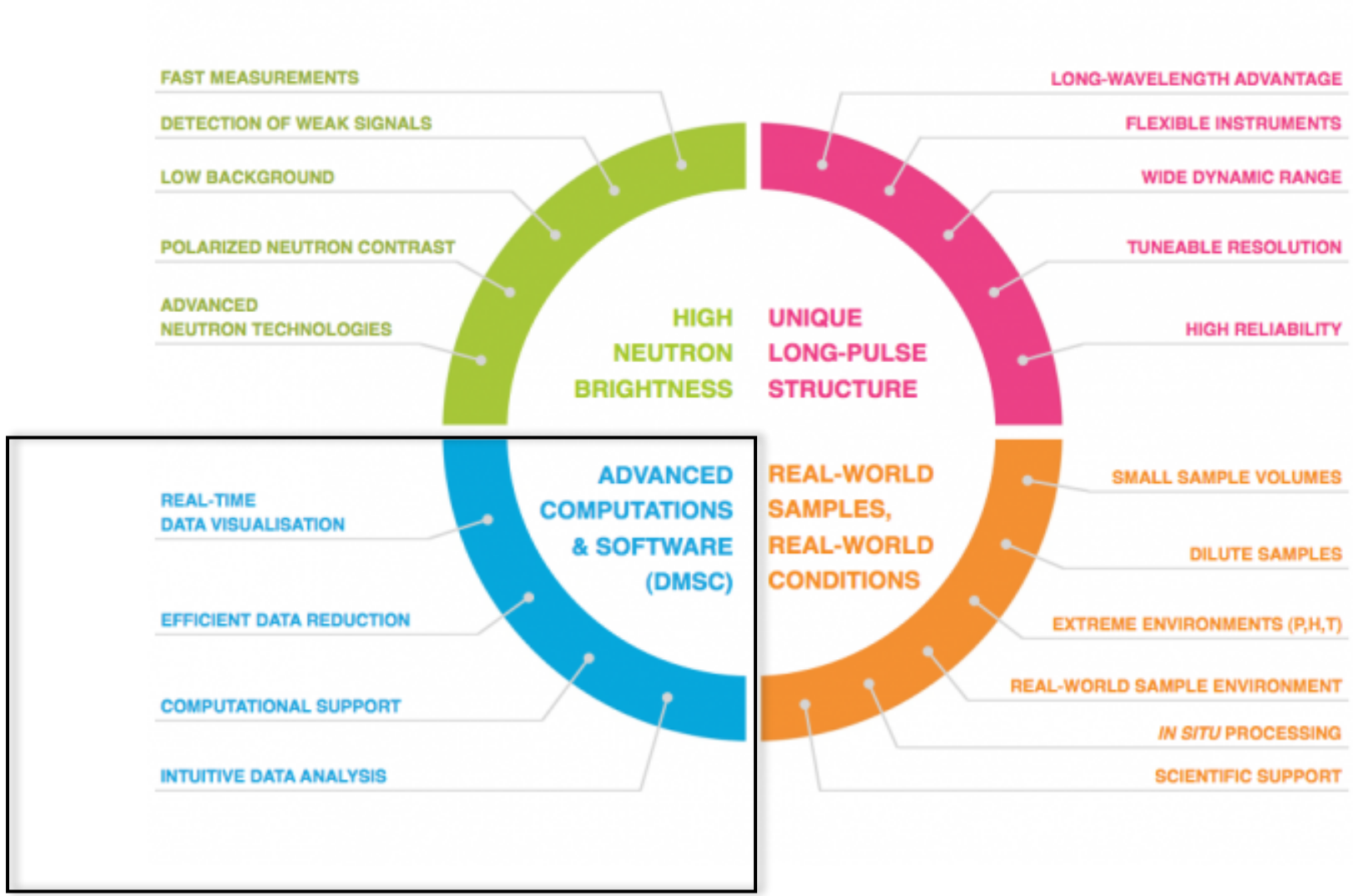
The ESS Data Management and Software Centre

Getting the most out of data

FREDRIK BOLMSTEN, GROUP LEADER SCIENTIFIC INFORMATION MANAGEMENT SYSTEMS

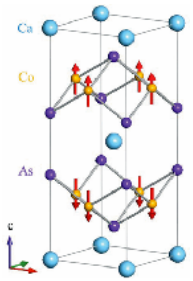
Technical Design Report (2013)

