

# HELIPORT

## A Potential Platform for Running Digital Twins FAIRly

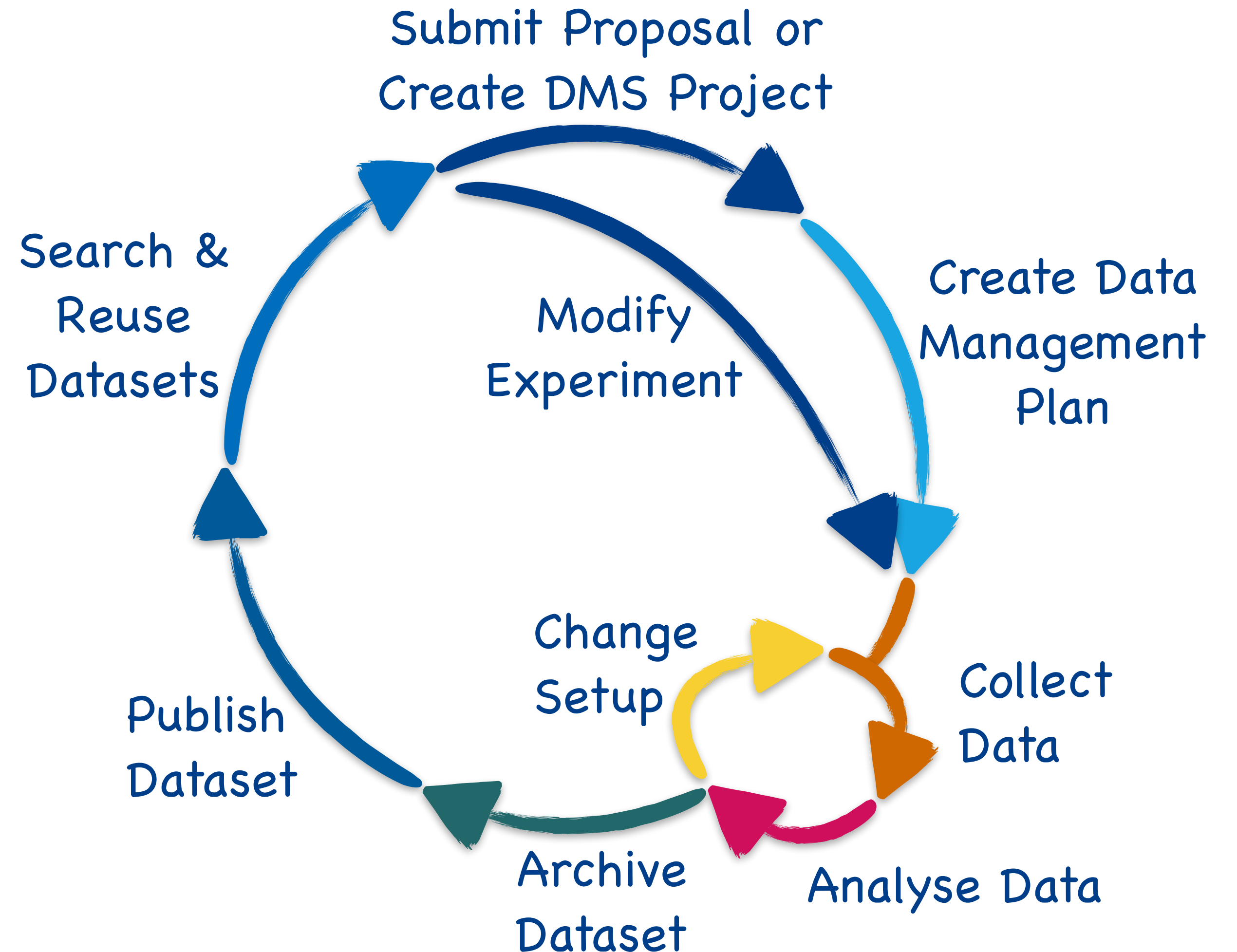
LEAPS Integrated Platform Workshop, Digital Twinning Session, 11 May 2021

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

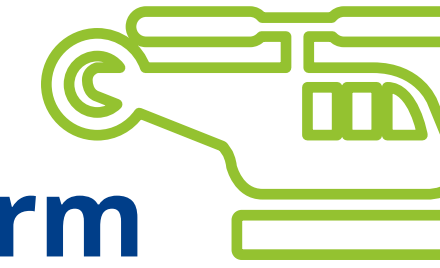


# Our Challenge: An End-to-End Digital Data Lifecycle

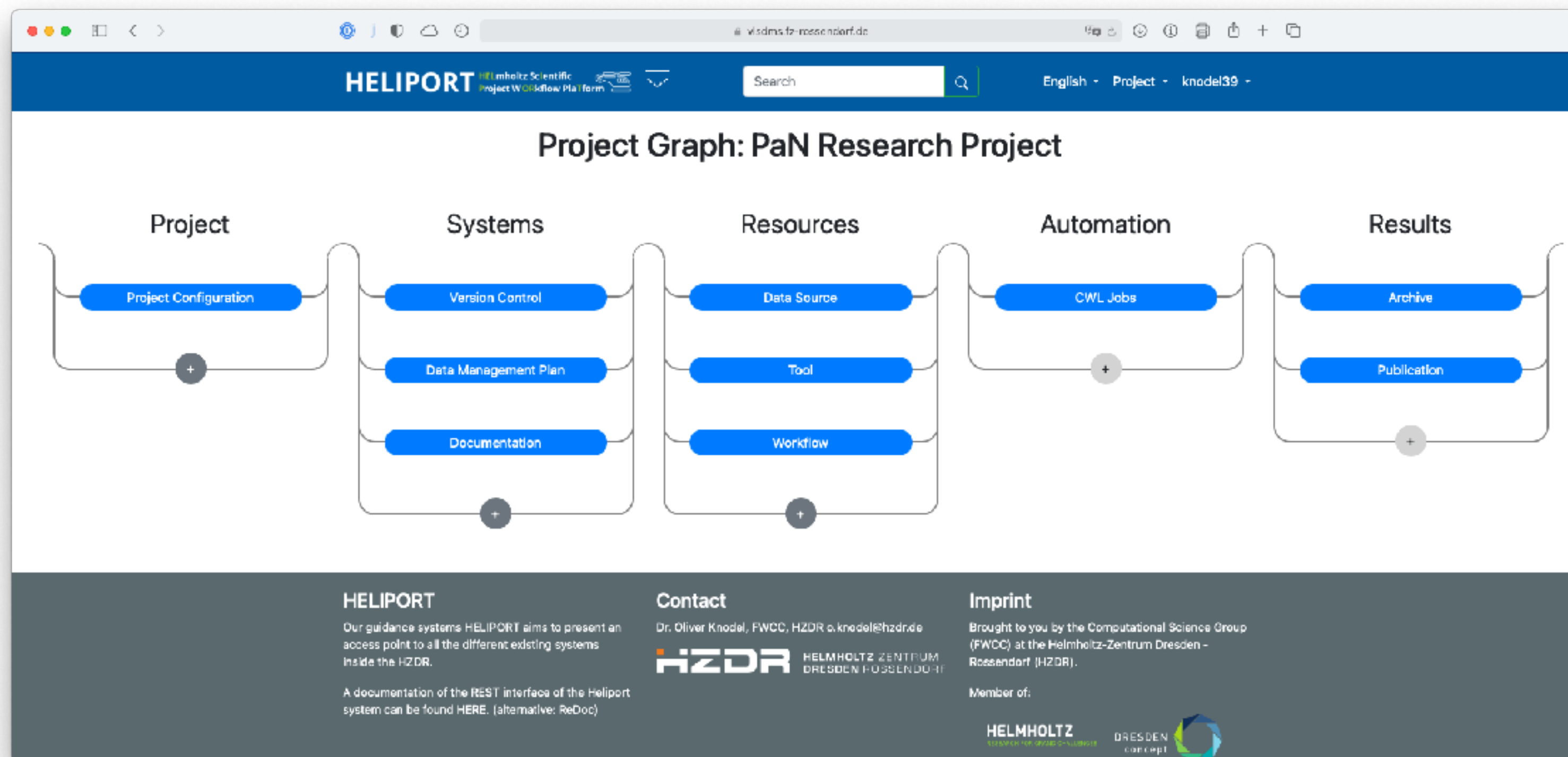
- We support many steps of a research experiment with tools:
  - electronic lab books,
  - interactive analysis,
  - publication of datasets,
  - scientific workflow management,
  - Handle generation and management.
- A uniform access to all services and systems is necessary.
- The documentation of all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



