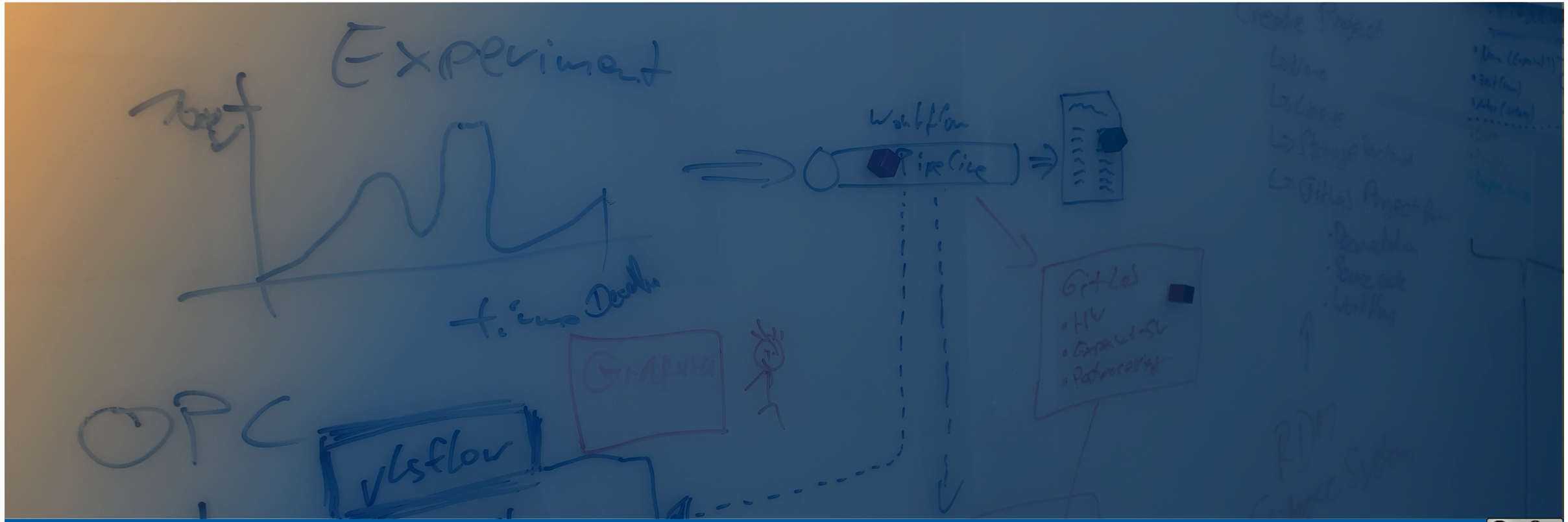
HELIPORT

HZDR
HELMHOLTZ ZENTRUM
DRESDEN ROSSENDORF



CC BY

HELmholtz ScIentific Project WORkflow PlaTform

Status and Roadmap

December 2020

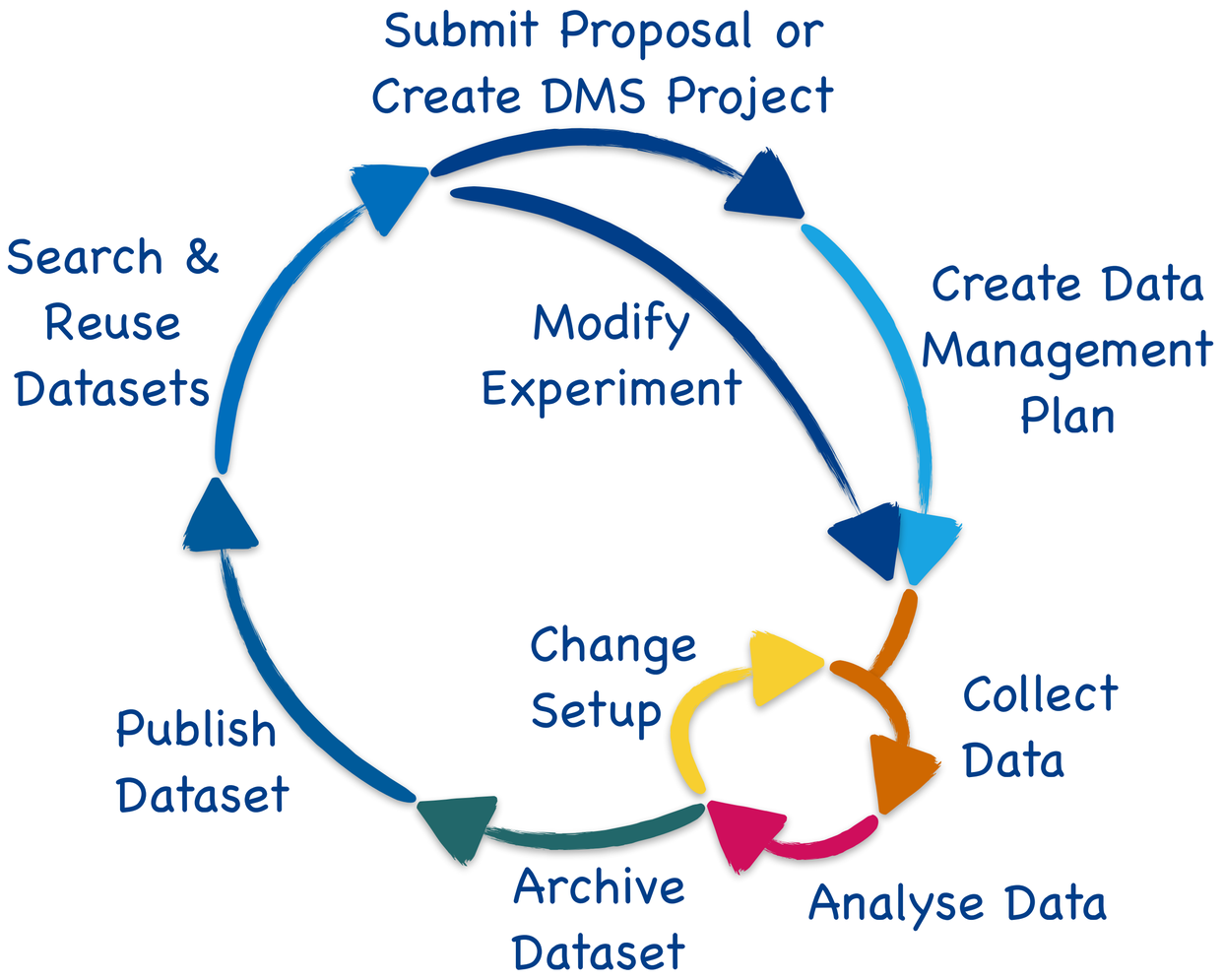
Oliver Knodel // contact: o.knodel@hzdr.de

DRESDEN concept

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

— We support all stages with tools:
  - electronic lab books,
  - interactive analysis,
  - (automated) publication,
  - scientific workflow management.

