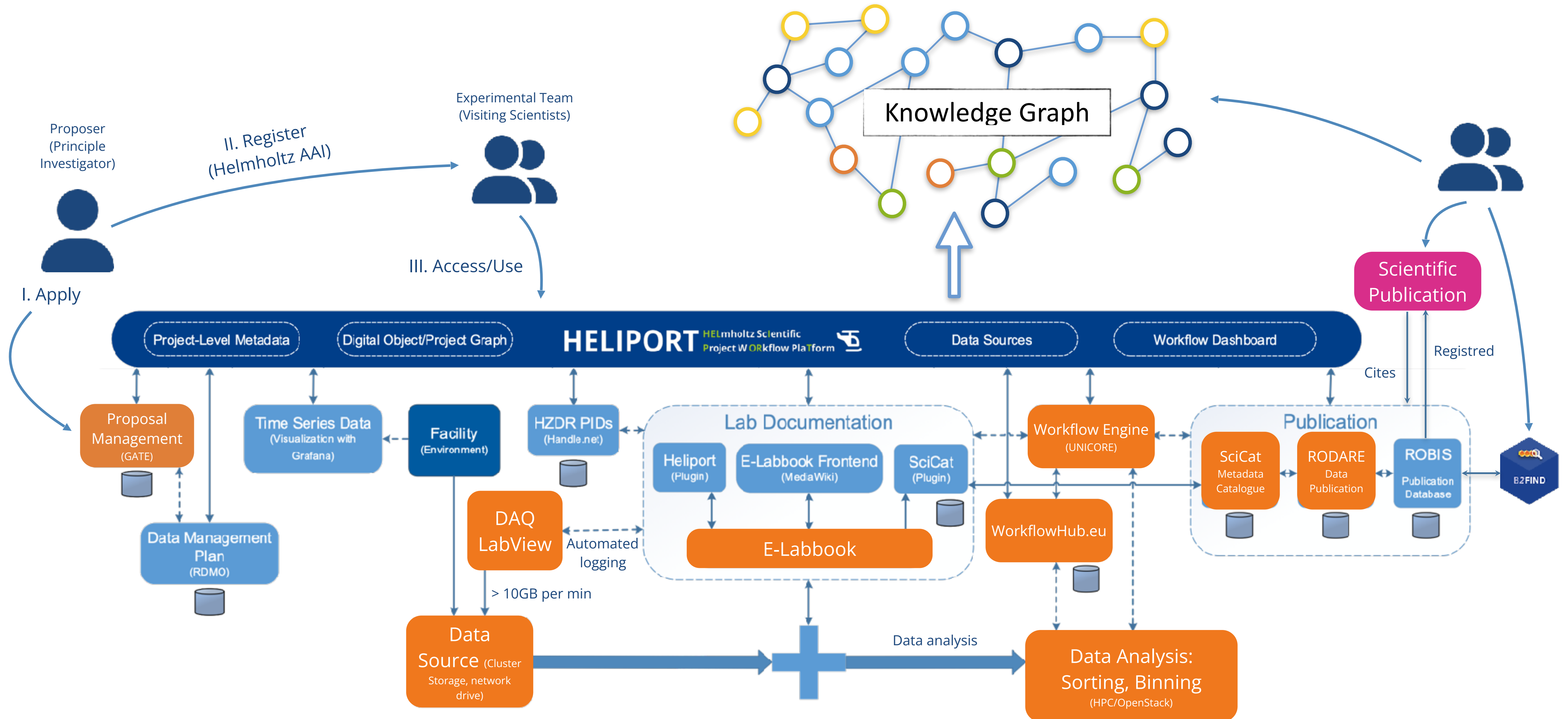
Example: TELBE Beamtime

TELBE Data Flow

- Terahertz facility at the ELBE center for High-Power Radiation Sources.
- In the future HELIPORT guides (external) scientists through the complete experiment.
- Submission of data analysis Jobs from LabView to our HPC Cluster with visualisation in HELIPORT