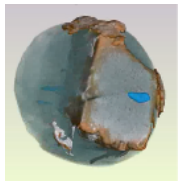
Importance of computing was emphasized already in design phase

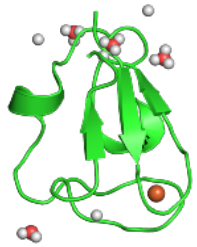


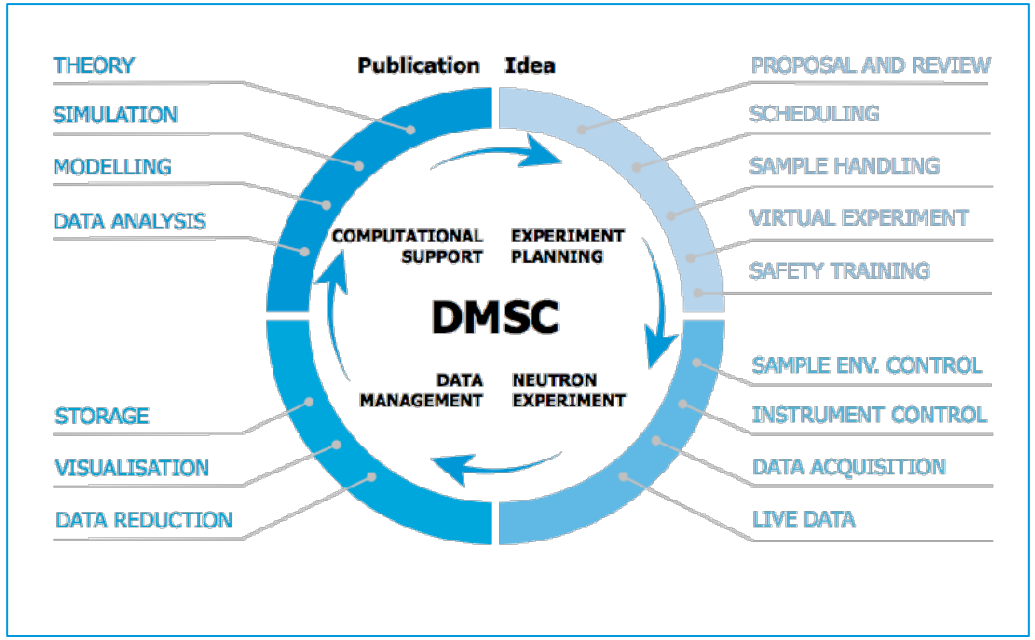
User journey

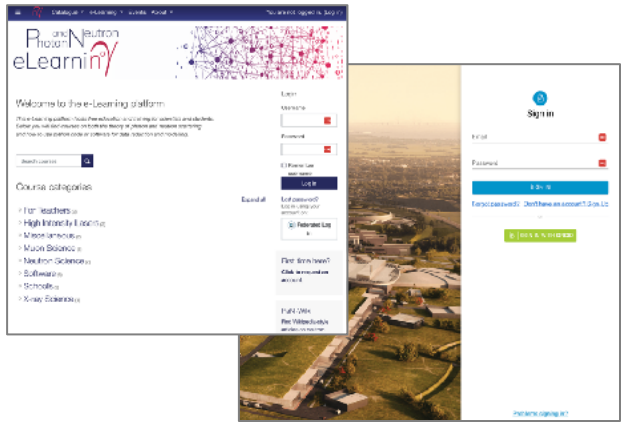
Support user from idea to publication with scientific computing tools & services