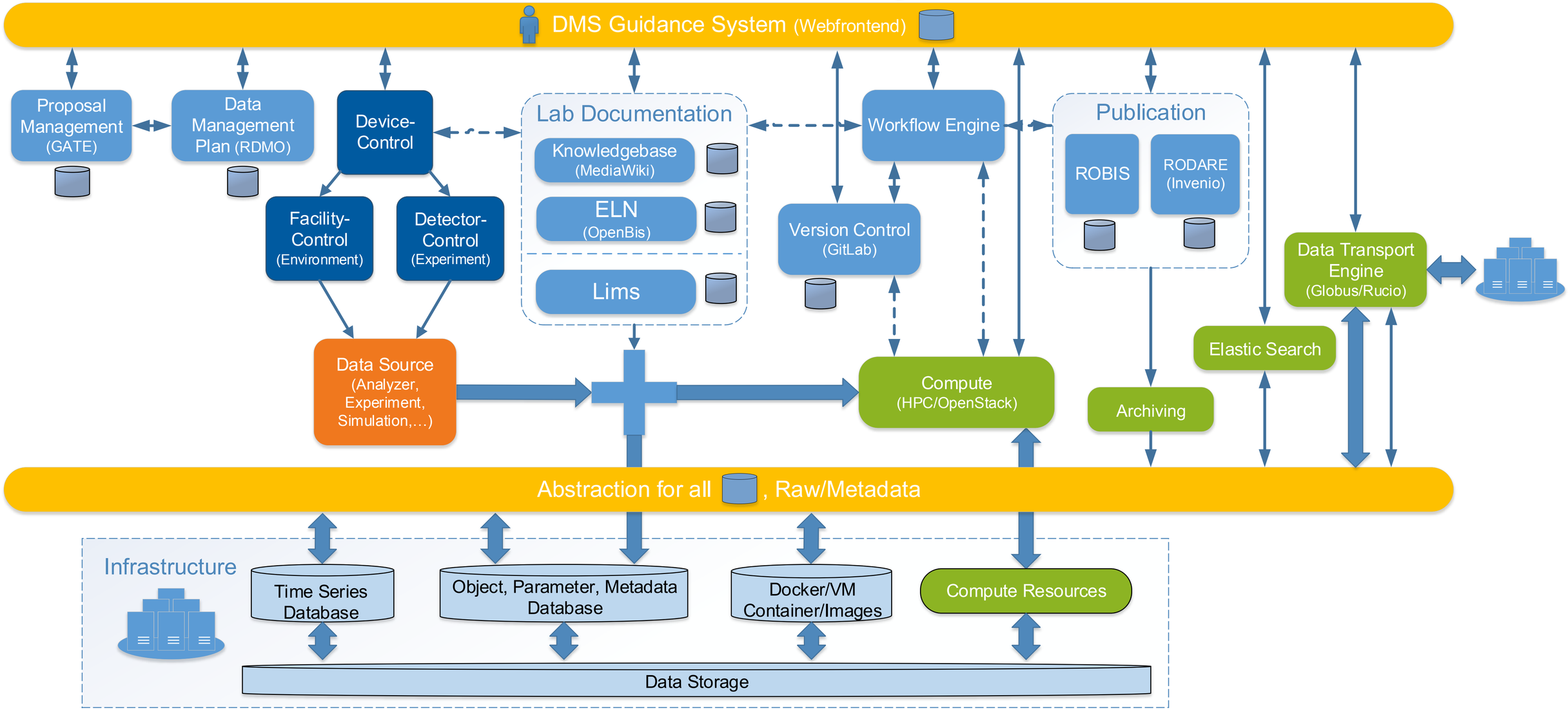
— **Requirement**: Transfer data as early as possible to our data center.



