

RIC and open science for FS

Research and innovation in scientific computing in central IT

Sophie Servan, Patrick Fuhrmann, Tim Wetzel, Paul Millar
Hamburg, 2023-11-10

Who is here today from RIC?



Patrick Fuhrmann
Head of RIC
Data management
Cloud computing



Philipp Heuser
Imaging
Machine learning



Paul Millar
Data management
FAIR data



Sophie Servan
Patrick's deputy
FAIR data



Tim Wetzel
Cloud computing
AAI
HPC

What does RIC do for FS?

Snapshot of the current actions

01 FAIR data management

- PaNET: techniques ontology
- NeXus
- Direct DOI data access
- Federated search API
- **VISA**

02 Image processing and ML

- Helmholtz Imaging
- Helmholtz-wide foundation model

03 Federated infrastructures

- **HIFIS** e.g. helmholtz.cloud
- Towards "Single Sign On" for federated Photon services
- EOSC German Node

04 DESY community

- PaNOSC @DESY
- NeXus @DESY
- **SciCat** @DESY
- Deep Learning @DESY

05 European Community

- Involved in the VISA and FAIR data management MoUs
- Leading LEAPS working group on IT
- Representing the PaN cluster in **OSCARS** with ESRF
- Technical board of DAPHNE4NFDI and EOSC-Future
- Representing DESY in the EOSC-Association
- Representing the NGI-DE

OSCARS

The next big project for PaN in EOSC

- After ExPaNDS, PaNOSC, EOSC-Future and in parallel to DAPHNE4NFDI
- Starting 01.01.2024 for 4 years and 25 M€ from the EU Commission
- **All sciences on the EU level**: 5 clusters
- **DESY** and **ESRF** for the PaN community
- Expected impacts:
 - increase **data re-use** in PaN community
 - increase adoption of VISA for open data analysis
 - promote PaN **standards**
 - provide a framework for managing data across facilities based on the "data lake" model from ESCAPE - e.g. outsourcing and disaster recovery
 - create a pool of **data stewards**
- **Open calls** to the whole community:
 - 100-250k€ for 1 to 2 years on a **focused** project towards FAIR and open science
 - 100% funded, lump sum, light-weight reporting
 - Total budget for open calls: **16.4 M€** for 2024-2025
 - First call in **March 2024**
 - Impact on multiple RIs - better on multiple sciences

FAIR data and open science at DESY

Are stars aligned?

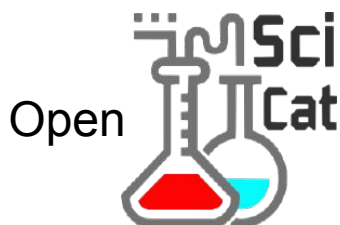
- In the PaN community, push towards FAIR and open and we're a bit behind
 - ExPaNDS raised the awareness at DESY
 - one of its main objectives: **metadata catalogue** linked to federated search between PaN facilities - final result: 12 out of 17 PaN facilities (not us)
 - DAPHNE4NFDI endorsed the choice of **SciCat** for DESY and picked up the process
- Metadata catalogue: Peter's group is now rapidly gaining expertise in deploying / running SciCat (cf. Jürgen's talk)
- Public access to data:
 - PETRA and FLASH data management not yet optimal for open data access after embargo and for VISA
 - **dCache Infinite Space** is now a HIFIS service, accessible via Helmholtz ID
- Open science: a prototype of **VISA** was developed during ExPaNDS
- Step-by-step approach:
 - Combine our know-how and publicly accessible tooling to enable beta testers to store and analyse open data with dCache, SciCat and VISA
 - Can be linked to European PaN federated search (EOSC)

FAIR data and open science at DESY

Composing existing components into a future repository service



HELMHOLTZ
ID



InFa

- Deposit open datasets
- Possible DOI minting
- Curate dataset metadata
- Access open datasets



RIC

- Visualise HDF5 datasets
- Run analysis tools on open datasets
- Publish analysis tools



Infinite Space by



SciComp & RIC

HIFIS

- HIFIS storage for open datasets

**Thank you.
Questions?**

6TH ROUND TABLE ON DEEP LEARNING AT DESY

08.12.2023, 10H-15:30H
FLASH SEMINAR ROOM

By now the use and development of deep learning methods found its way in all areas of research and all institutes and labs on DESY campus. For the 6th time we invite you for a meeting dedicated to the exchange of ideas and innovations across the campus.

The idea behind this meeting is to bring together the people interested in and working on machine learning, deep learning and artificial intelligence methods on DESY Campus.

CALL FOR CONTRIBUTIONS:
YOU ARE INVITED TO PRESENT HOW YOU ARE USING DEEP LEARNING TO PUSH YOUR RESEARCH DOMAIN FORWARD IN A SHORT FLASH TALK. PLEASE SUBMIT A SHORT ABSTRACT (JUST ONE SHORT PARAGRAPH) IN INDICO BY 26TH NOV 2023

Please register at:
<https://indico.desy.de/event/42212/>



Contact:
philipp.heuser@desy.de
engin.eren@desy.de