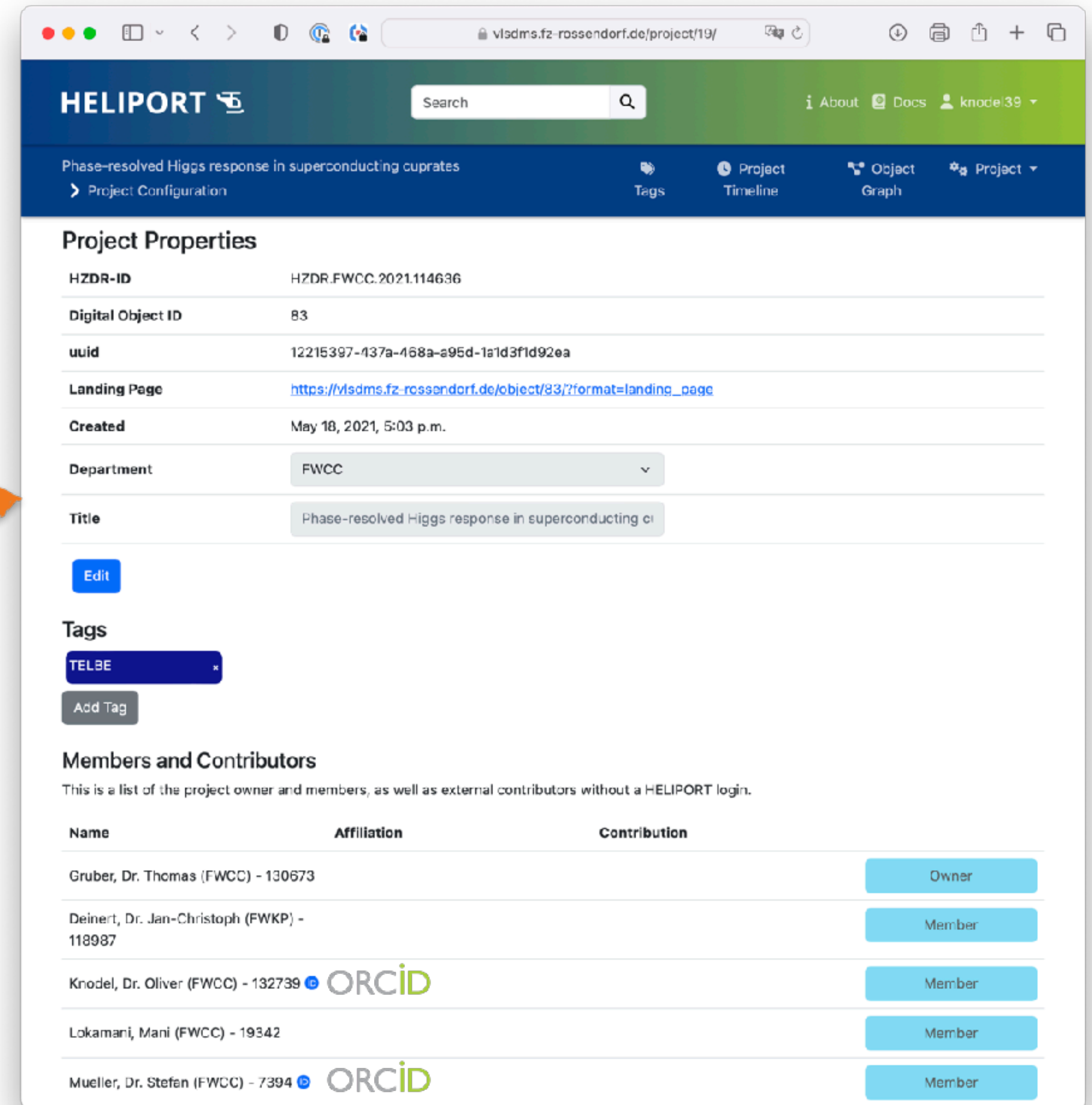
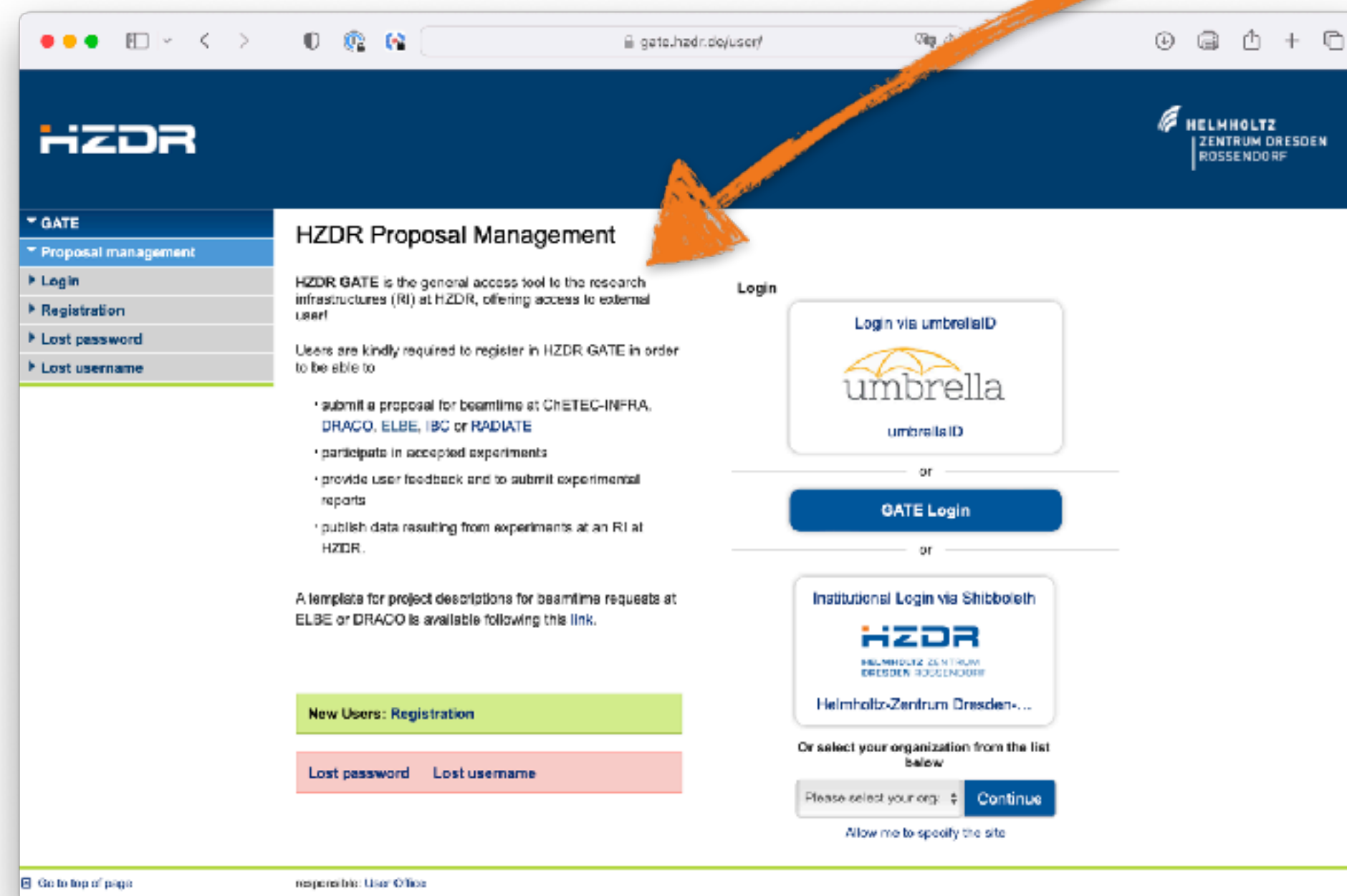
Mapping of the TELBE Resources to HELIPORT



I. Proposal Submission

Automated transfer of project metadata from the proposal system (GATE) into HELIPORT:

- Title, Authors, Description,
- Beamtime schedule,
- Large-scale facility used,
- Scientific method (PaNET)



II. Project List and Dashboard

- Typically, a beam line scientist is the owner of a HELIPORT project and the proposer has the role of the manager and can add additional project members.
- Tags and sub-projects including inheritance are possible in the project list.