HELIPORT DRESDEN concept HZDR

HZDR Data Management Strategy

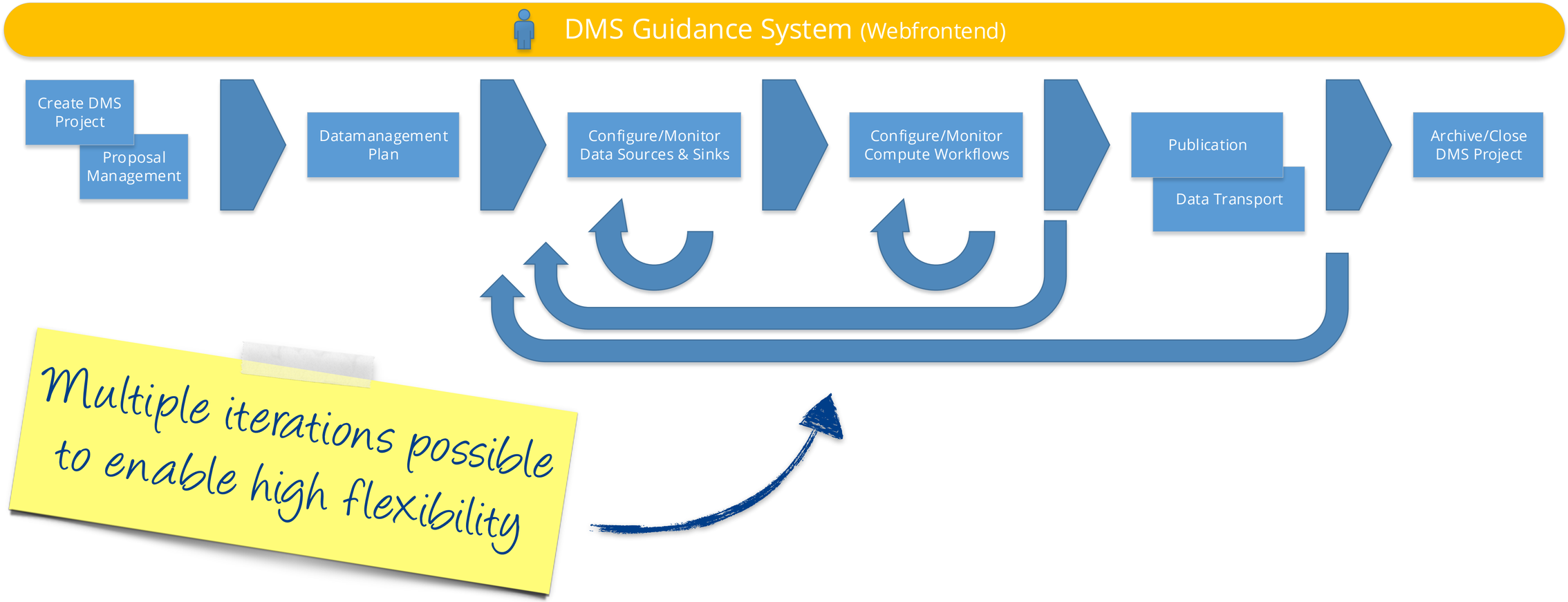
HELIPORT



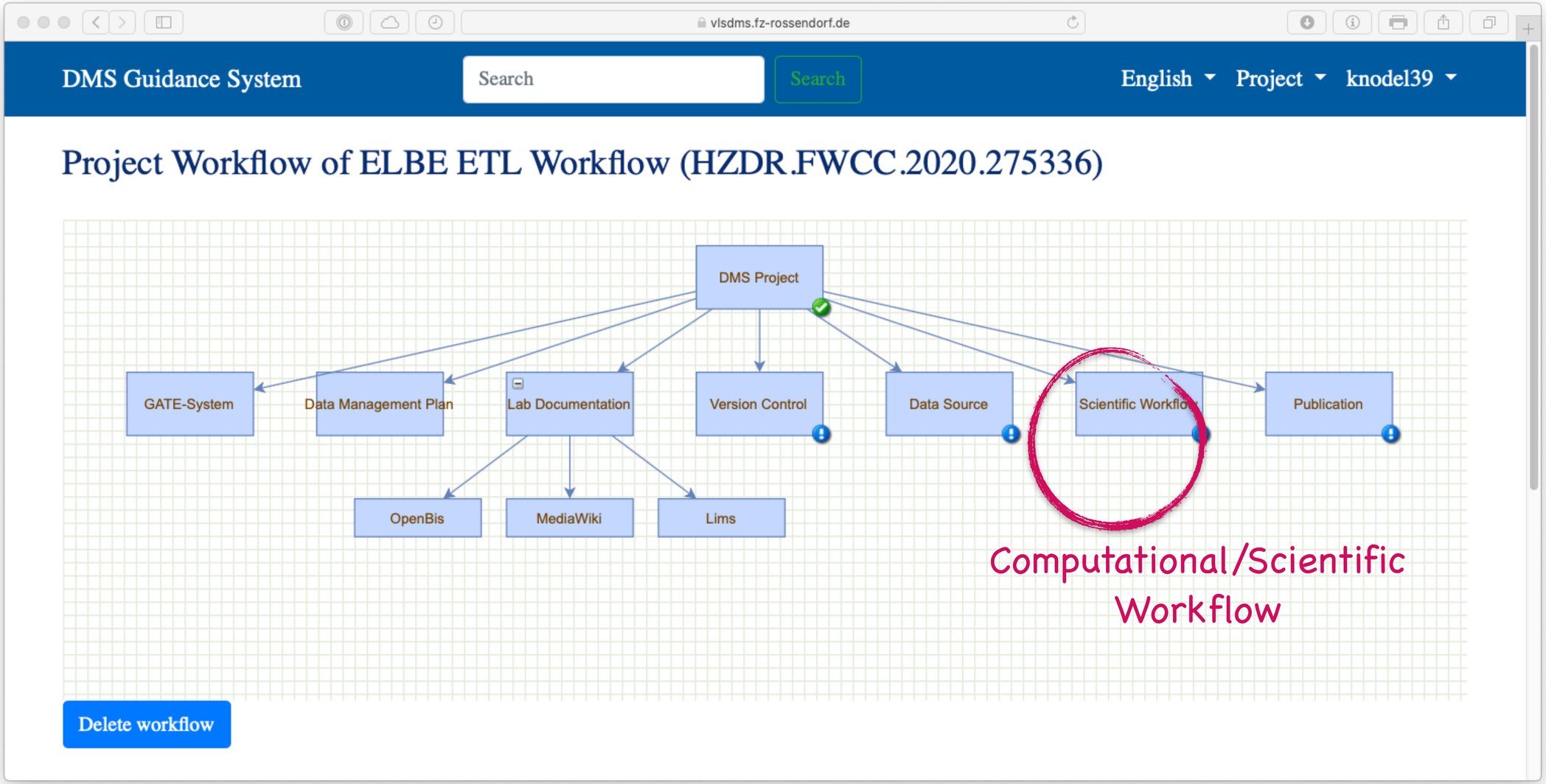
HELIPORT DRESDEN concept HZDR

HZDR Data Management Strategy

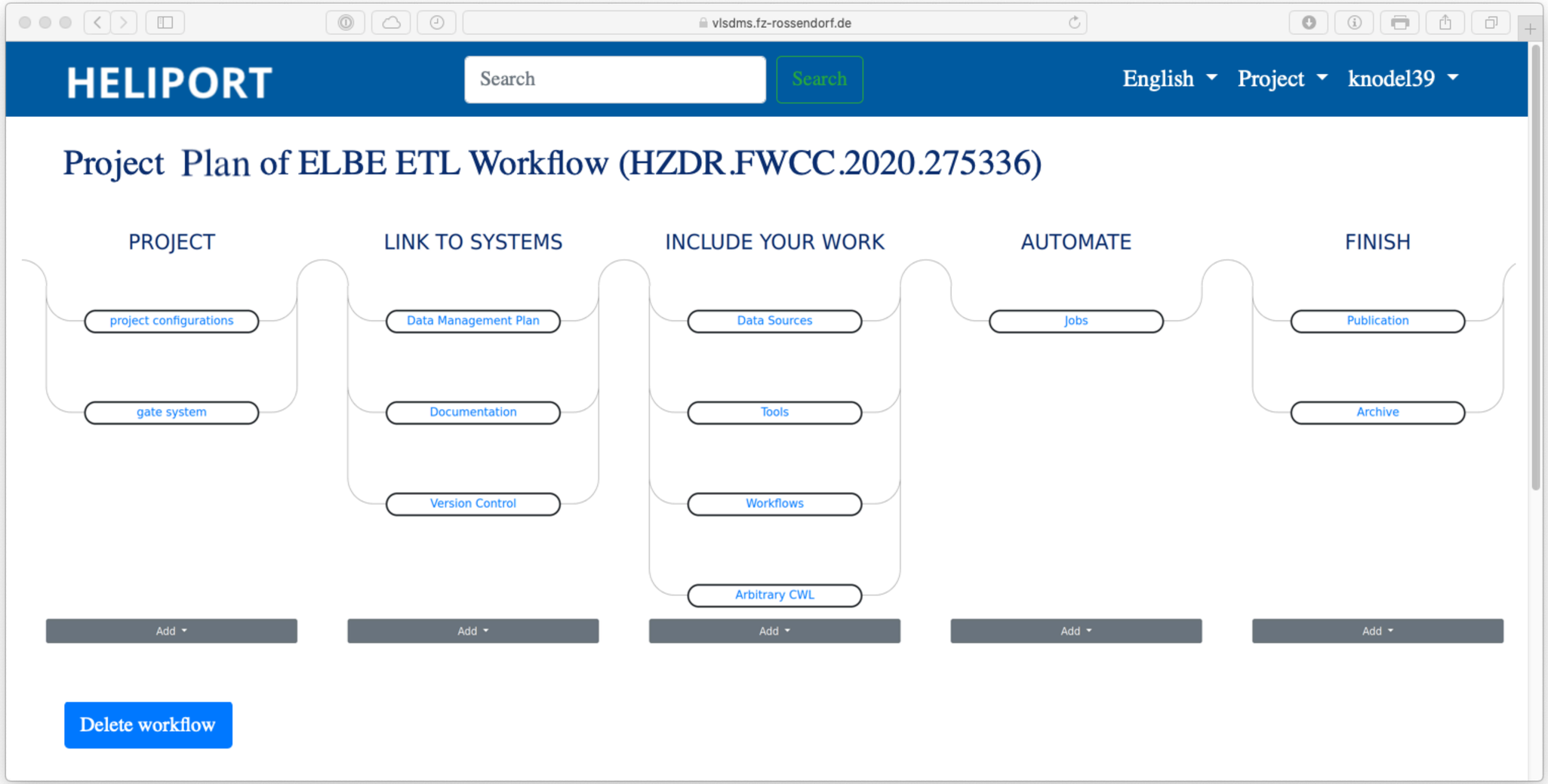
A Guidance System Should Enable a **Workflow on Project Level**