# HELIPORT HELmholtz Scientific Project WORKflow PlaTform

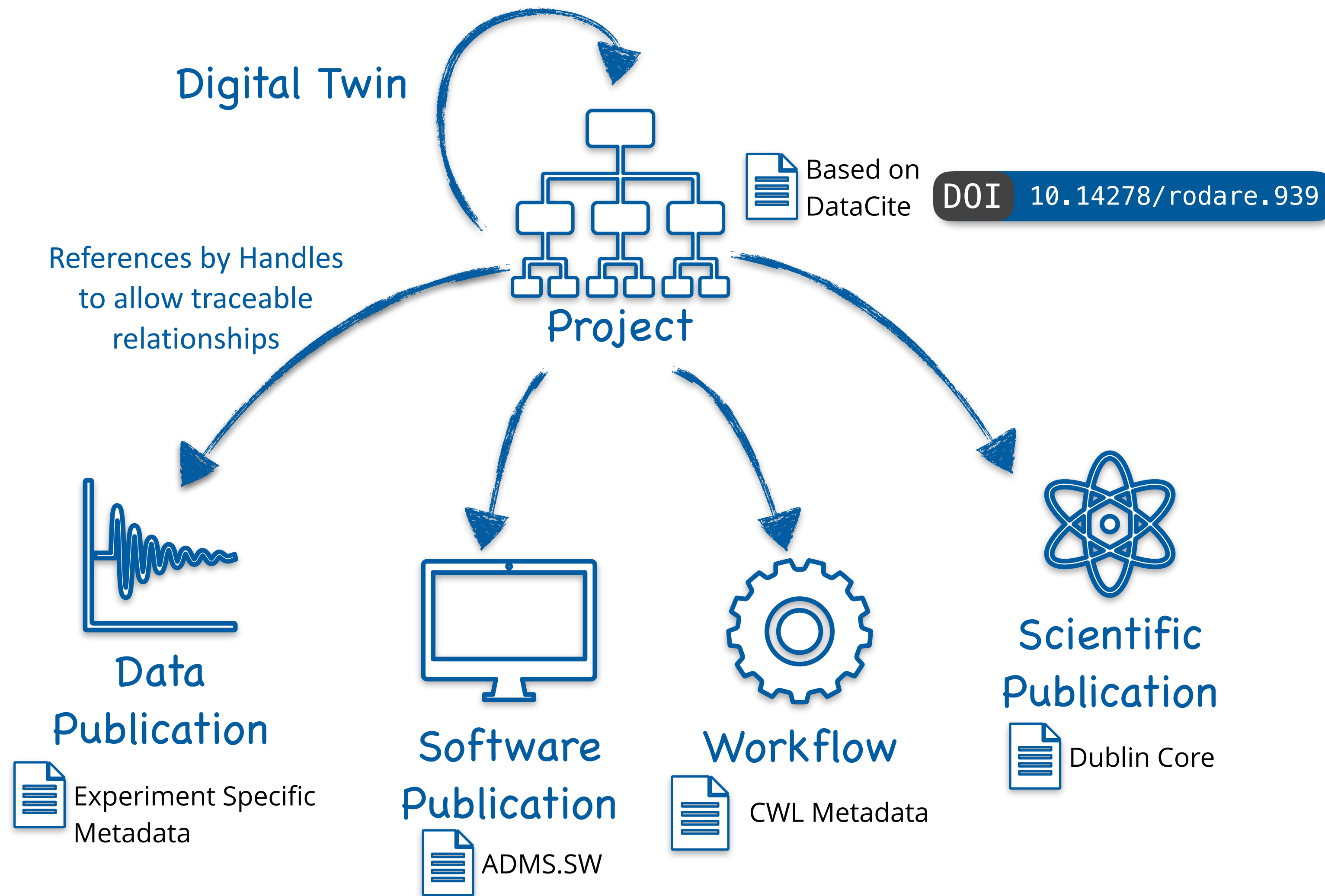


“ The HELIPORT project aims at developing a platform which accommodates the **complete life cycle** of a scientific project and links all corresponding programs, systems and workflows to create a more **FAIR** and comprehensible project description.



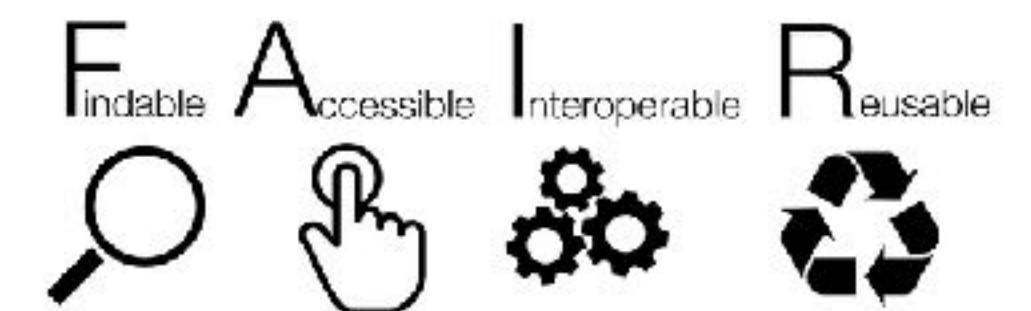
```
{
  "namespaces": {
    "datacite": "http://purl.org/spar/datacite/",
    "rdfs": "http://www.w3.org/2000/01/rdf-schema#",
    "heliport": "https://heliport/schema/",
    "time": "http://www.w3.org/2006/time#",
    "dc": "http://purl.org/dc/terms/"
  },
  "heliport:project_id": 28,
  "datacite:hasIdentifier": "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 09:14:34.296524-00:00",
  "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://dd",
      "rdfs:label": "Test"
    }
  ],
  "heliport:has_DataManagementPlan": [
    {
      "heliport:data_management_plan_id": 6,
      "datacite:uri": "https://dddd",
      "datacite:hasDescription": "dddd"
    }
  ],
  "heliport:has_Documentation": [
    {
      "heliport:documentation_id": 7,
      "datacite:uri": "https://dddd",
      "heliport:documentation_system": "MediaWiki",
      "datacite:hasDescription": "dddd"
    }
  ],
  "heliport:has_DataSource": [
    {
      "heliport:data_source_id": 11,
      "datacite:uri": "http://ddd",
      "heliport:use_computer": null,
      "rdfs:label": "ddd",
      "datacite:hasDescription": ""
    }
  ],
  "heliport:has_Archive": [
    {
      "heliport:archive_id": 4,
      "datacite:hasDescription": "ret"
    }
  ],
  "heliport:has_Publication": [
    {
      "heliport:publication_id": 6,
      "datacite:hasDescription": "ret"
    }
  ]
}
```