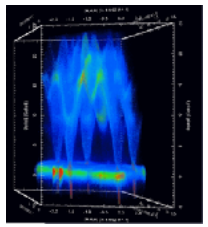


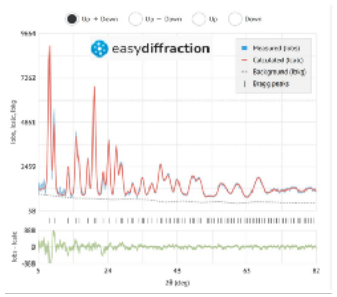


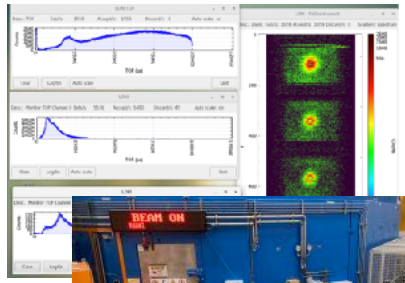


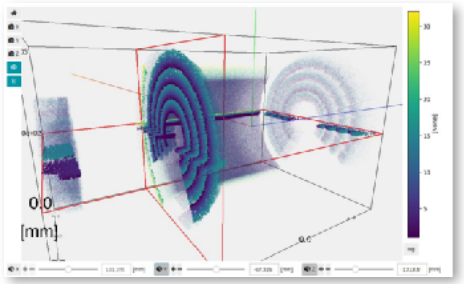


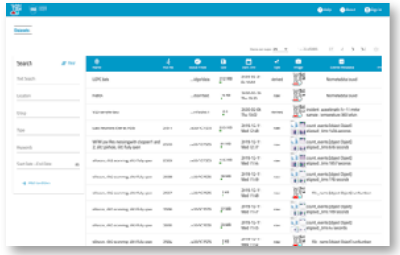


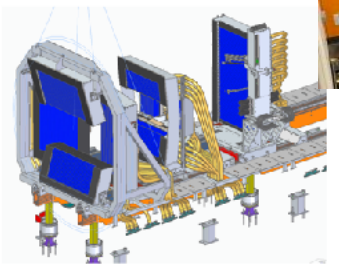






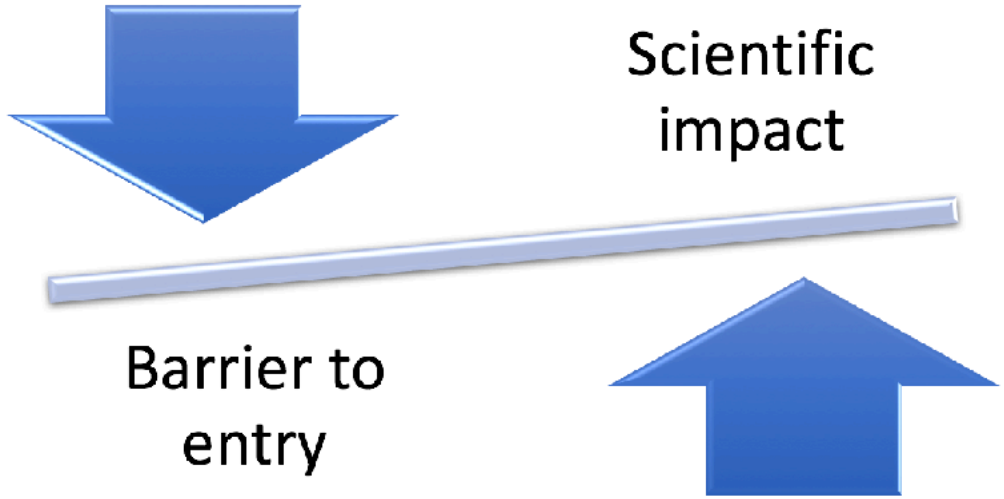






DMSC objective

Minimize the time it takes to analyze and interpret experimental data



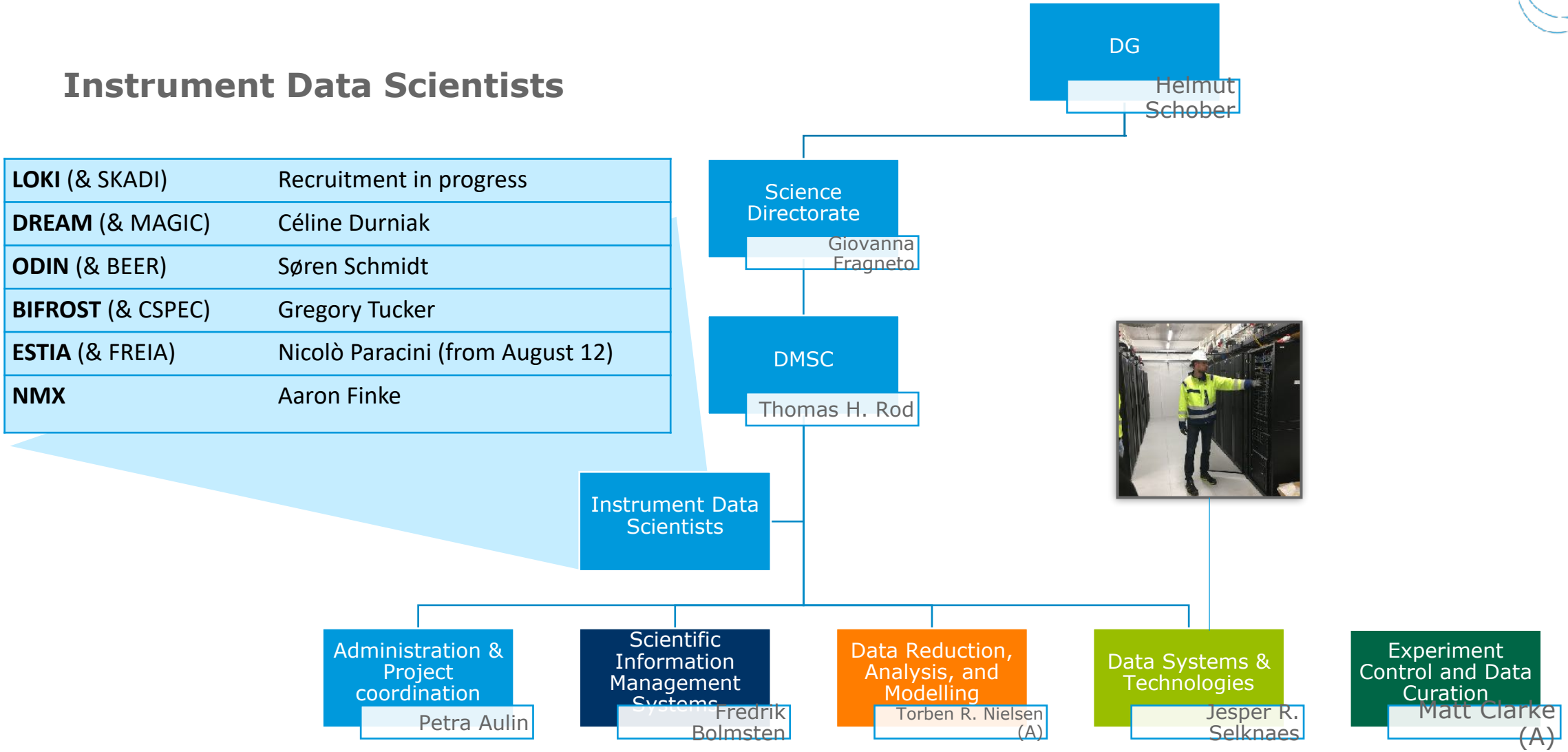
This is particular important for neutron sources due to the cost of producing neutrons

Maximise the scientific impact and success of ESS by serving the needs of both non-expert and advanced users

DMSC organisation

Instrument Data Scientists

LOKI (& SKADI)	Recruitment in progress
DREAM (& MAGIC)	Céline Durniak
ODIN (& BEER)	Søren Schmidt
BIFROST (& CSPEC)	Gregory Tucker
ESTIA (& FREIA)	Nicolò Paracini (from August 12)
NMX	Aaron Finke



Integrated data pipeline

for each instrument

- Interface to instrument teams
- Customization to instruments

UOS | EPICS | NICOS | kafka | scipp | Technique specific | MSci Cat | NeXus | python