Project List

Project Name	Last Modified	Owner	Action
Semantic x-Lab	Jul 11, 2023	Voigt, Martin (FWCC-D) - 141575	Open
gELBE Projects	Apr 24, 2023	Mueller, Dr. Stefan (FWCC) - 7394	Open
gELBE beamtime 21102205-ST	Sep 11, 2023	Mueller, Dr. Stefan (FWCC) - 7394	Open
gELBE beamtime 21202619-ST	Sep 11, 2023	Mueller, Dr. Stefan (FWCC) - 7394	Open
Example parent project	Apr 24, 2023	Voigt, Martin (FWCC-D) - 141575	Open
ML Ops Project	Jun 06, 2023	Knodel, Dr. Oliver (FWCC) - 132739	Open
SOTA on Uncertainties	May 23, 2023	Pape, David (FWCC) - 139658	Open
Phase-resolved Higgs response in superconducting cuprates	May 23, 2023	Gruber, Thomas (FWCC-D) - 141575	Open
Digital Twin Showcase	Jun 07, 2023	Voigt, Martin (FWCC-D) - 141575	Open
Beamtime Dashboard Test	May 31, 2022	Voigt, Martin (FWCC-D) - 141575	Open
Rodare Data Publication Project	Aug 09, 2022	Knodel, Dr. Oliver (FWCC) - 132739	Open

[Create Project](#)

« < 1 2 3 > »

Project Dashboard: Phase-resolved Higgs response in superconducting cuprates

Tags Project Timeline Object Graph Project

Systems

- Version Control
- Data Management Plan
- Documentation
- Digital Objects

Resources

- Data Source
- SSH Files/Directories
- UNICORE Storages

Automation

- UNICORE Jobs

Results

- Archive
- Publication

Powered by HZDR, FZJ, HJ & eHMC Member of Helmholtz Imprint Privacy Policy Terms of Use

III. Resources: Documentation and Repositories

The documentation section is typically used to refer to all internal and external systems or services used:

- E-Labbook (Mediawiki),
- GitLab, Github, Workflowhub, ...

The collage displays four overlapping web interfaces related to scientific data management and documentation:

- HELIPORT Documentation:** A web page titled "HELIPORT" with a search bar and navigation links. It features a table of documentation entries with columns for ID, Description, and System. The table lists an entry with ID 57, Description "Project documentation in Mediawiki", and System "MediaWiki".
- WorkflowHub:** A page titled "Sorting and registration of Terahertz ELBE raw data" (Version 1). It includes a description of the radiation source ELBE, a list of creators (Thomas Gruber), a license (Creative Commons Attribution 4.0), and a version history section.
- Mediawiki Dataset Page:** A page titled "FWKP:22 DAQ CdAs 120degs WP 45degs SHG V polar 01". It displays three plots: "FWKP direct plot of 22 DAQ CdAs 120degs WP 45degs SHG V polar", "FWKP:ftt power linescale plot of 22 DAQ CdAs 120degs WP 45degs SHG V polar", and "FWKP:ftt power logscale plot of 22 DAQ CdAs 120degs WP 45degs SHG V polar". A table below the plots lists data files and workflow details.
- Version Control Interface:** A page titled "HELIPORT" with a search bar and navigation links. It features a table of version control entries with columns for ID, Name, and System. The table lists an entry with ID 7, Name "Telbe sorting script", and System "MediaWiki".