# Heliport Metadata Ecosystem



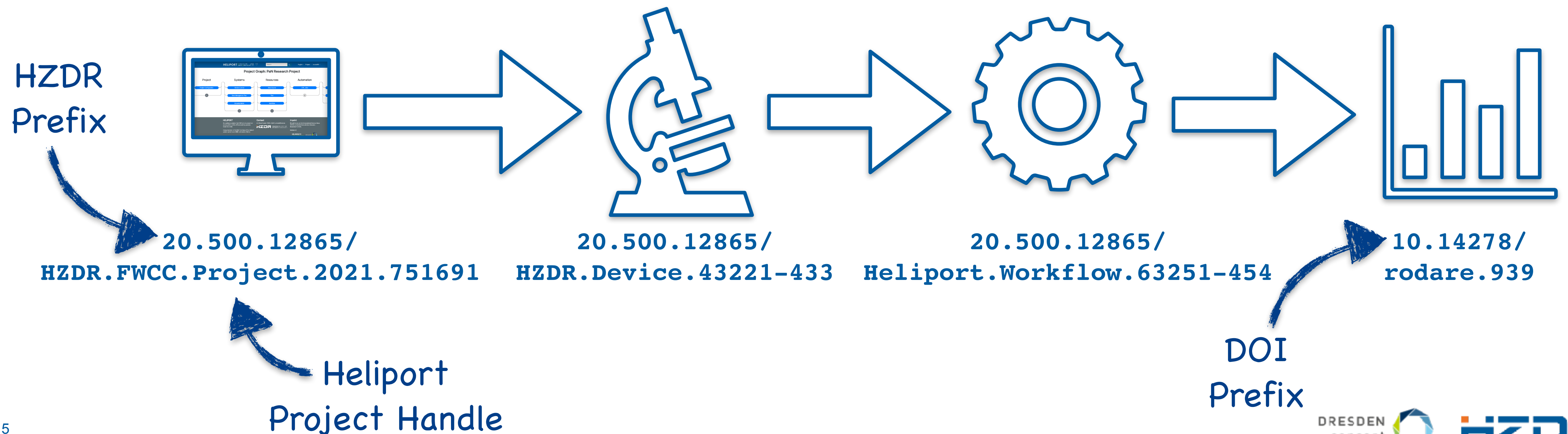
## Our Objective

- In all stages of an experiment Heliport combines information about involved services with PIDs.
- Metadata (stored *near* the PID) is used to transfer information between different systems and a documentation of the project-level workflow is possible.
- In the end every digital object should have an uniform PID, describing metadata in an open and widely used format to be



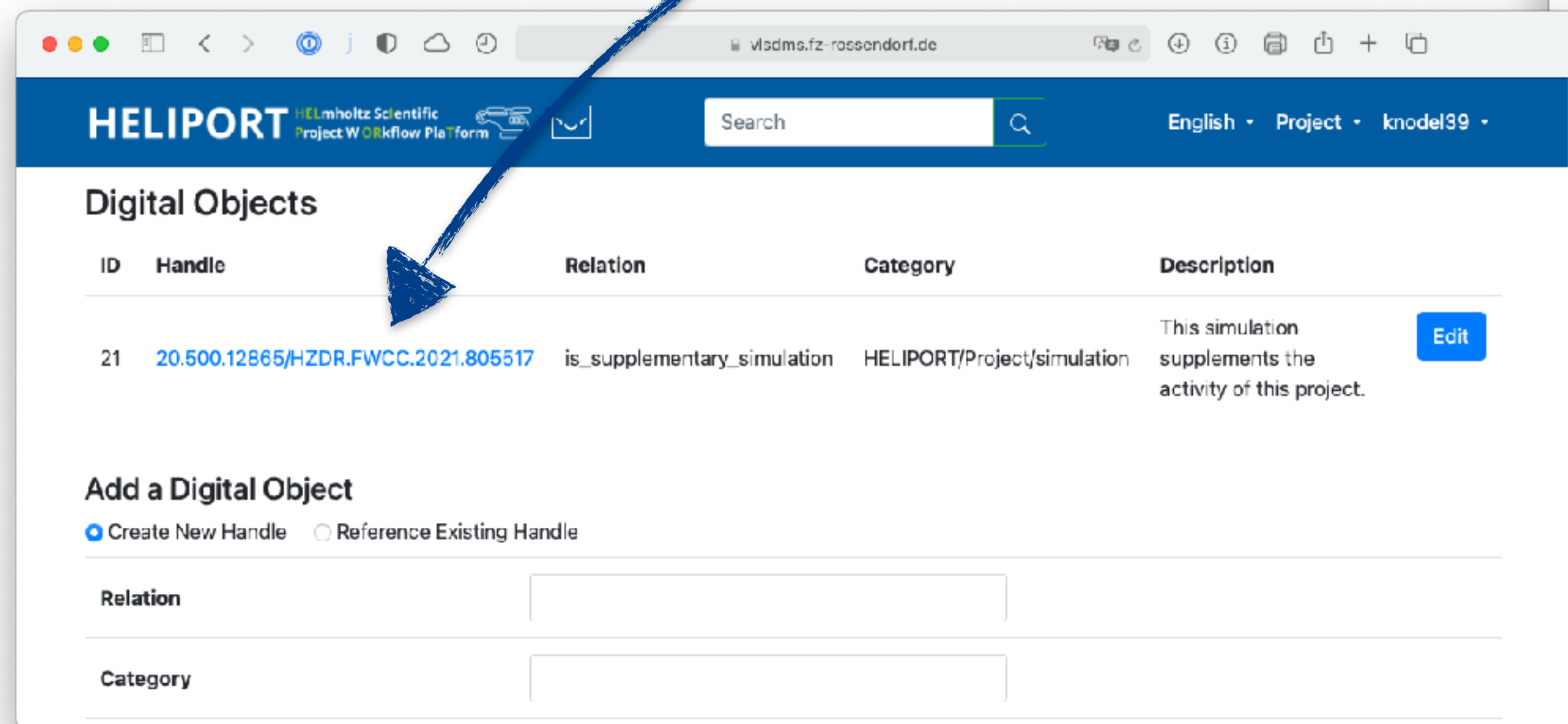
# Handle Management Support in Heliport

Heliport is linked with our local Handle-Server ([handle.hzdr.de](http://handle.hzdr.de)) **hdlenabled** and generates uniform PIDs (resolvable using [hdl.handle.net](http://hdl.handle.net)) from and for various systems and services. Associated information can be changed as needed without changing the identifier.



# Supporting Digital Twins in Heliport

- We can map all kinds of digital workflows to Heliport and we can manage simulations,
- In Heliport a project can reference a digital twin (`is_supplementary_simulation`) as digital object with a unique PID.