HELmholtz Scientific Project WORkflow PlaTform — Version 0.2.0



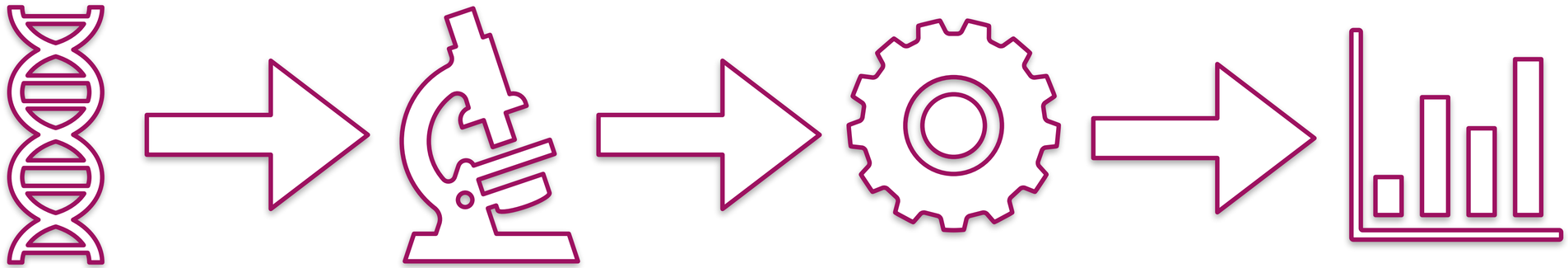
HELmholtz Scientific Project WORkflow PlaTform — Version 0.3.4



Scientific Workflow

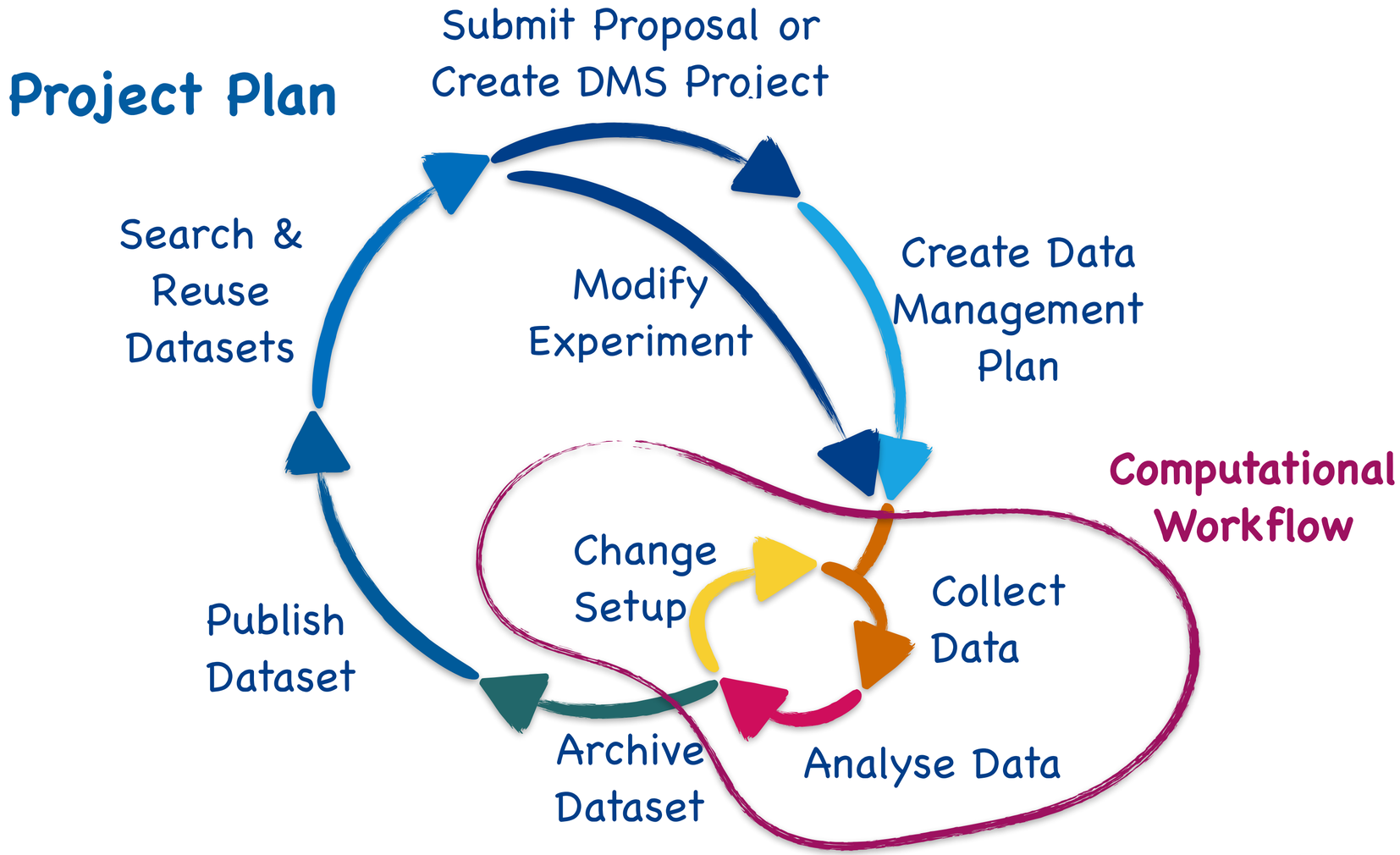
Computational workflows describe the complex multi-step methods that are used for data collection, data preparation, analytics, predictive modelling, and simulation that lead to new data products.