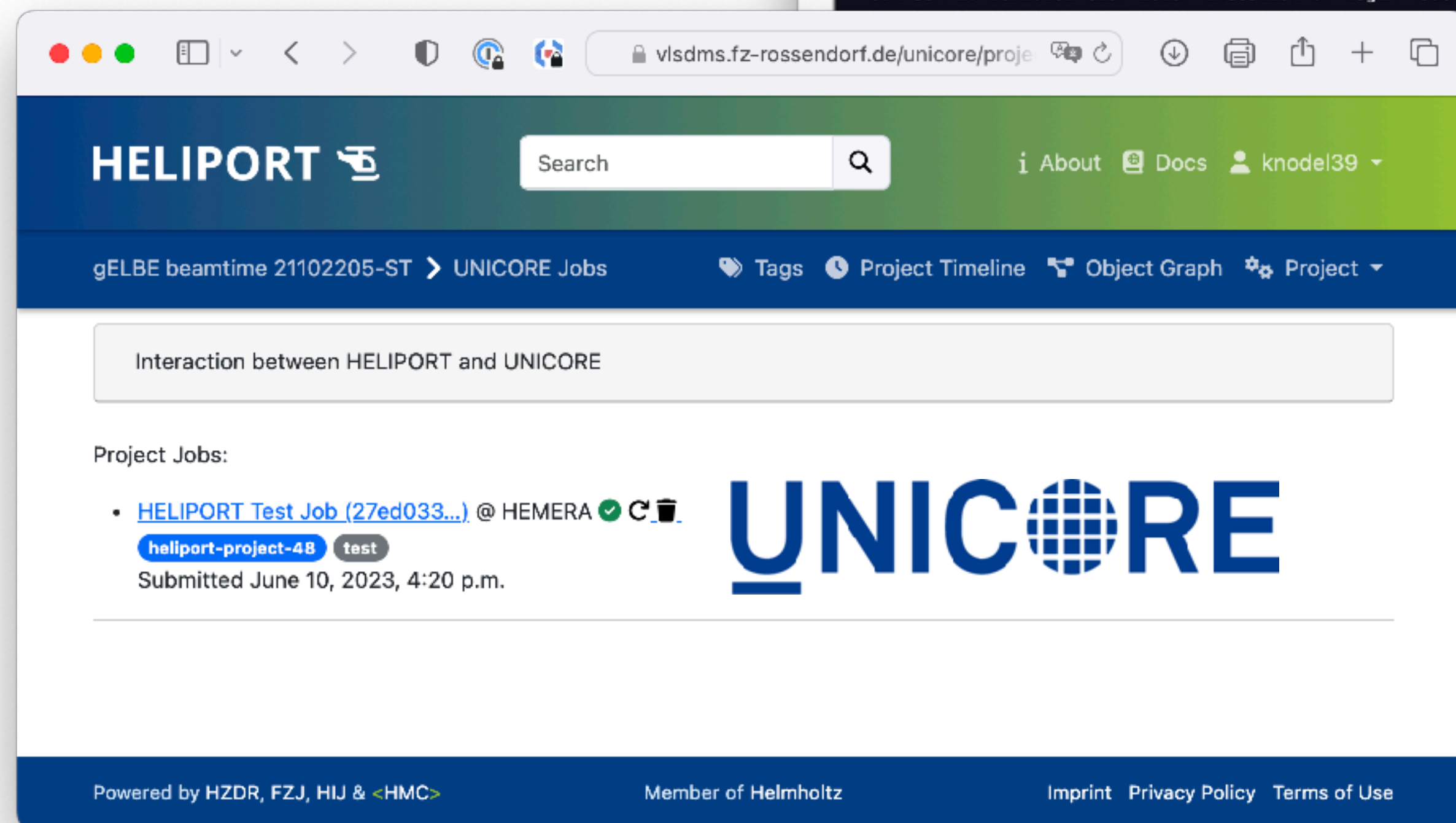
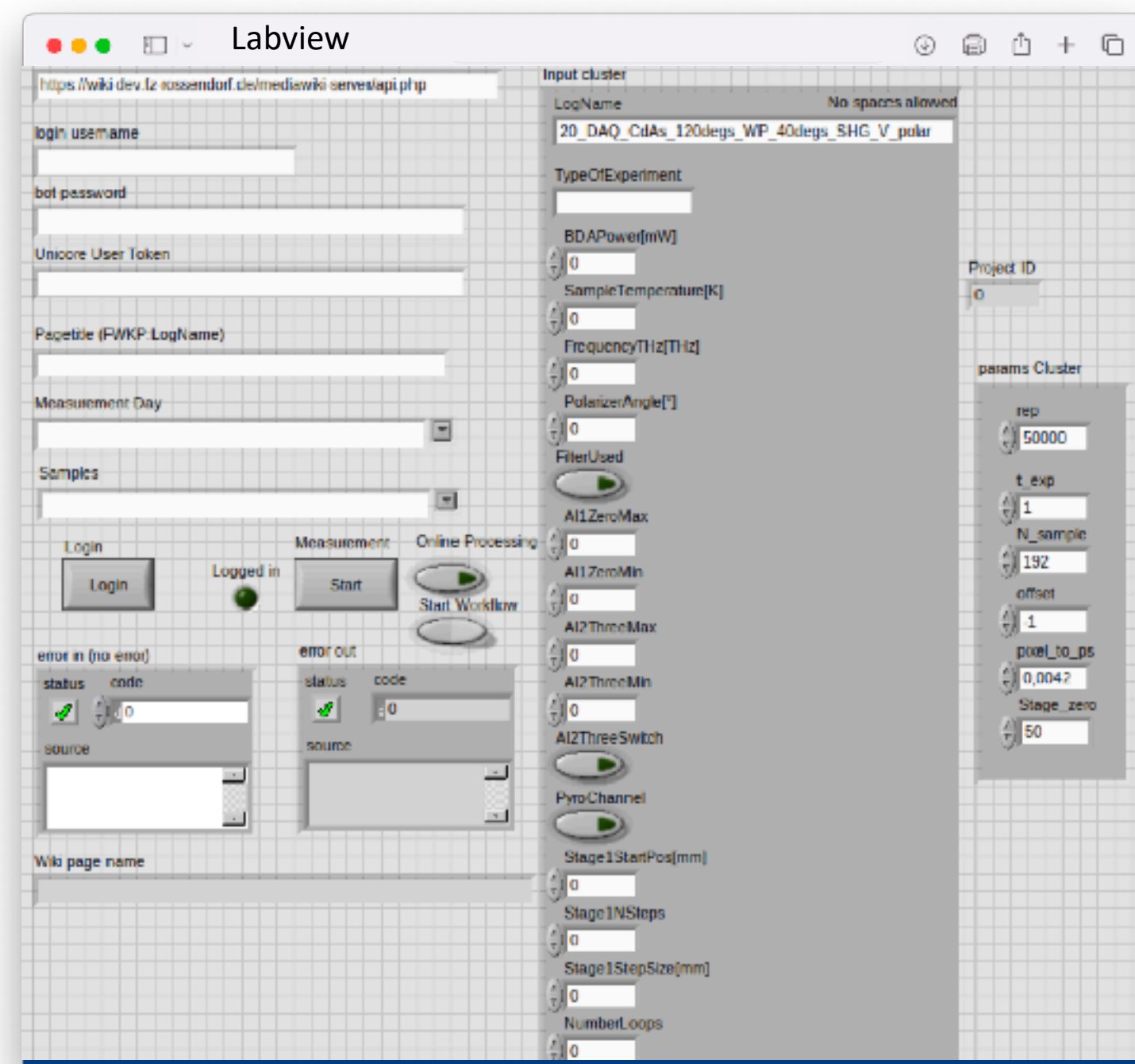
ID	Description	System
57	Project documentation in Mediawiki	MediaWiki