**HELIPORT** HELmholtz Scientific Project Workflow Platform

Search English Project knodel39

### Digital Objects

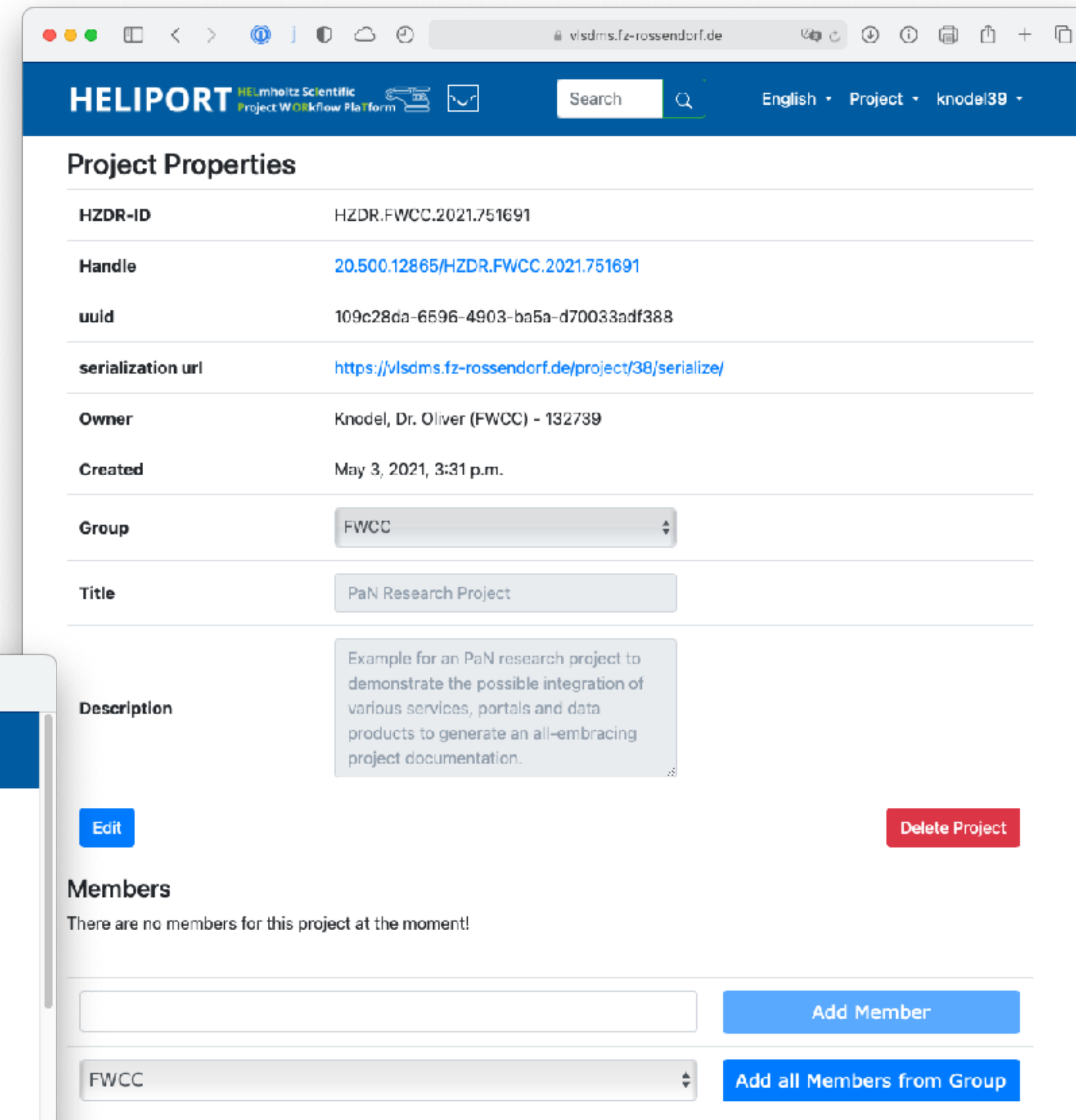
ID	Handle	Relation	Category	Description	
21	<a href="#">20.500.12865/HZDR.FWCC.2021.805517</a>	is_supplementary_simulation	HELIPORT/Project/simulation	This simulation supplements the activity of this project.	<a href="#">Edit</a>

**Add a Digital Object**

Create New Handle  Reference Existing Handle

Relation

Category



**HELIPORT** HELmholtz Scientific Project Workflow Platform

Search English Project knodel39

### Project Properties

HZDR-ID	HZDR.FWCC.2021.751691
Handle	<a href="#">20.500.12865/HZDR.FWCC.2021.751691</a>
uuid	109c28da-6596-4903-ba5a-d70033adf388
serialization url	<a href="https://visdms.fz-rossendorf.de/project/38/serialize/">https://visdms.fz-rossendorf.de/project/38/serialize/</a>
Owner	Knodel, Dr. Oliver (FWCC) - 132739
Created	May 3, 2021, 3:31 p.m.
Group	FWCC
Title	PaN Research Project

**Description**

Example for an PaN research project to demonstrate the possible integration of various services, portals and data products to generate an all-embracing project documentation.

[Edit](#) [Delete Project](#)

### Members

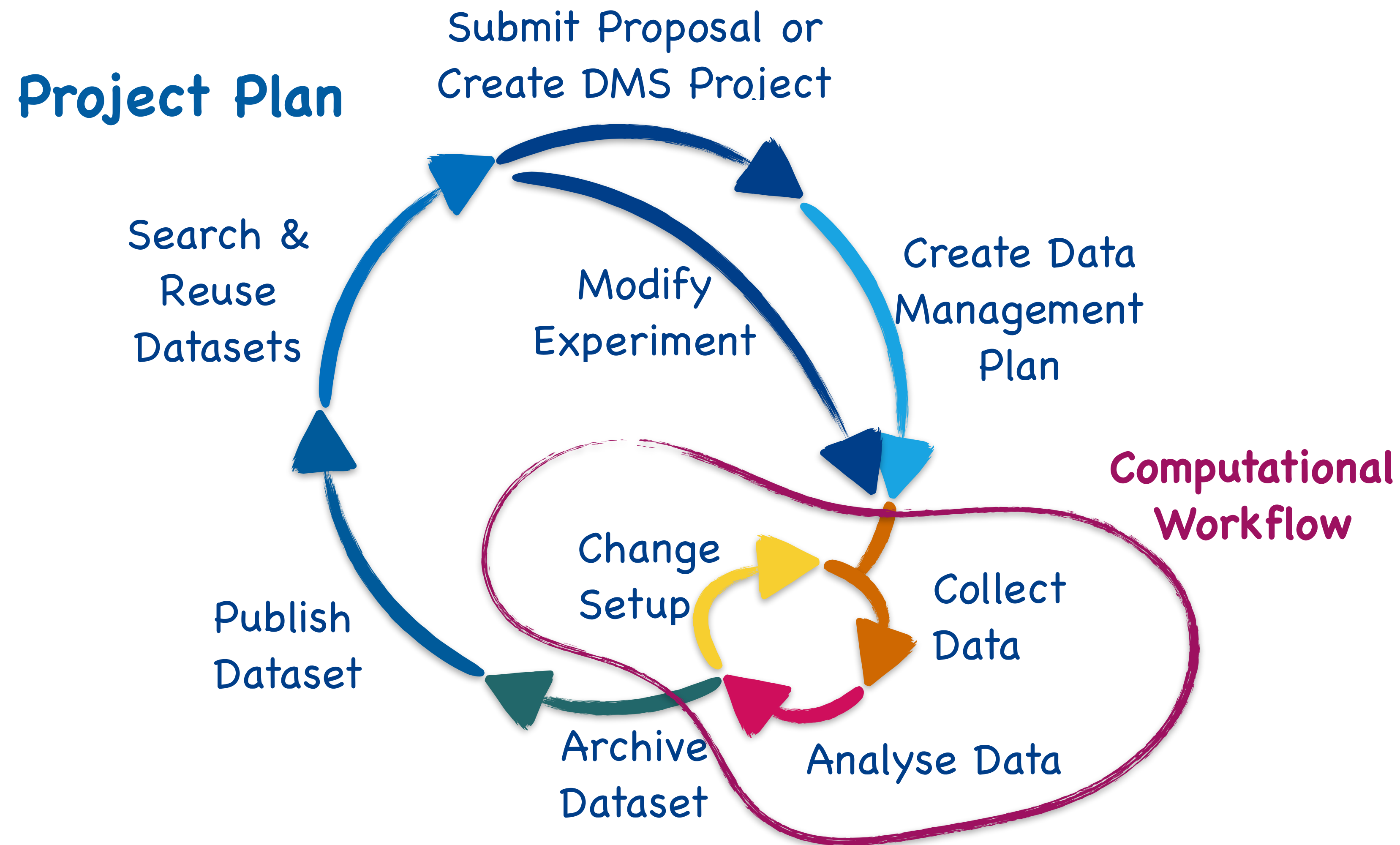
There are no members for this project at the moment!