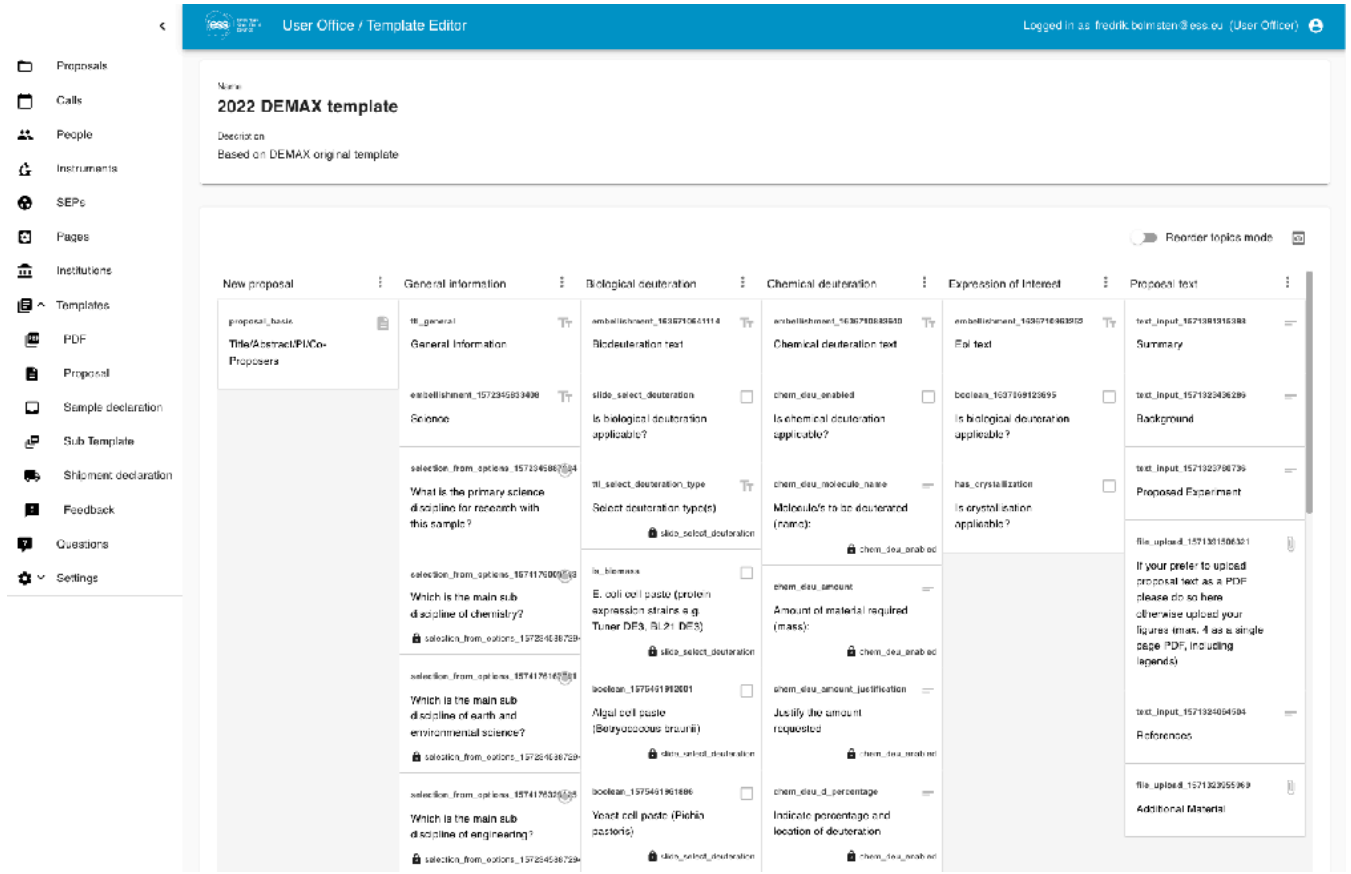
Modelling & simulations (planning) → User office software (proposal) → Experiment control → Stream events & meta data → Data reduction & visualization → Data analysis & modelling → FAIR Data management

Data Systems & Technologies:

VISA | jupyter

Remote Access

User Office



DYNAMIC

Virtually nothing is hard coded. Templating tools are simple to use by non-software developers giving flexibility to changes.

EUROPEAN SPALLATION SOURCE

Science and Technology Facilities Council

eli

Proposal Folder and File Structure

Proposed structure to be mounted in the same place on all relevant machines. The view of this structure should be consistent across machines regardless of location.

TODO: Permissions are for the embargo period. What is opened after, if subject to checks with the policy.