ID	Name
7	Telbe sorting script

Data Files	File:FWKP:22 DAQ CdAs 120degs WP 45degs SHG V polar all loops.dat
Workflowhub URL	https://workflowhub.eu/workflows/458/versions/1/
Workflowhub Version	1
Repetition Rate	50000 Hz
time on single step measurement	1

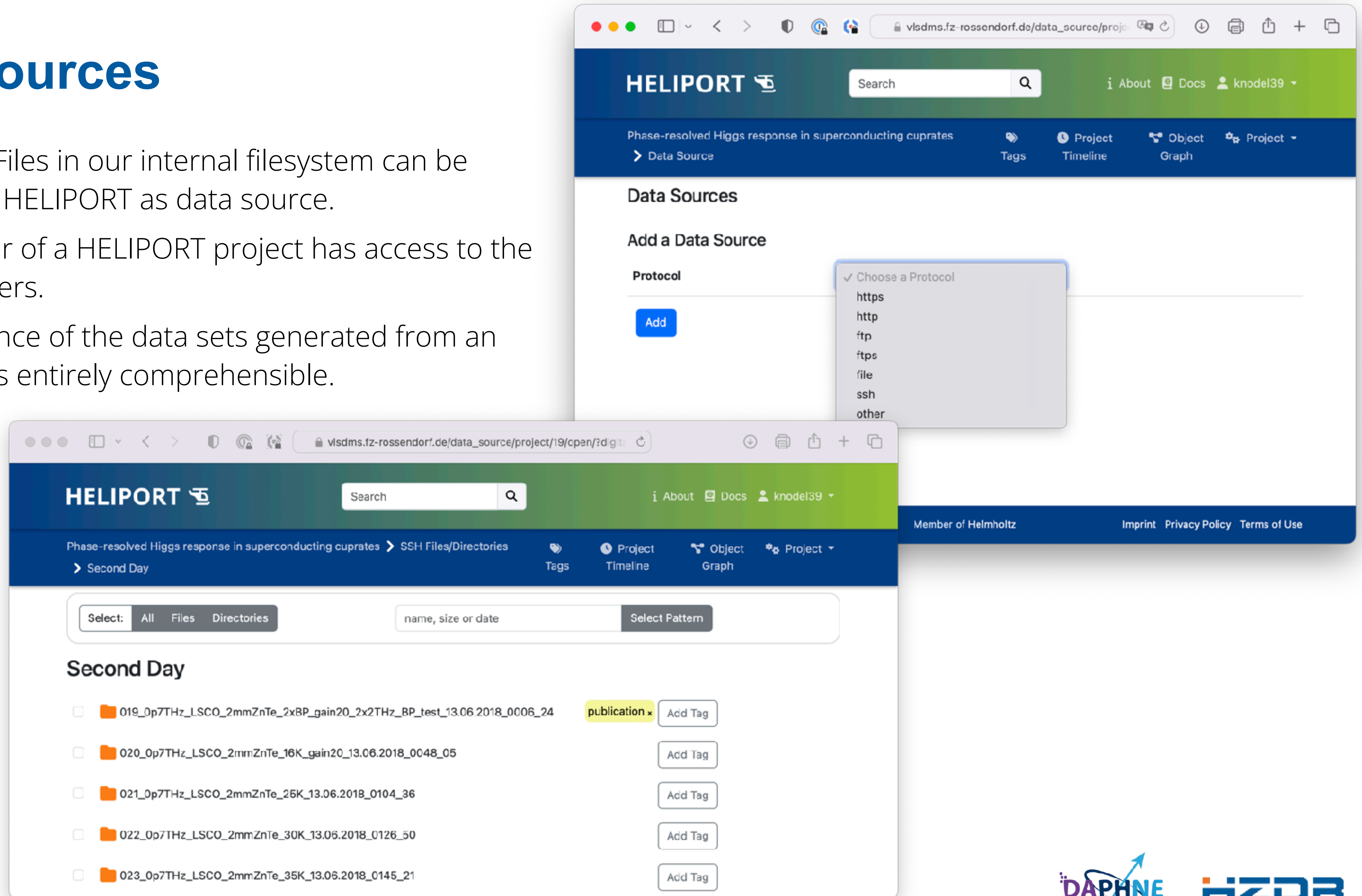
IV. Detector Control and Workflows

- The HELIPOINT REST-API enables the transfer of metadata between HELIPOINT and external systems (e.g. detector control in LabView).
- The integrated workflow management system (e.g. UNICORE) provides metadata for the provenance information required by HELIPOINT.
- Workflows (on our HPC cluster) can be viewed by any project member directly in the HELIPOINT web frontend.



V. Data Sources

- Folders and Files in our internal filesystem can be registered in HELIPORT as data source.
- Each member of a HELIPORT project has access to the files and folders.
- The provenance of the data sets generated from an experiment is entirely comprehensible.



VI. Integration in an Overall Data Publication Workflow

The workflow is shown through four overlapping screenshots:

- HELIPORT Project Properties:** Shows project details for "Phase-resolved Higgs response in superconducting cuprates". Key fields include HZDR-ID (HZDR.FWCC.2021.114636), Digital Object ID (83), and a landing page URL.
- HELIPORT Systems and Resources:** A diagram showing the integration of various systems (Version Control, Data Management Plan, Documentation, Digital Objects) and resources (Data Source, SSH Files/Directories, UNICORE Storages).
- HELIPORT Data Source:** A screenshot of the "Second Day" data source interface, showing a list of files and folders with checkboxes and "Add Tag" buttons.
- RODARE Publication:** A screenshot of the RODARE publication interface, showing the title "Phase-resolved Higgs response in superconducting cuprates", a list of authors, a preview of the data, and publication statistics (2,980 views, 10,619 downloads).