[Add Member](#)

FWCC [Add all Members from Group](#)

# Scientific Workflow

HELIPORT has an build-in Integration of Scientific Workflows



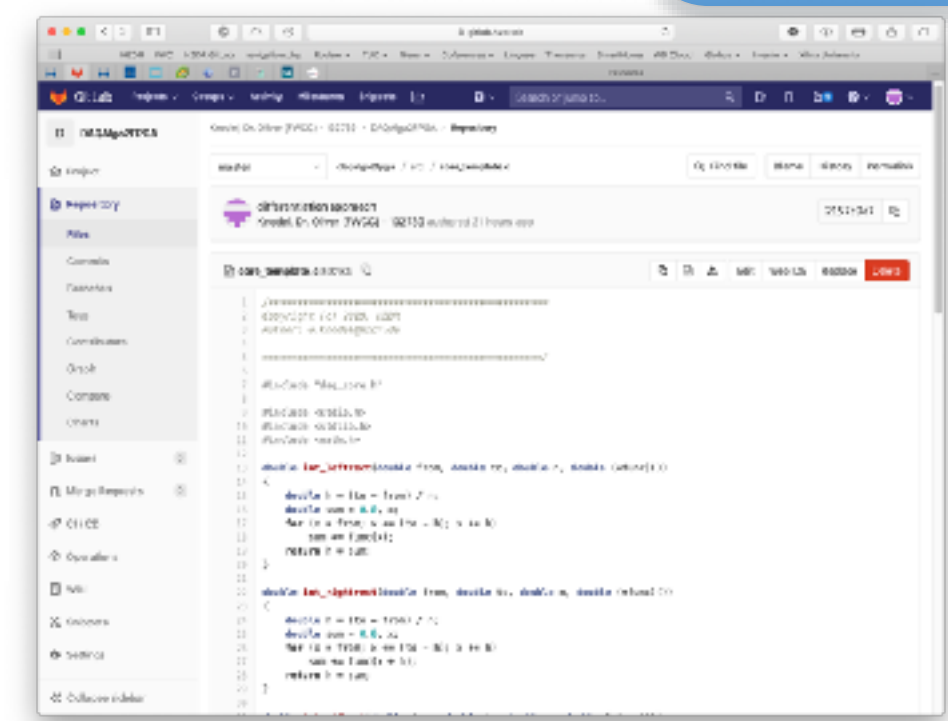
# Scientific Software Development and Reproducible Workflows

ID	Name	Cluster Login	Directory on Cluster	Status
46	cat echo	hensera	~/heliport_jobs	✓
44	echo cat sleep	Choose a Login	~/heliport_jobs	✓
44	echo cat sleep	hensera	~/heliport_jobs	✓
51	one bad disc per work	Choose a Login	~/heliport_jobs	✗
51	one bad disc per work	hensera	~/heliport_jobs	✗
41	sleep 5 seconds	Choose a Login	~/heliport_jobs	⚠
41	sleep 5 seconds	hensera	~/heliport_jobs	⚠

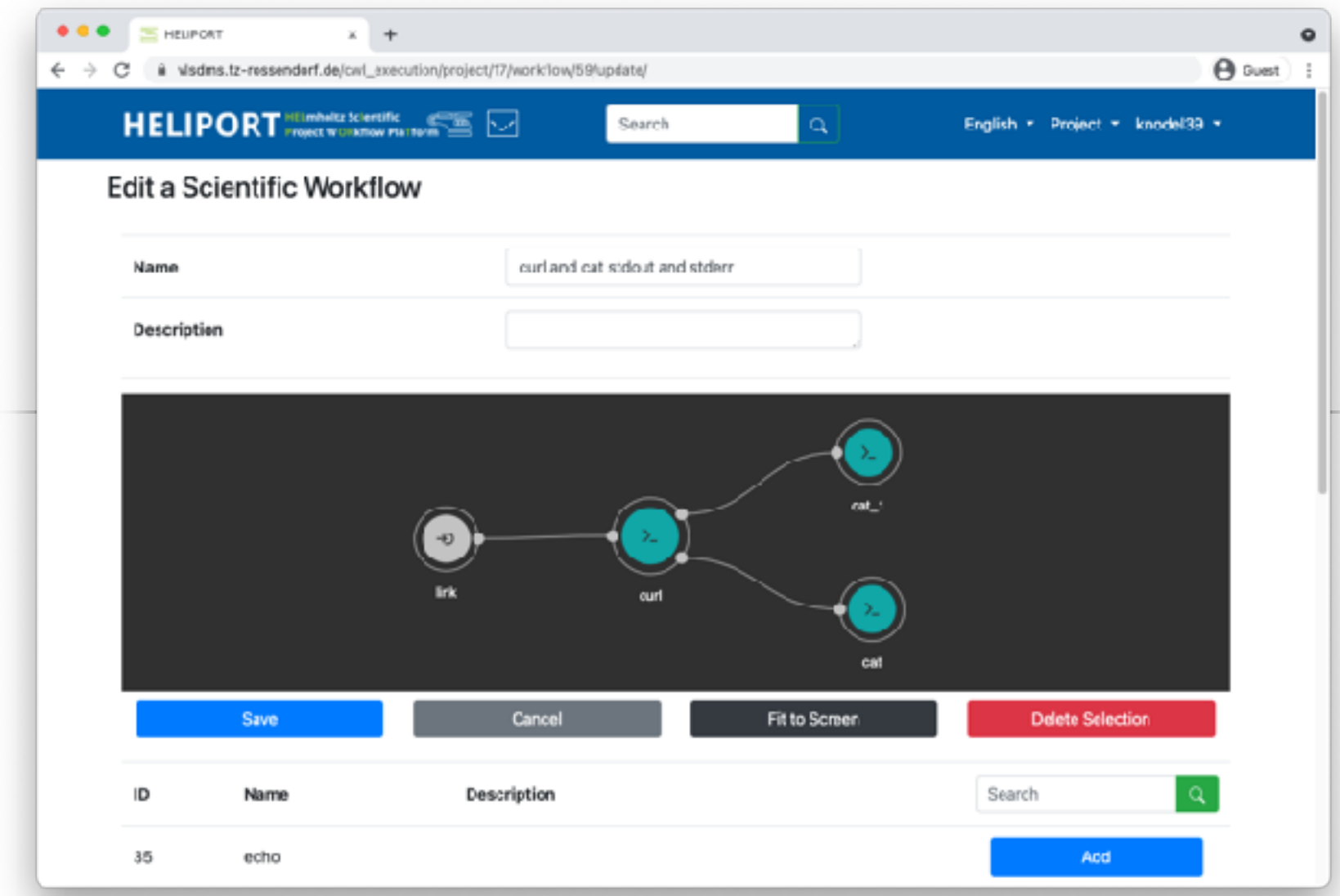
Workflow Engine

Version Control (GitLab)