There will be one proposal folder per instrument per year.
The top level path for the proposal is set to:

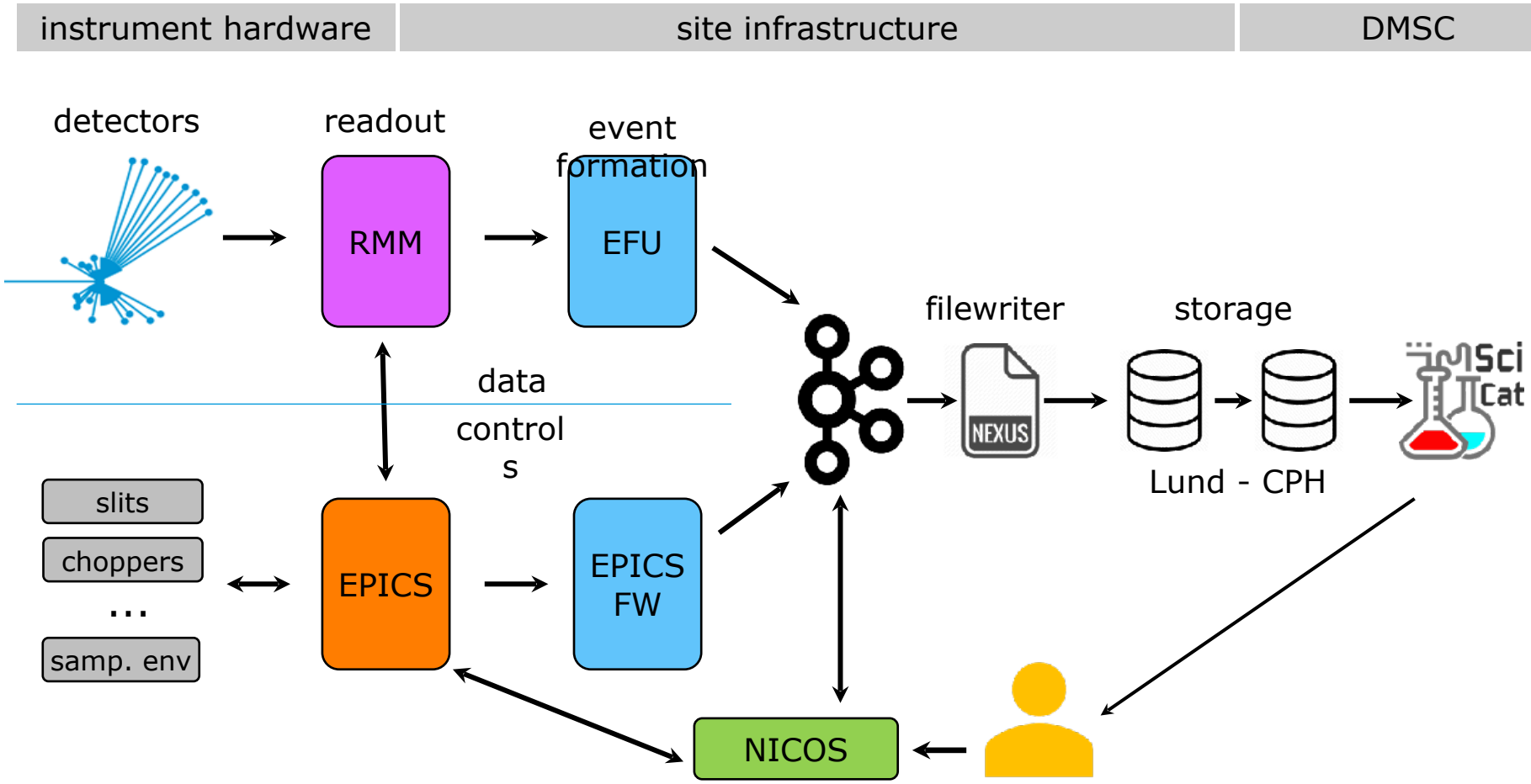
/ess/data/<instrument>/<year>/<proposal_id>

In this folder, the following sub folders will be created:

- **raw**
  This folder will contain the raw data files created by the file writer with the data collected during the experiment time.
  Files/Folder are only writable by the file writer (ECDC staff), read-only for the proposal team + instrument team.
- **reduced**
  This folder will contain the reduced data produced by the reduction pipeline specific for the instrument. The reduction pipeline should be running automatically, although the data instrument and

