

Integrated Research Data Management at HZDR

Thomas Gruber, Jeffrey Kelling, Oliver Knodel, Mani Lokamani, Stefan Müller, David Pape, and Guido Juckeland
Helmholtz-Zentrum Dresden-Rossendorf, Computational Science Department, Data Management and HPC Group

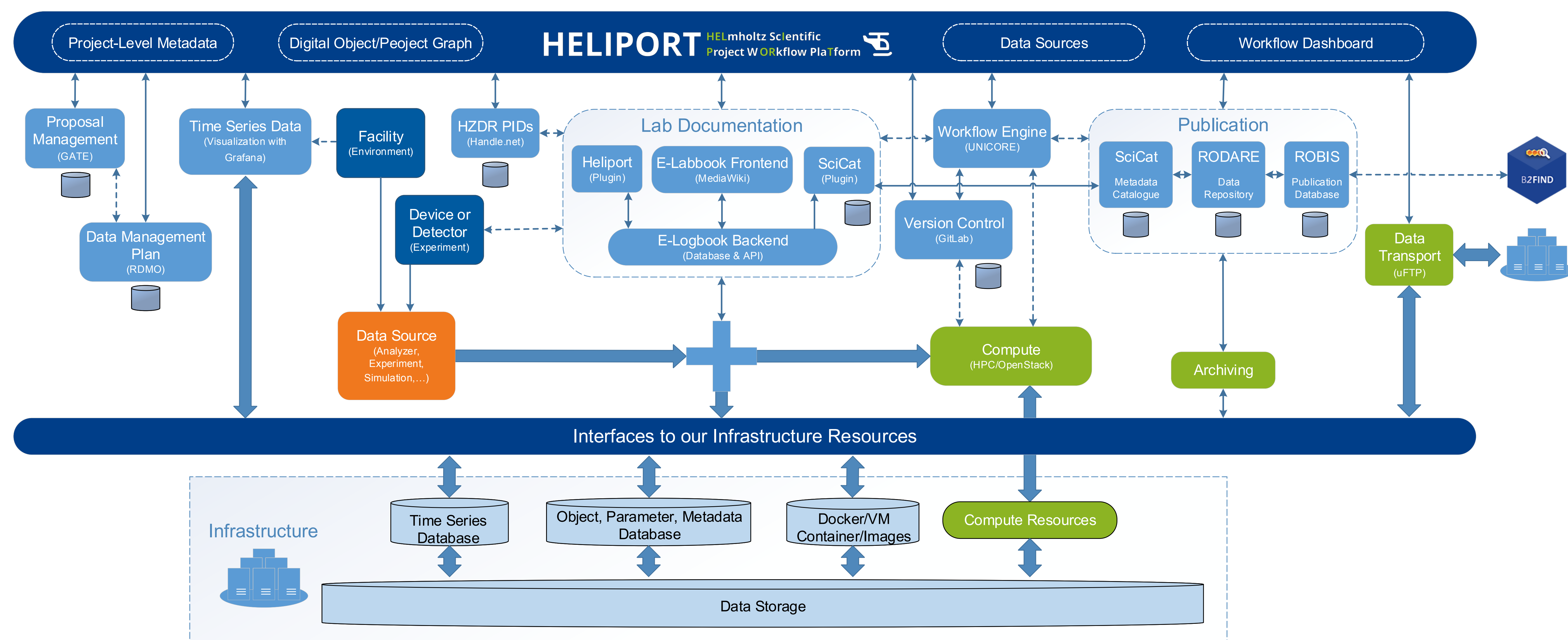
DOI 10.14278/rodare.193

heliport@hzdr.de

www.hzdr.de/fwcc-d



HZDR Data Management Strategy — Top-Level Architecture



Components

- **HELIPORT** (heliport.hzdr.de): The future overall entry point for every (new) user (**GATE** and **RDMO**) to access tools and particular datasets.



- **Lab Documentation** (wiki.hzdr.de): The E-Logbook backend database contains all necessary data as well as APIs for automated documentation of the experiment itself. We provide **MediaWiki** as primary web frontend for manual Documentation.