Carole Goble , Sarah Cohen-Boulakia, Stian Soiland-Reyes, Daniel Garijo, Yolanda Gil, Michael R. Crusoe, Kristian Peters, Daniel Schober: “FAIR Computational Workflows,” Data Intelligence, vol. 23, pp. 0–2, 2019.



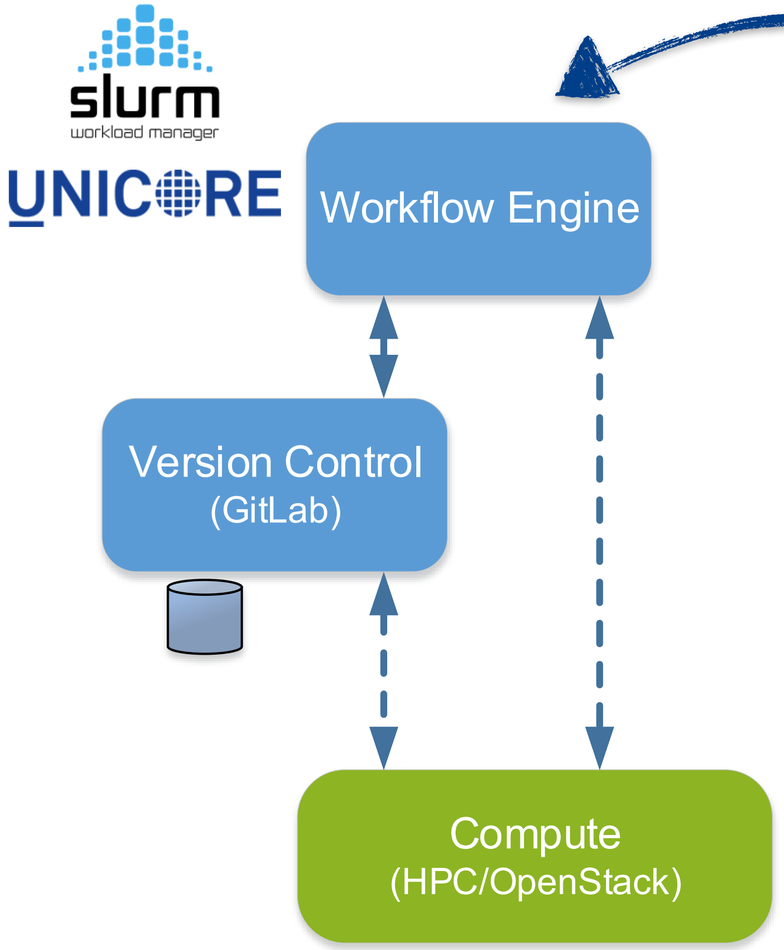
Scientific Workflow

HELIPORT has an build-in Integration of Scientific Workflows



HELIPORT DRESDEN concept HZDR

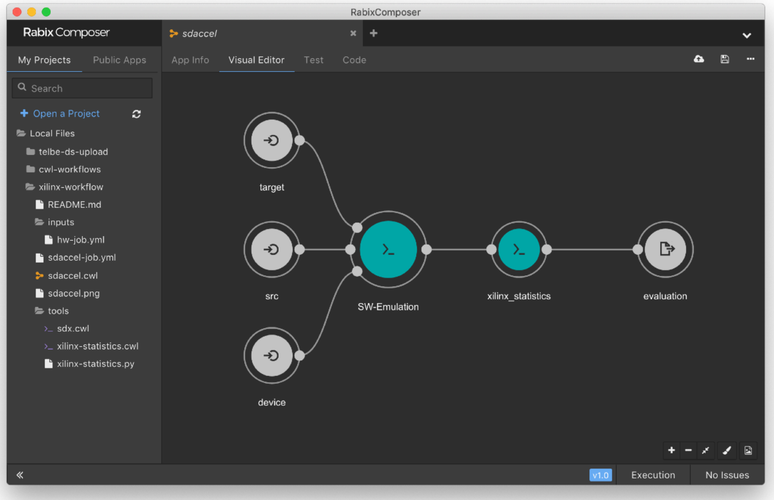
Scientific Software Development and Workflows

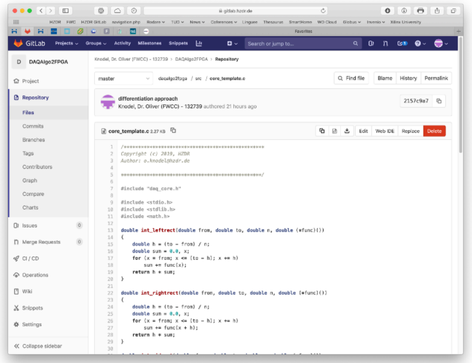


COMMON
WORKFLOW
LANGUAGE

— Analysis and Pre-/Postprocessing steps needs to be:
  - Documented and
  - Reproducible

— Capsuling every step in a workflow adapts the **FAIR** principles.