Compute (HPC/OpenStack)

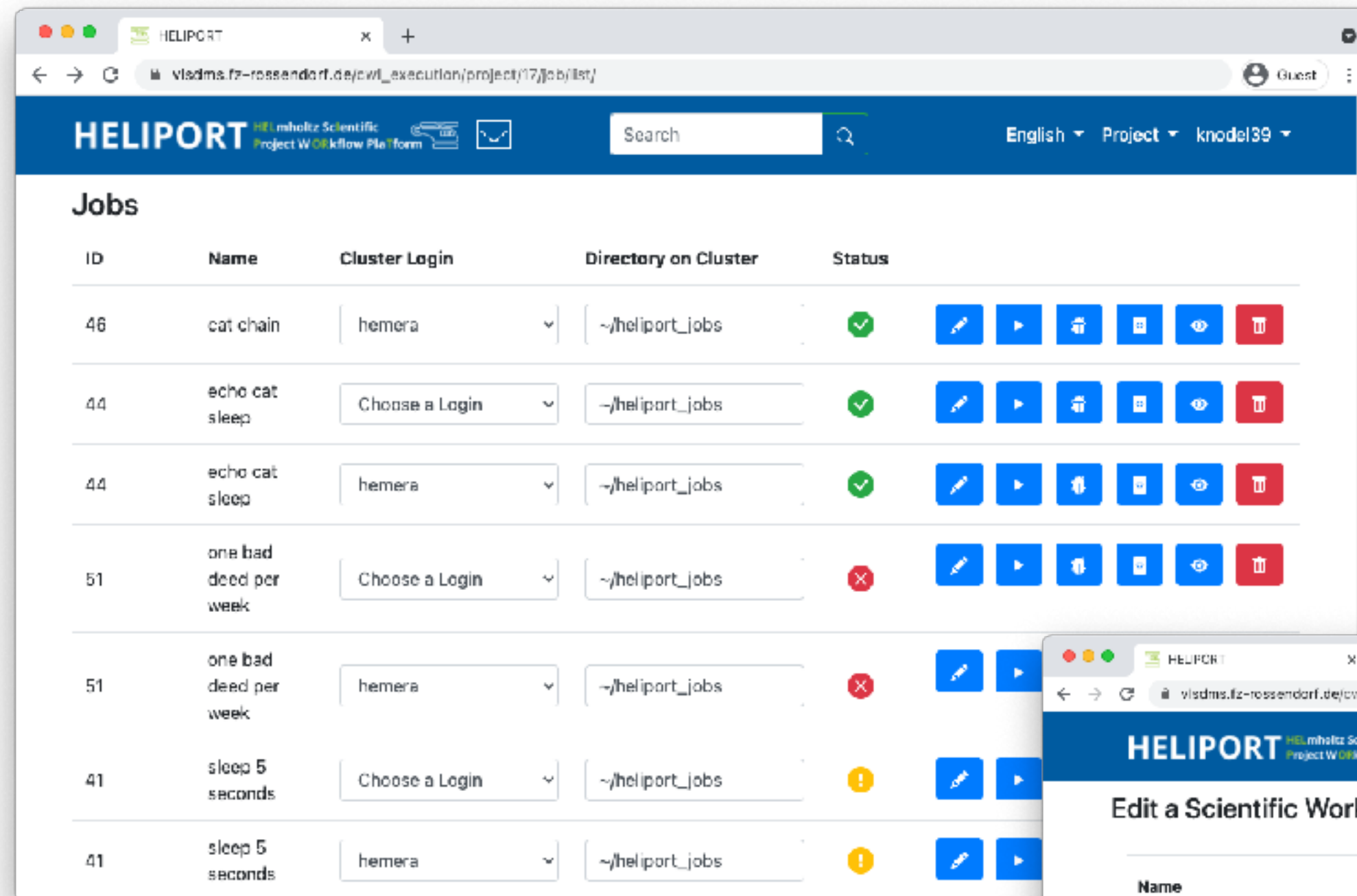


- Analysis and Pre-/Postprocessing steps needs to be:
  - Documented and
  - Reproducible
- Capsuling every step in a workflow adapts the **FAIR** principles.





