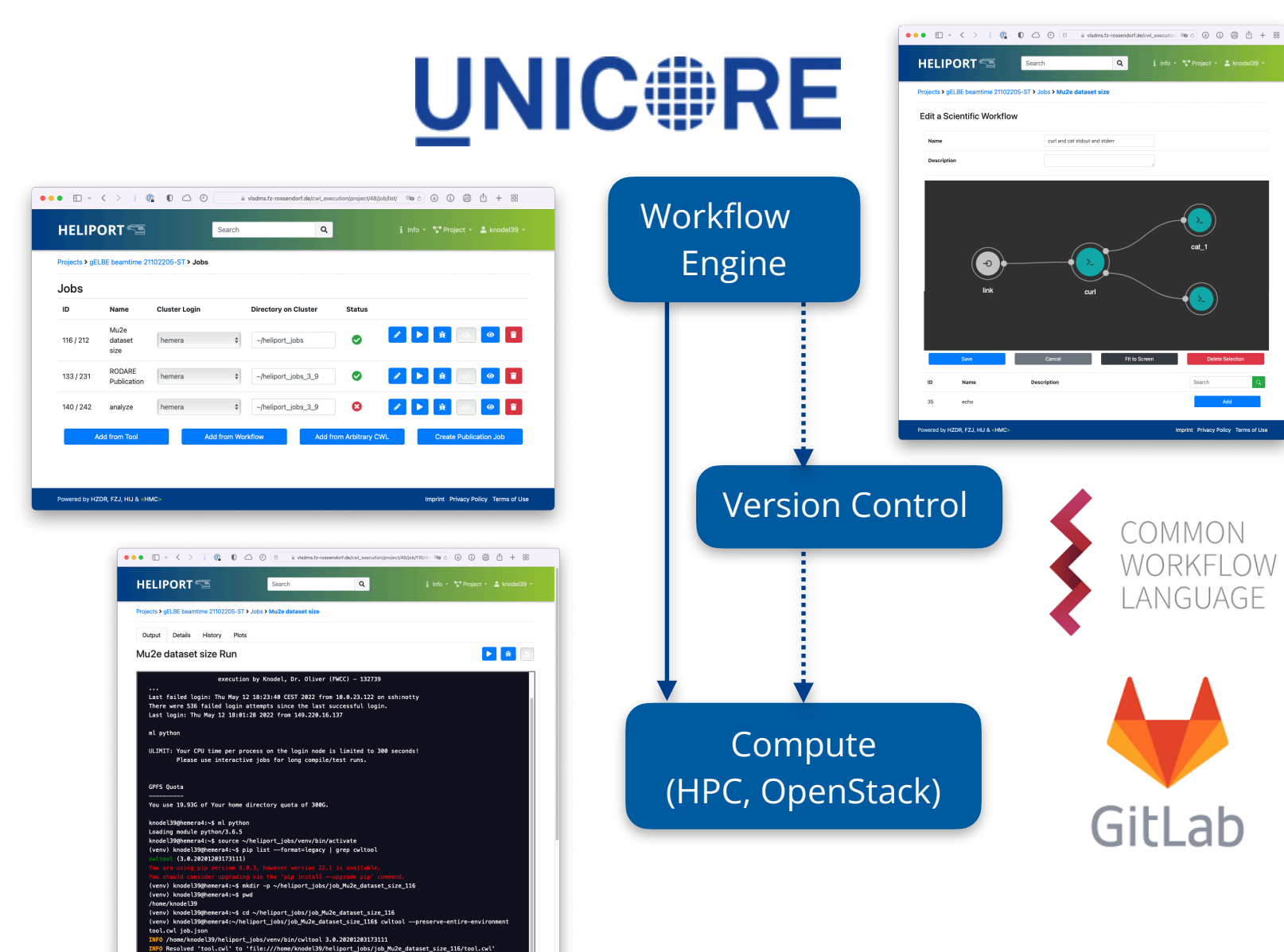


- **Metadata Catalogue SciCat** (scicat.hzdr.de): Additional view which combines metadata from MediaWiki, GATE and the experiment in a searchable application.



- **Data Management Plan (rdmo.hzdr.de)**: The Research Data Management Organizer (RDMO) is a web service to assist structured planning, implementation and administration of the data in a scientific project.

- **Workflows Engine and Monitoring**: We provide UNICORE (Uniform Interface to Computing Resources) as cluster authentication and workflow system:
 - HELIPORT's encapsulation of individual steps in a computational workflow follows the FAIR principles, enables reusability and
 - Analysis and pre-/post-processing steps can be documented and reproduced.