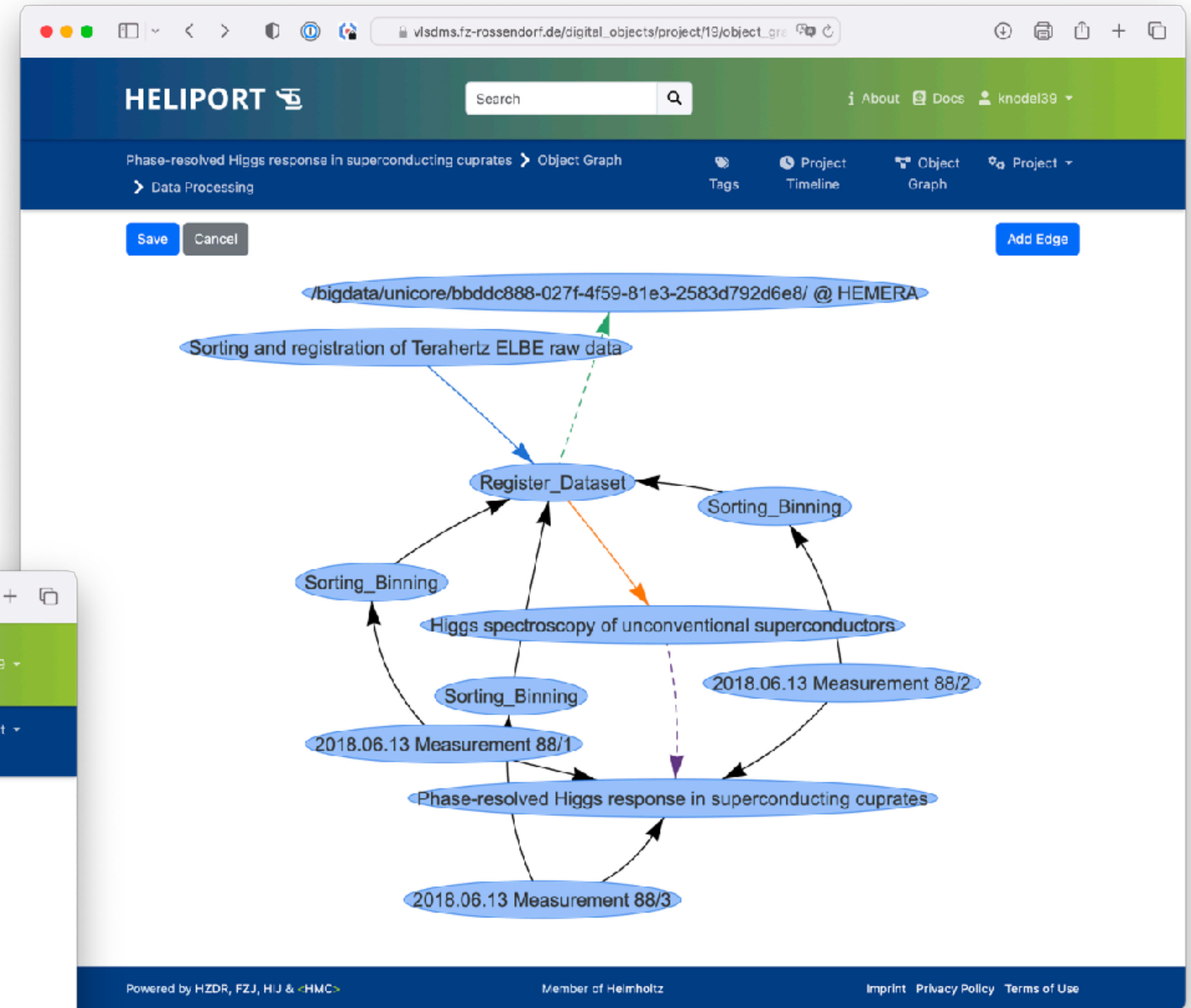
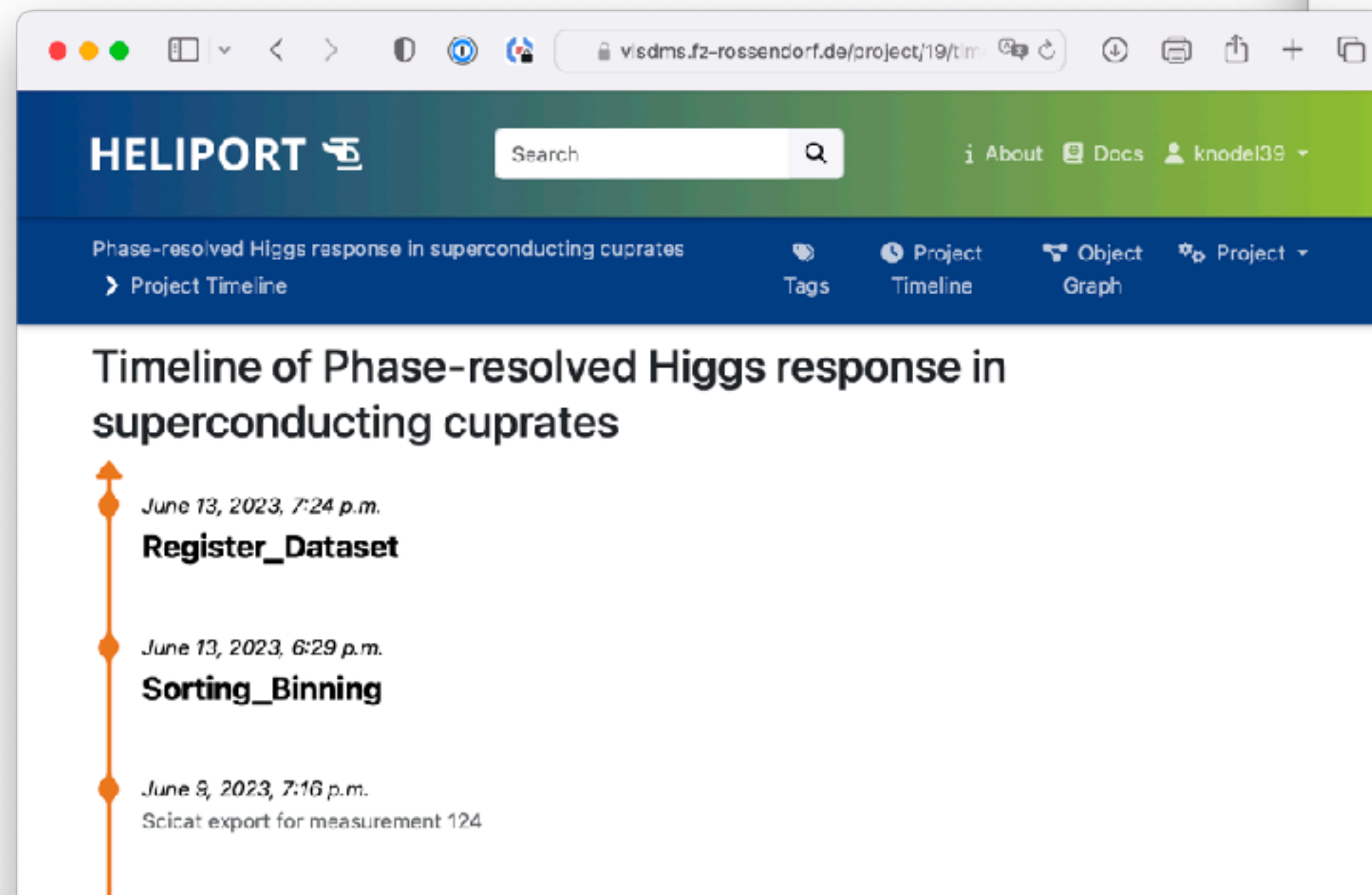
Orange arrows indicate the flow of data from the Proposal System (HELIPORT) to the Data Source (HELIPORT) and finally to the publication (RODARE).