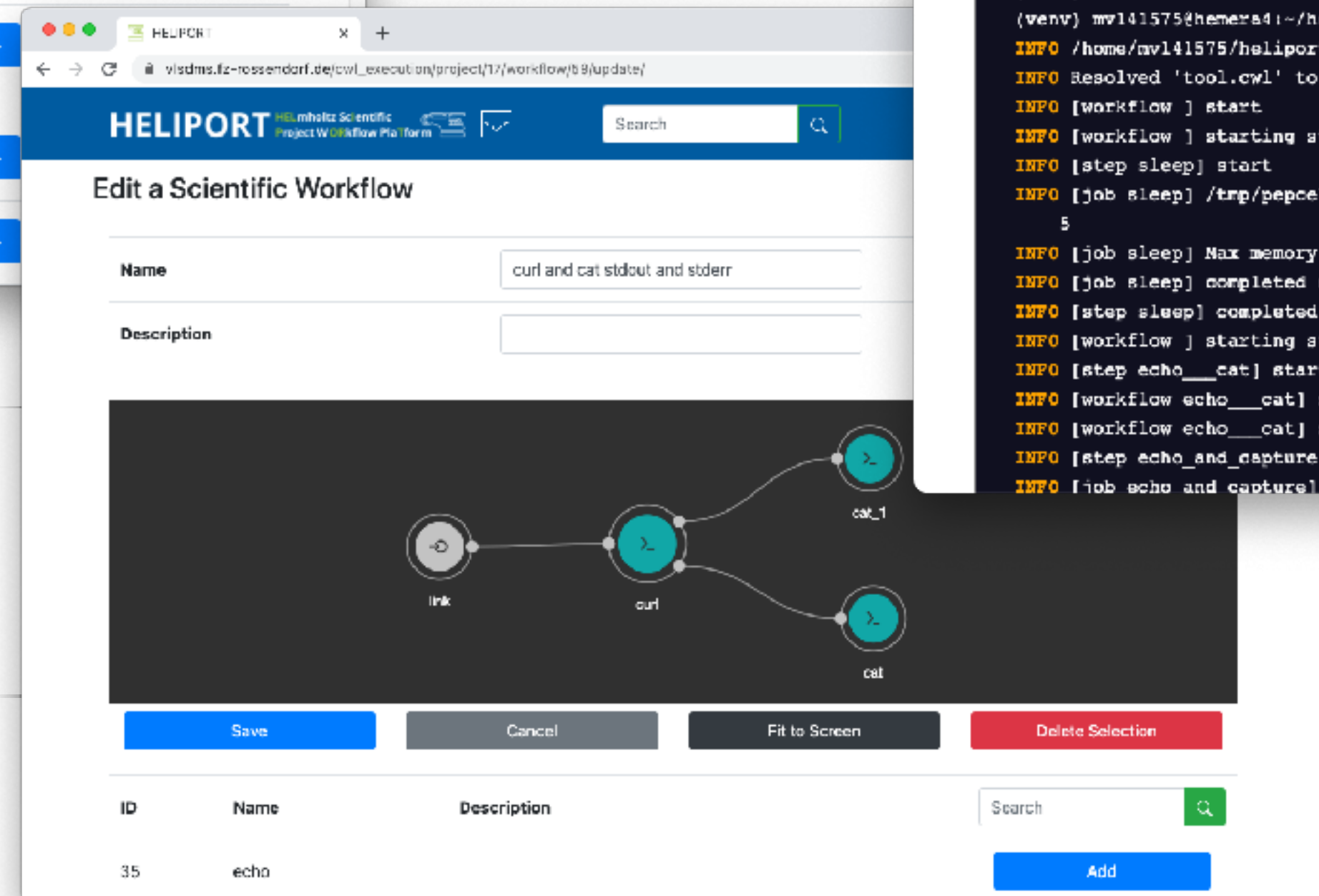


# Heliport Tracks Workflows, their Execution and Metadata



HELIPORT Jobs

ID	Name	Cluster Login	Directory on Cluster	Status	
46	cat chain	hemera	~/heliport_jobs	✓	    
44	echo cat sleep	Choose a Login	~/heliport_jobs	✓	    
44	echo cat sleep	hemera	~/heliport_jobs	✓	    
51	one bad deed per week	Choose a Login	~/heliport_jobs	✗	    
51	one bad deed per week	hemera	~/heliport_jobs	✗	 
41	sleep 5 seconds	Choose a Login	~/heliport_jobs	!	 
41	sleep 5 seconds	hemera	~/heliport_jobs	!	 



HELIPORT Edit a Scientific Workflow

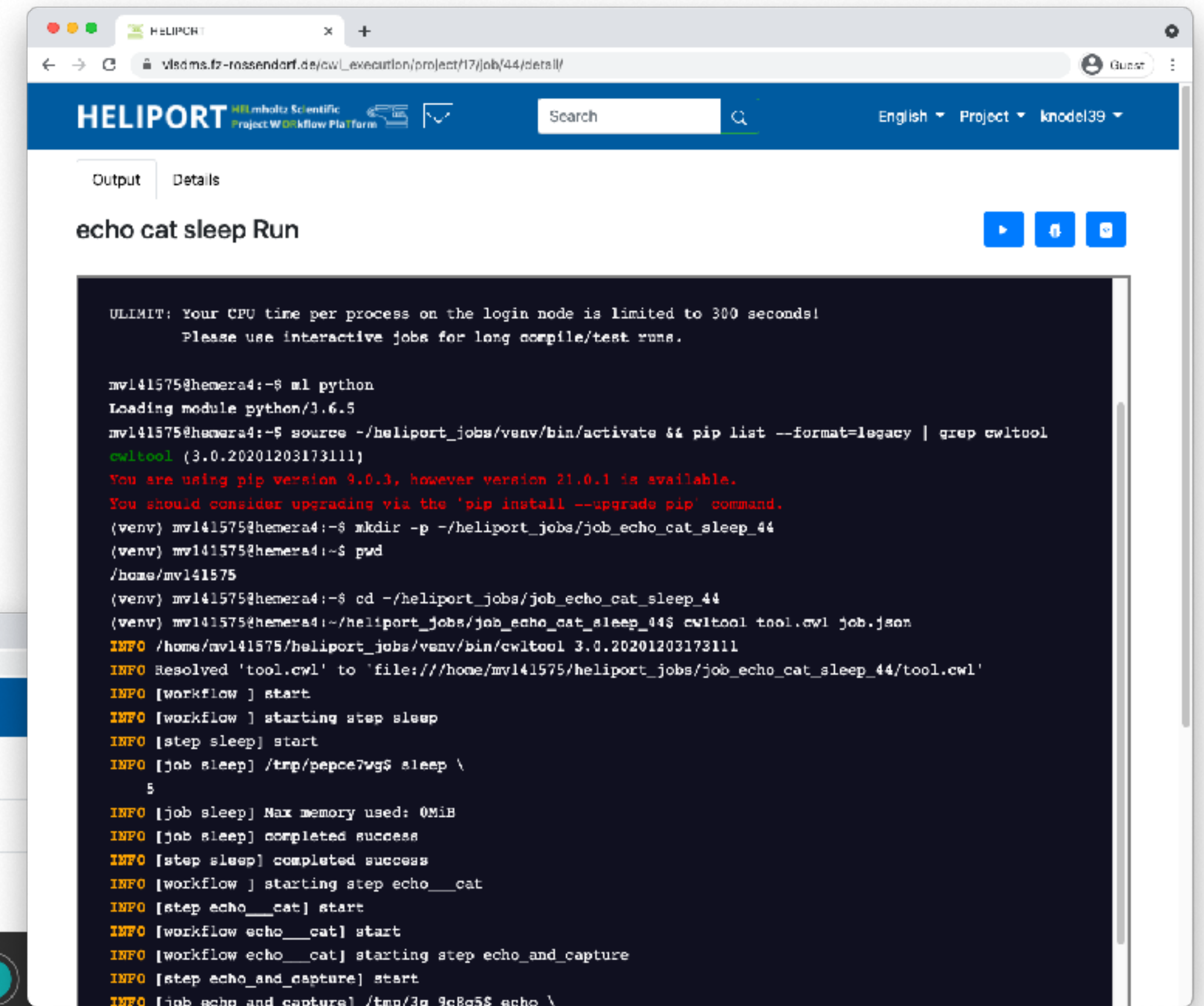
Name: curl and cat stdout and stderr

Description:

```
graph LR;
  In[ ] --> curl[ ];
  curl --> cat1[cat_1];
  curl --> cat[cat];
```

Buttons: Save, Cancel, Fit to Screen, Delete Selection

ID	Name	Description
35	echo	



HELIPORT Output Details

### echo cat sleep Run

```
ULIMIT: Your CPU time per process on the login node is limited to 300 seconds!
Please use interactive jobs for long compile/test runs.

mv141575@hemera4:~$ ml python
Loading module python/3.6.5
mv141575@hemera4:~$ source ~/heliport_jobs/venv/bin/activate && pip list --format=legacy | grep cwltool
cwltool (3.0.20201203173111)
You are using pip version 9.0.3, however version 21.0.1 is available.
You should consider upgrading via the 'pip install --upgrade pip' command.
(venv) mv141575@hemera4:~$ mkdir -p ~/heliport_jobs/job_echo_cat_sleep_44
(venv) mv141575@hemera4:~$ pwd
/home/mv141575
(venv) mv141575@hemera4:~$ cd ~/heliport_jobs/job_echo_cat_sleep_44
(venv) mv141575@hemera4:~/heliport_jobs/job_echo_cat_sleep_44$ cwltool tool.cwl job.json
INFO /home/mv141575/heliport_jobs/venv/bin/cwltool 3.0.20201203173111
INFO Resolved 'tool.cwl' to 'file:///home/mv141575/heliport_jobs/job_echo_cat_sleep_44/tool.cwl'
INFO [workflow ] start
INFO [workflow ] starting step sleep
INFO [step sleep] start
INFO [job sleep] /tmp/pepce7vg$ sleep \
5
INFO [job sleep] Max memory used: 0MiB
INFO [job sleep] completed success
INFO [step sleep] completed success
INFO [workflow ] starting step echo__cat
INFO [step echo__cat] start
INFO [workflow echo__cat] start
INFO [workflow echo__cat] starting step echo_and_capture
INFO [step echo_and_capture] start
INFO [job echo and capture] /tmp/3a_9c8g5$ echo \
```

# Heliport REST API

- The API provides access to our full Heliport infrastructure:
  - Proposal access (GATE),
  - Handle management,
  - CWL execution and monitoring,
  - Project metadata export,
  - Digital Object and
  - Lifecycle management.
- API documentation (ReDOC) available.