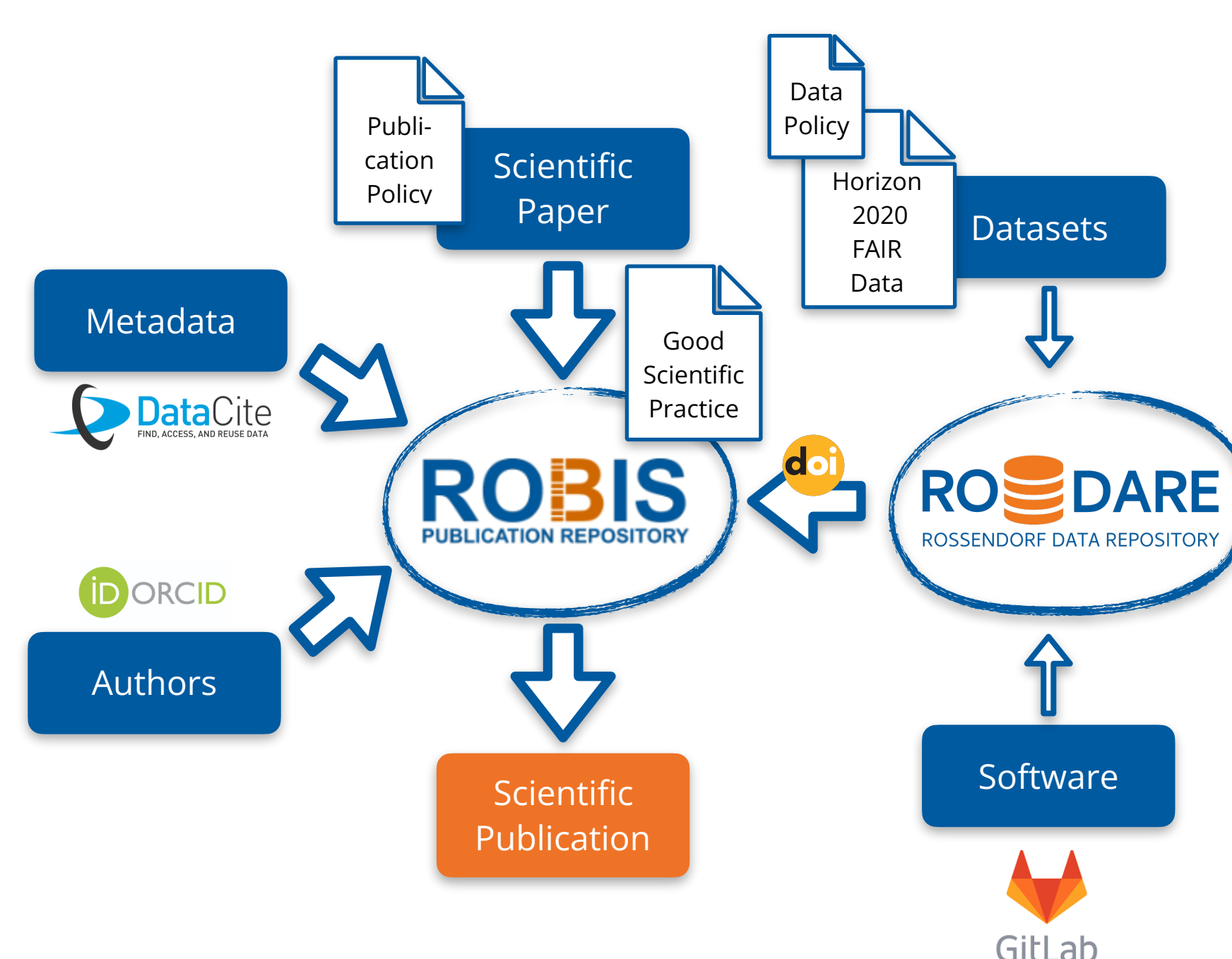
- **Version Control GitLab** (codebase.helmholtz.cloud): Code repository for software and all files under version control.

- **Data Analysis**: Live data analysis with our **jupyter** notebook service at HZDR's Hemera cluster.

- **PaN-Training** (pan-training.eu): The training catalogue for the Photon and Neutron community, provides a portal for trainers and trainees to discover online information and content, including training materials, events and interactive tutorials.