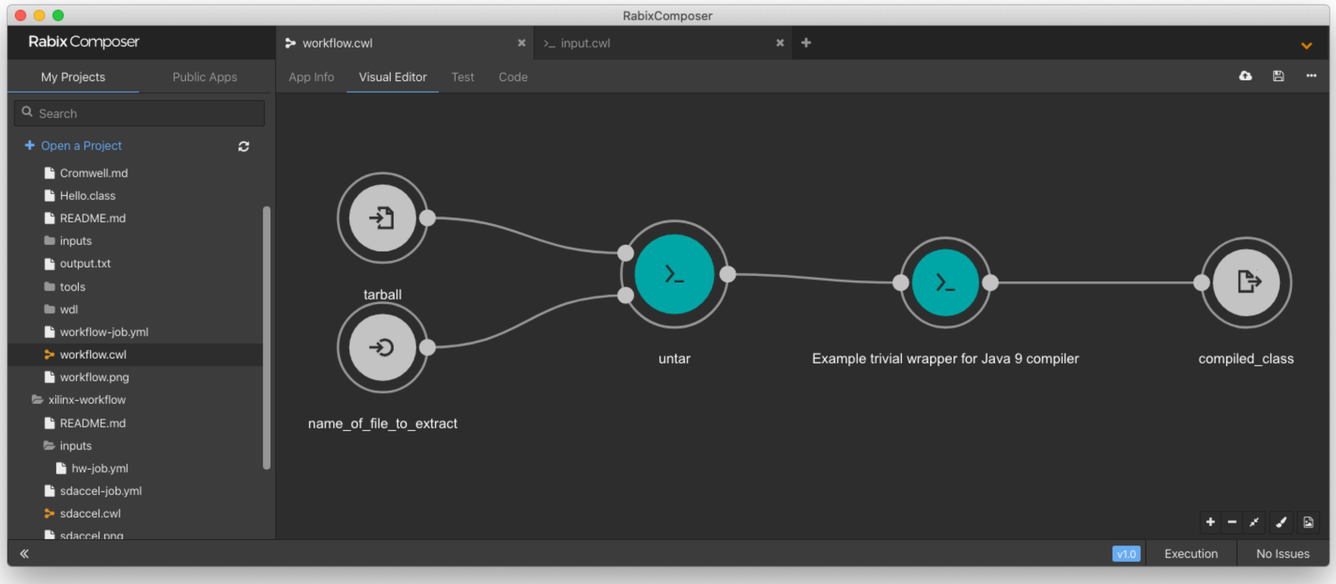




GitLab

HELIPORT DRESDEN concept HZDR

Scientific Workflow Example

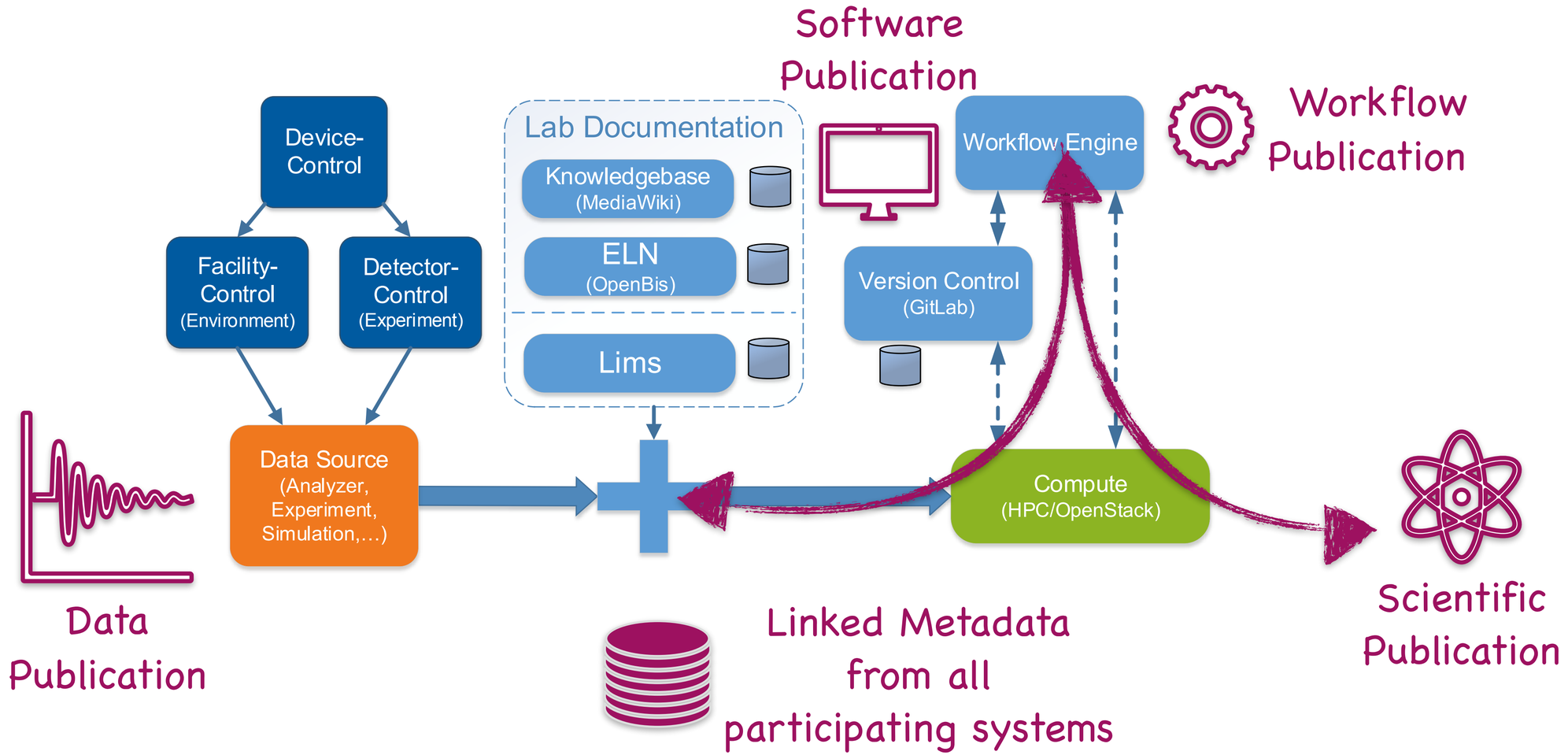


The description of the workflow itself...