The screenshot displays the Heliport REST API documentation interface. The browser address bar shows the URL: `vlsdms.fz-rossendorf.de/redoc/#operation/createDigitalObject`. The interface is divided into several sections:

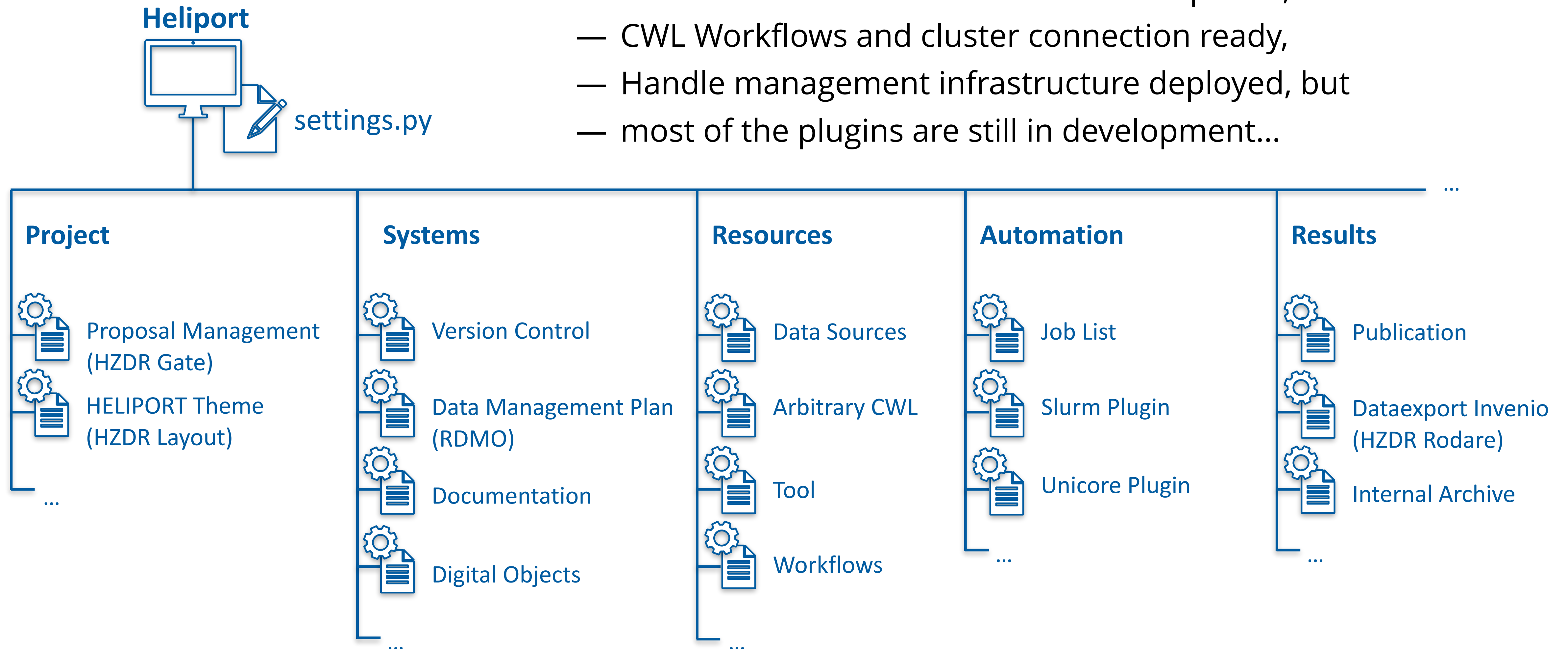
- Left Sidebar:** A search bar and a list of API endpoints. The 'digital-objects' section is expanded, showing endpoints like `listDigitalObjects`, `createDigitalObject` (highlighted), `retrieveDigitalObject`, `updateDigitalObject`, `partialUpdateDigitalObject`, and `destroyDigitalObject`.
- Main Content Area:**
  - Endpoint:** `createDigitalObject` (POST) under the 'Digital Objects' group.
  - Request Body Schema:** `application/json`. The schema includes:
    - `project`: integer, required.
    - `handle`: string, `<= 100 characters`, Nullable.
    - `relation`: string, required.
    - `category`: string, required.
    - `description`: string, required.
  - Responses:** A section for response codes, currently showing `> 201`.
- Right Panel:** A dark-themed sidebar with two sections:
  - Request samples:** A 'Payload' section showing a JSON example:

```
{  "project": 0,  "handle": "string",  "relation": "string",  "category": "string",  "description": "string"}
```
  - Response samples:** A '201' section showing a JSON example:

```
{  "digital_object_id": 0,  "project": 0,  "handle": "string",  "relation": "string",  "category": "string",  "description": "string"}
```

# Modular Heliport Design (Django Apps)

- Initial structure and infrastructure completed,
- CWL Workflows and cluster connection ready,
- Handle management infrastructure deployed, but
- most of the plugins are still in development...



# Heliport (Project) Roadmap

## First Draft: Project Plan (August 2020)

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

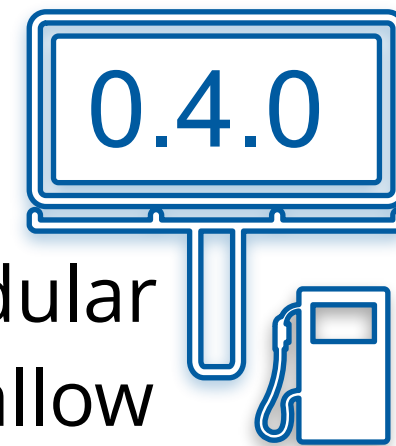
## Modular Structure (July 2021)

- Subdivision of the stages into modular and configurable Django apps to allow individual extensions
- Refactoring of the project
- Official start of the HMC founded Heliport project:



## Documentation of a TELBE user experiment

- Integration of all related data sources
- Automated workflow initiation
- Publication of all data products



## 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)

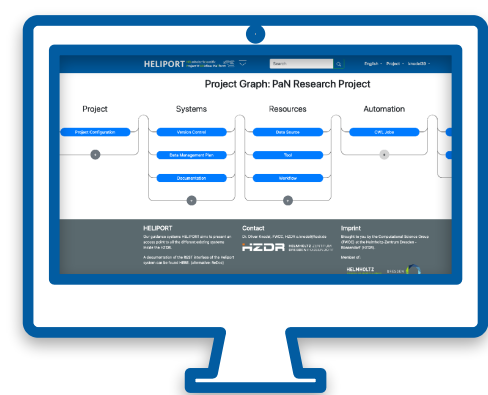
- Fully configurable stages and modules
- Infrastructure and database updates
- Daily proposal database update
- CI pipeline for test and deployment
- Advanced logging and monitoring

## Integration of various Apps and Features

- Export for (different) Metadata Schemas
- Computational/Scientific workflow execution
  - Workflow management and monitoring
  - CWL support
- Documentation using GitLab pages
- (Global) **Handle management**
- Extended **Support for Digital Twins**
- Data Management Plan Export for RDMO

# HELIPORT

HELMholtz Scientific  
Project WORKflow PlaTform



Heliport Prototype  
(Only available within  
the HZDR network)



Heliport source code v0.4.0  
available on Rodare

DOI 10.14278/rodare.947

<https://vlsdms.fz-rossendorf.de/>

<https://rodare.hzdr.de/record/947>