Automated data publication with:

- Metadata from Proposal System,
- Files and folders registered and selected in HELIPORT.

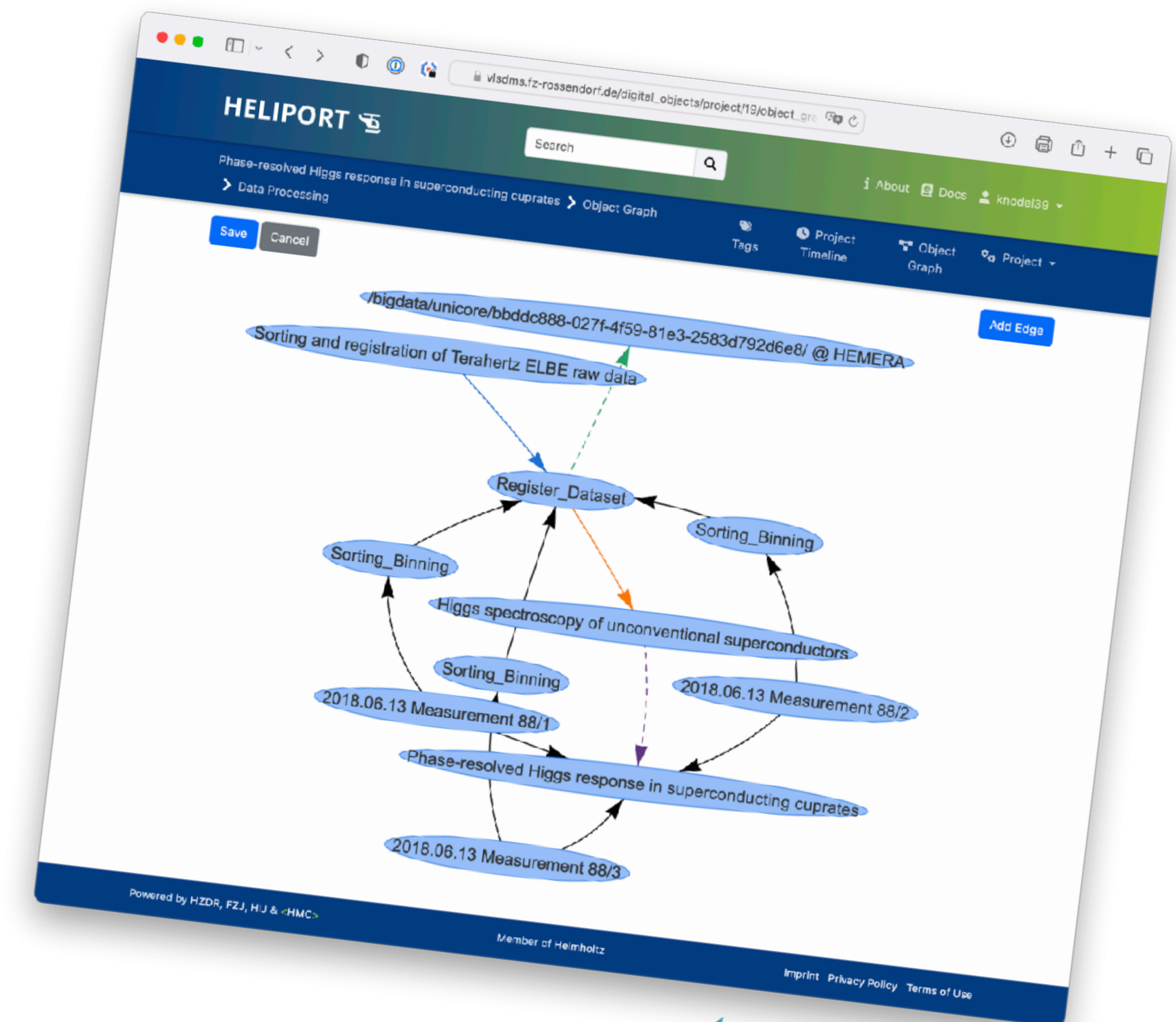
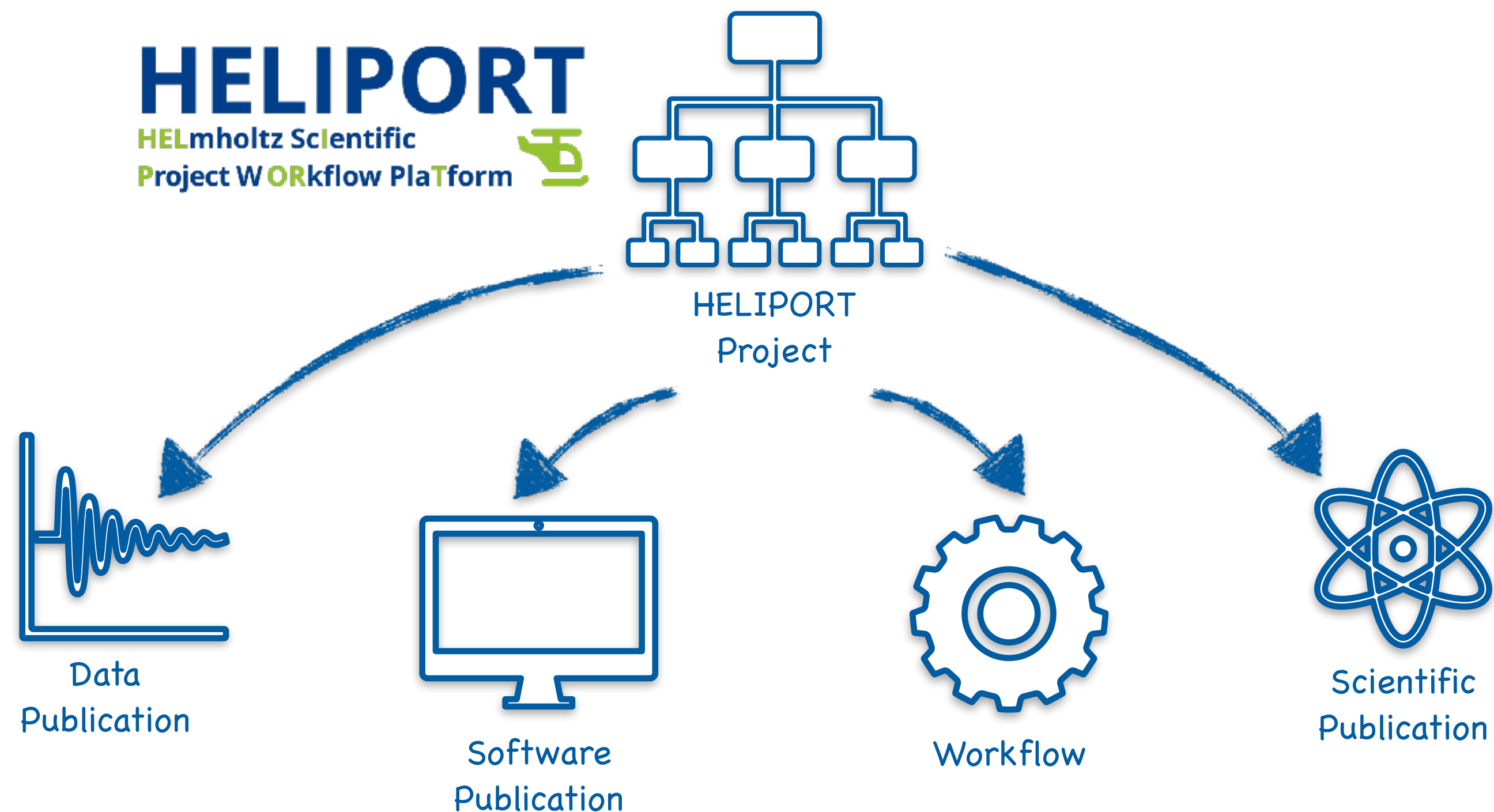
VII. Relations Between Digital Objects and

- Relations between digital objects are visualized to provide a top-level view on the project with dependencies.
- The relationships between simulation (surrogate model) and experiment can also be demonstrated.
- The versioning of an experiment is an essential extension, and first approaches via a timeline are being evaluated.



Conclusions

- HELIPORT describes and collects all metadata from all services and systems involved in an scientific experiment.
- Such an approach is desirable and leads us to a fully **FAIR** and **comprehensible** research project.
- Data sources and workflows are essential to keep track of everything what happened during an experiment.



Resources

Website: heliport.hzdr.de

Repository: codebase.helmholtz.cloud/heliport



The collage displays various resources for the HELIPORT project. At the top left is a QR code. The central part features the HELIPORT website homepage, which describes the system as a guidance system for project lifecycle management. To the left of the homepage is a screenshot of the API documentation, showing a search bar and a list of endpoints. Below the homepage is a screenshot of the 'Response samples' for a GET request to '/api/projects/', showing a JSON response with project details. To the right of the homepage is a screenshot of the HELIPORT GitHub repository page, showing the project's commit history and storage. At the bottom right is a screenshot of a workshop presentation slide titled 'HELIPORT: A Portable Platform for FAIR {Workflow | Metadata | Scientific Project Lifecycle} Management and Everything' by Oliver Knodel et al. The presentation slide also features the DOI 10.1145/3456287.3465477.

API Doc: heliport.hzdr.de/redoc/