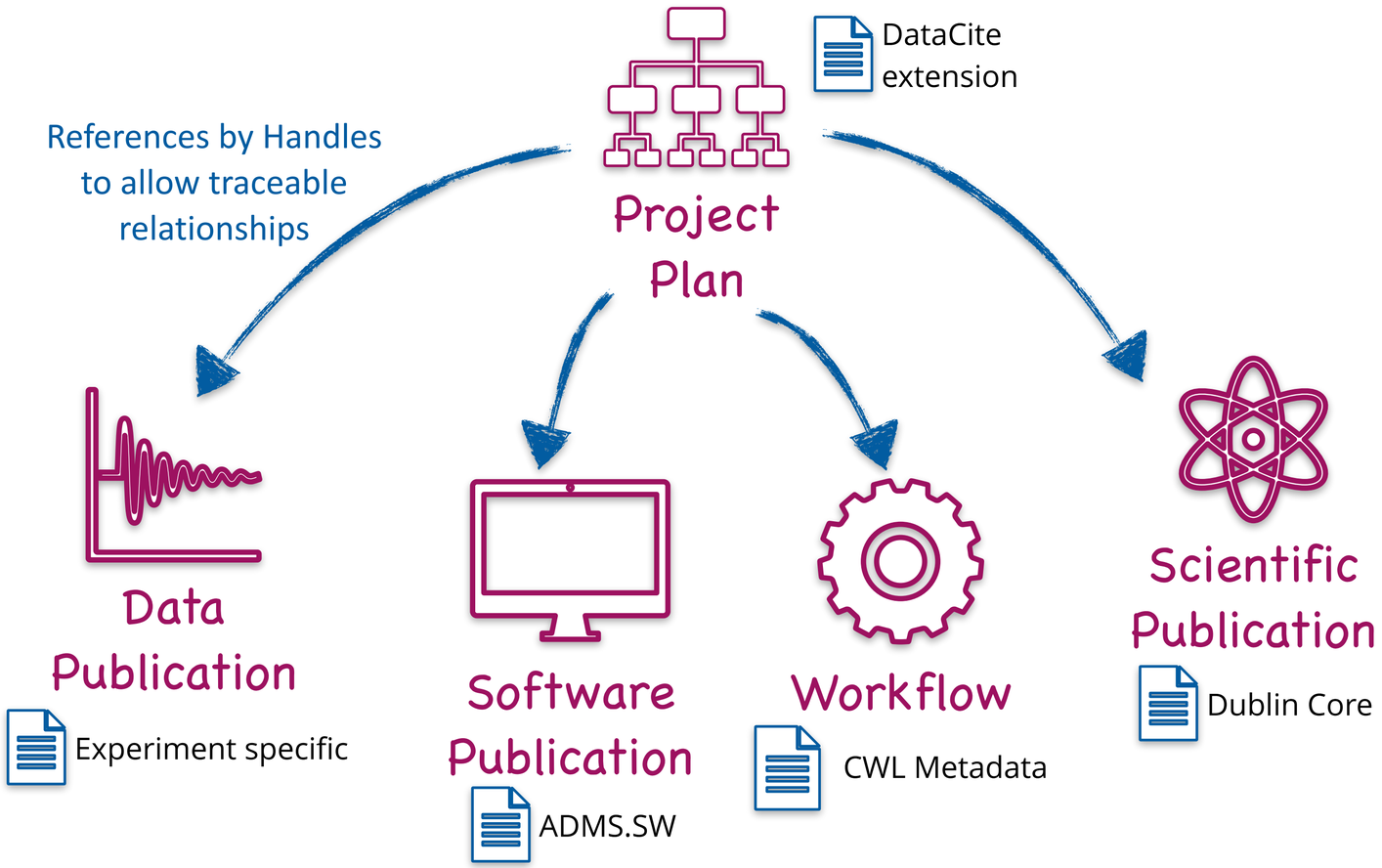
HELIPORT DRESDEN concept HZDR

Different Types of Publications

An Example for a Complete FAIR Research Experiment

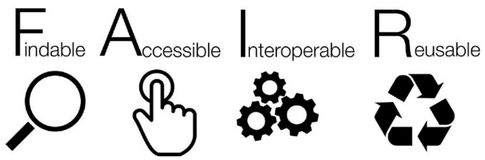


Possible Metadata Standards



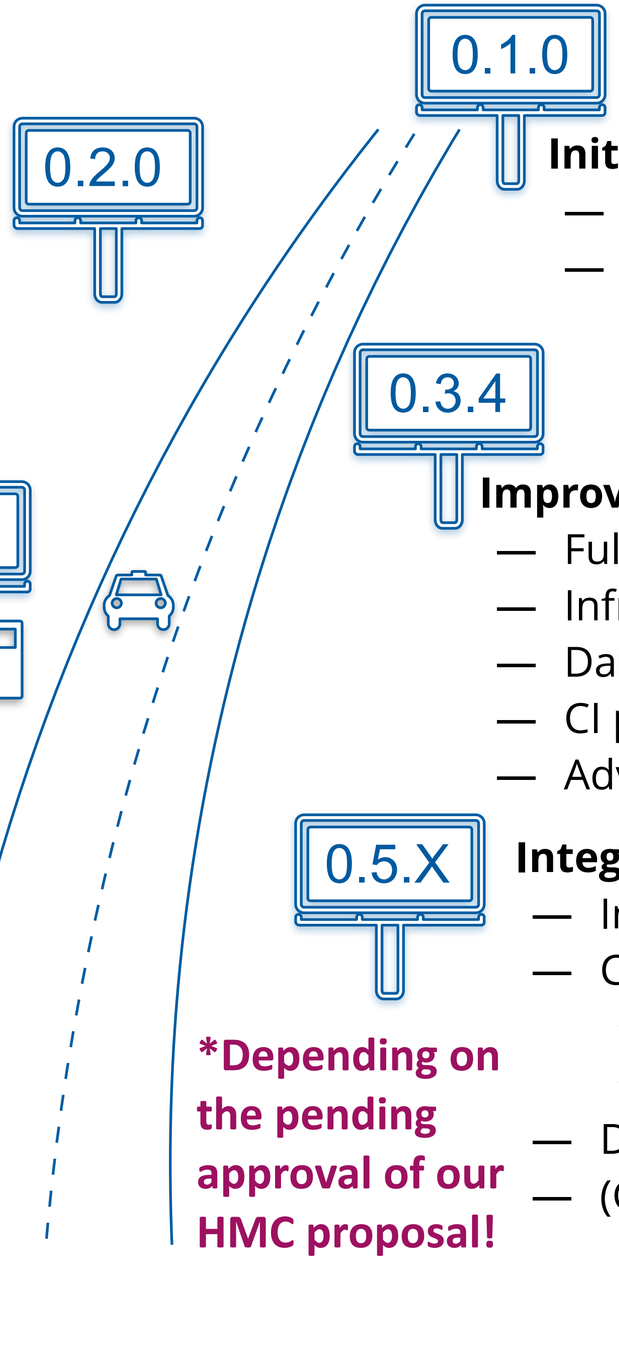
Our Objective

- In all stages as much information as possible is stored in our HELIPORT PostgresDB.
- Metadata is used to transfer the information between different (unknown) systems.
- HELIPORT is designed to export metadata for known schemes.
- Every published object should have describing metadata in an open and widely used format to be



HELIPORT DRESDEN concept HZDR

HELIPORT Roadmap



First Draft: Project Plan (August 2020)
- Project and user management
- Configurable stages
- REST API for proposal information
- CWL visualization prototype

Initial Version (June 2020)
- Webinterface with user authentication (LDAP)
- DMS Projects and additional proposal information from the HZDR GATE database

Improved Project Plan (December 2020)
- Fully configurable stages and modules
- Infrastructure and database updates
- Daily proposal database update
- CI pipeline for test and deployment
- Advanced logging and monitoring