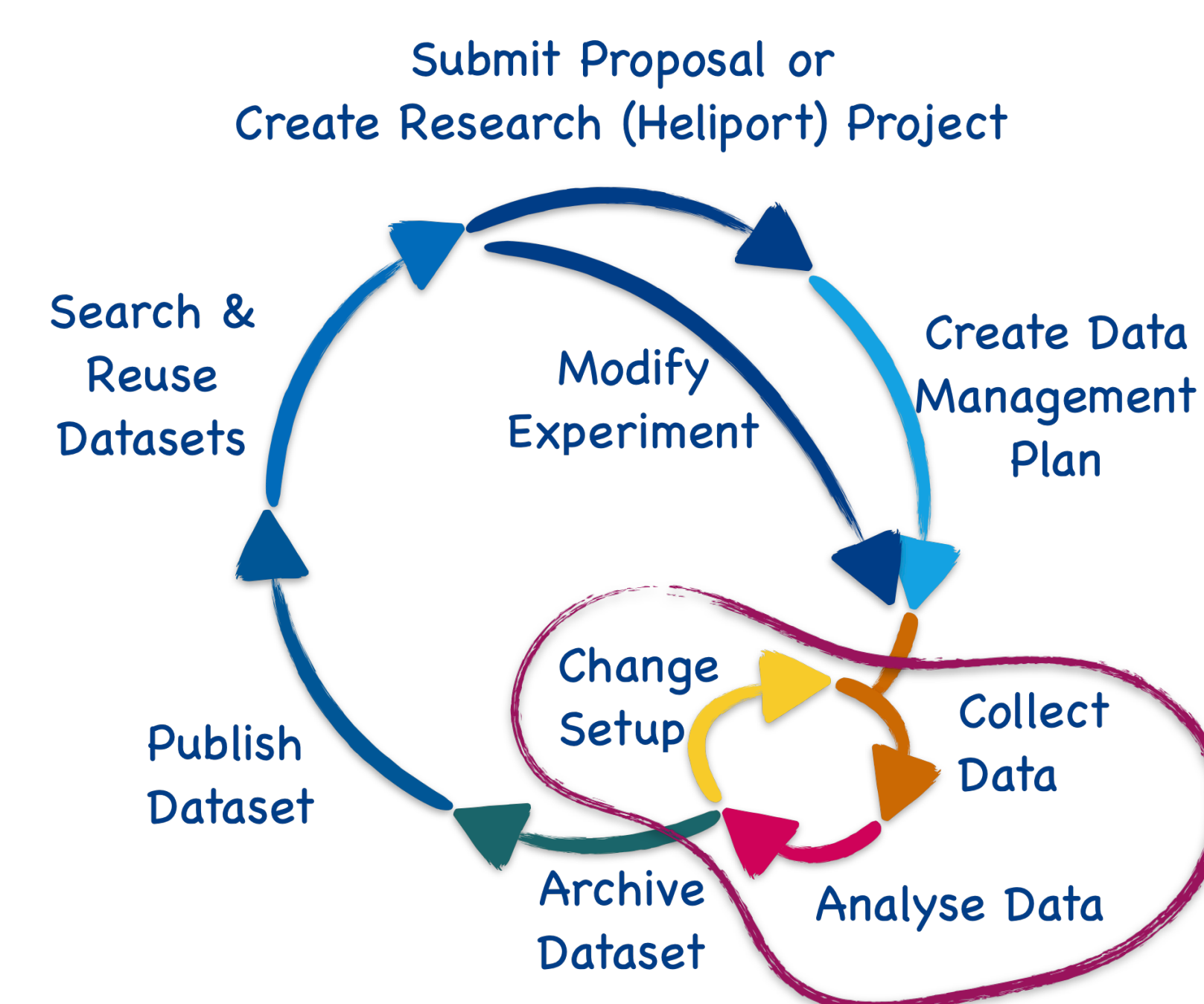
- **(Data) Publication**: At the end of the experiment datasets (raw data, results, surrounding ecosystem,...) can be published using Rodare (rodare.hzdr.de) — even software or workflows...



- **Digital Object and Handles**: HELIPORT interfaces with a local handle.net **hdl**enabled server (e.g. handle.hzdr.de) to enable sustainability.

- Automated generation of uniform, globally unique PIDs for digital objects of all systems, jobs, services, ...
- With digital objects, object relations and landing pages, HELIPORT improves **provenance** and **comprehensibility**.

Full Lifecycle Management



- **Findable Accessible Interoperable Reusable** data management for modern scientific experiments with a wide variety of appropriate software tools:

- electronic lab books,
- interactive analysis,
- publication repositories for code and data,
- scientific workflow management,
- various databases and storages,
- and many more.

- **Uniform and smooth access** to and between all services and systems in the IT ecosystem is necessary to ensure:

- comprehensibility,
- machine-actionability and
- collaborative teamwork.

- **HELIPORT** is designed to be configurable and adaptable for the IT infrastructure of a research center to offer a holistic view of an experiment.