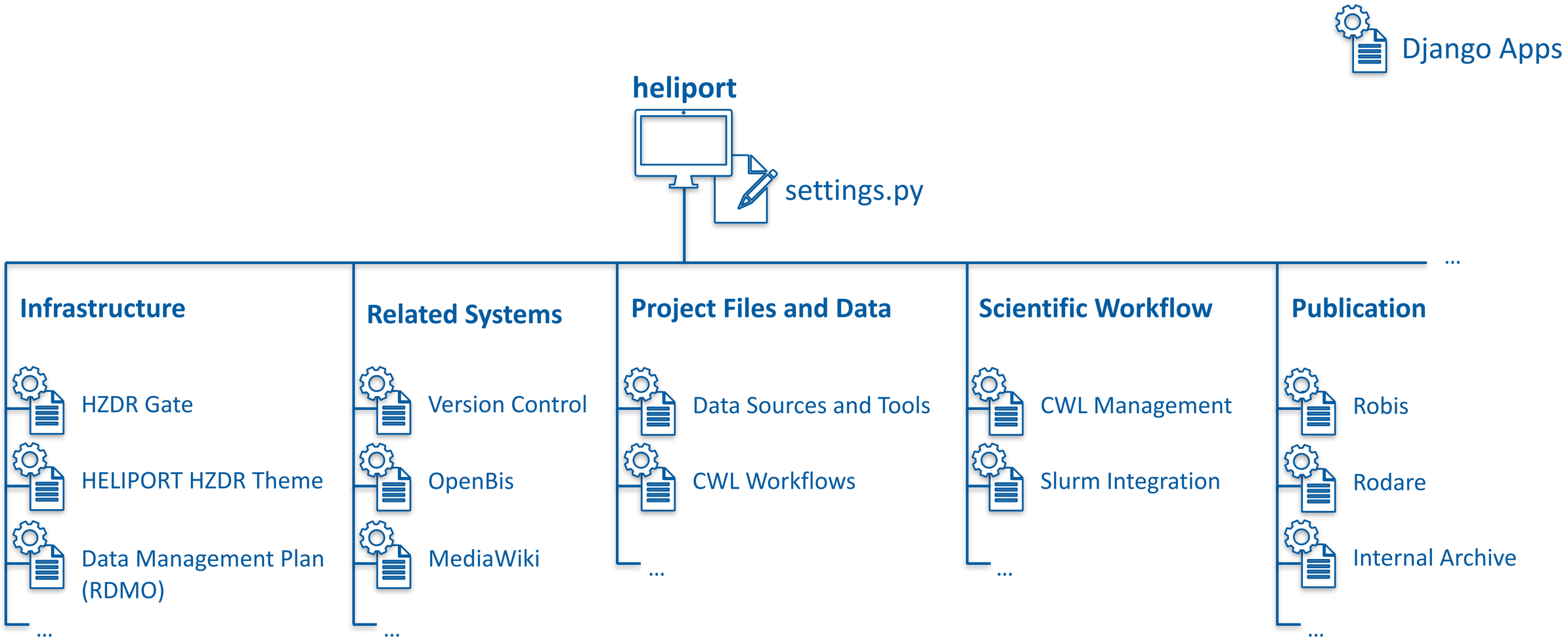
Modular Structure (February 2021)
- Subdivision of the stages into modular and configurable Django apps to allow individual extensions
- Refactoring of the project
- Documentation using GitLab pages

Integration of various Apps and Features*
- Integration of (different) Metadata Schemas
- Computational/Scientific workflow execution
  - Workflow management and monitoring
  - CWL support
- Data Management Plan support
- (Global) Handle management

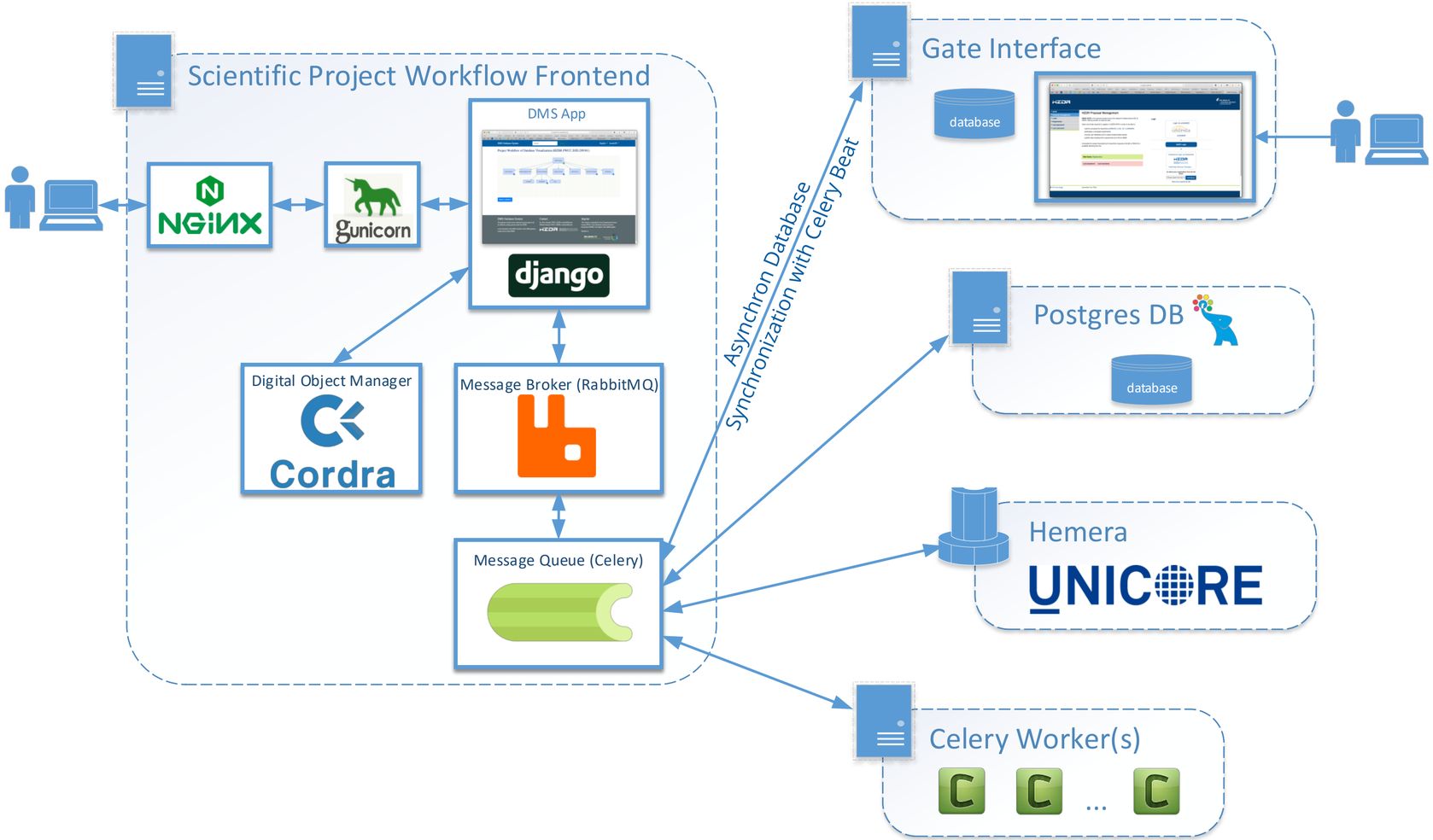
Documentation of a TELBE user experiment*
- Integration of all related data sources
- Automated workflow initiation
- Publication of all data products

***Depending on the pending approval of our HMC proposal!**

Modular HELIPORT Design — Draft for Version 0.4.0



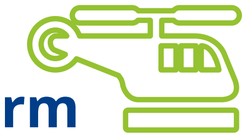
HELIPORT Architecture Concept



HELIPORT DRESDEN concept HZDR

HELIPORT

HELmholtz Sc**I**entific
Project W**OR**kflow Pla**T**form



https://vlsdms.fz-rossendorf.de

HELIPORT DRESDEN concept HZDR