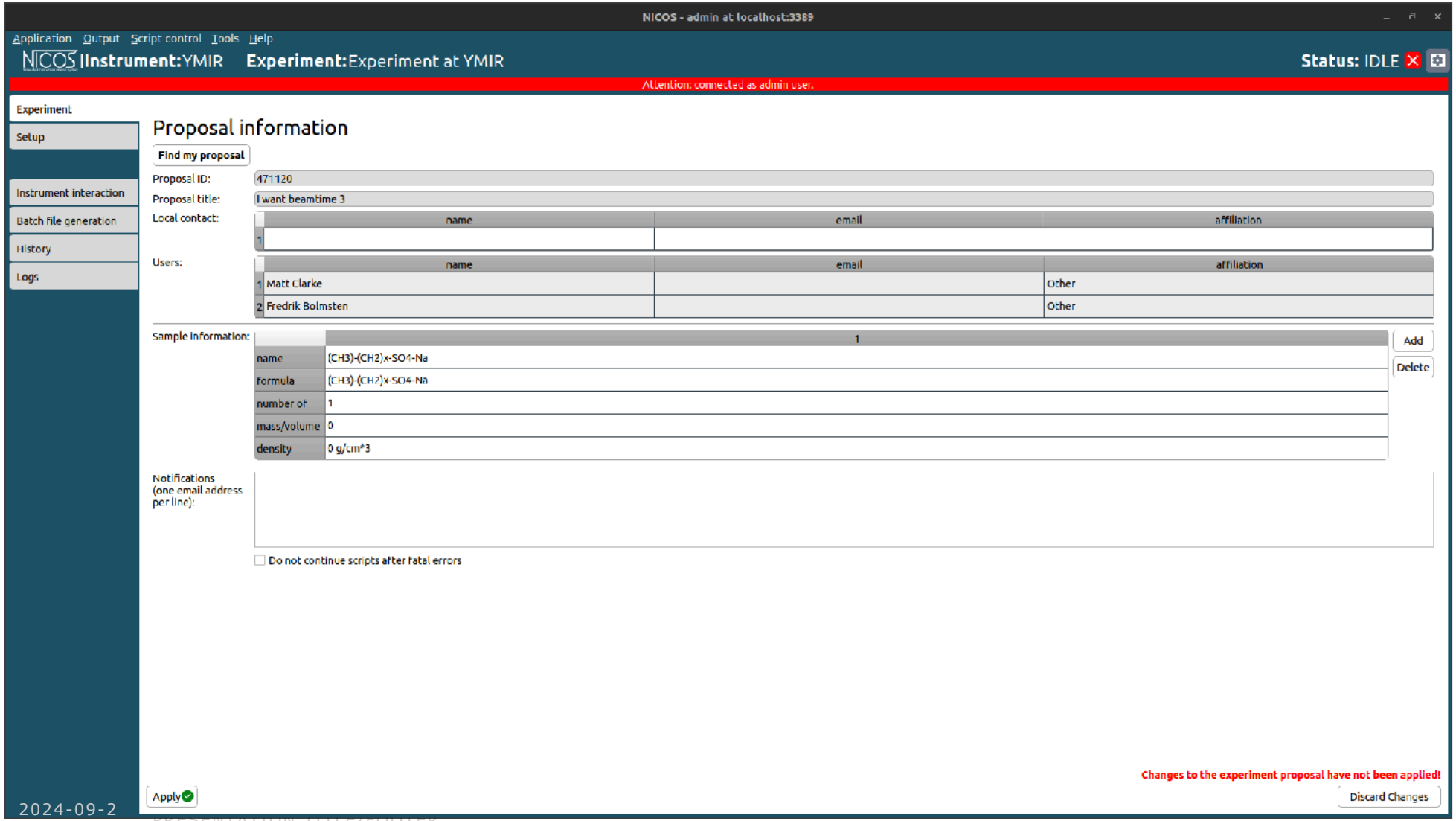
```
[root@sftpserver2 ~]# sudo -u nexus ls -l /sftproot/ess/data/ymir/2024/451867/
total 4
drwxr-x---     2 nexus ess_proposal_451867 4096 May  6 15:26 aux
dr-xr-x---     2 nexus ess_proposal_451867 4096 May  6 15:26 derived
drwx------ 65535 root  root               4096 May  6 15:26 ingestor
drwx------ 65535 root  root               4096 May  6 15:26 raw
dr-xr-x---     2 nexus ess_proposal_451867 4096 May  6 15:26 reduced
drwxr-x---     2 nexus ess_proposal_451867 4096 May  6 15:26 tmp
drwxr-x---     2 nexus ess_proposal_451867 4096 May  6 15:26 upload
```

Experiment Control



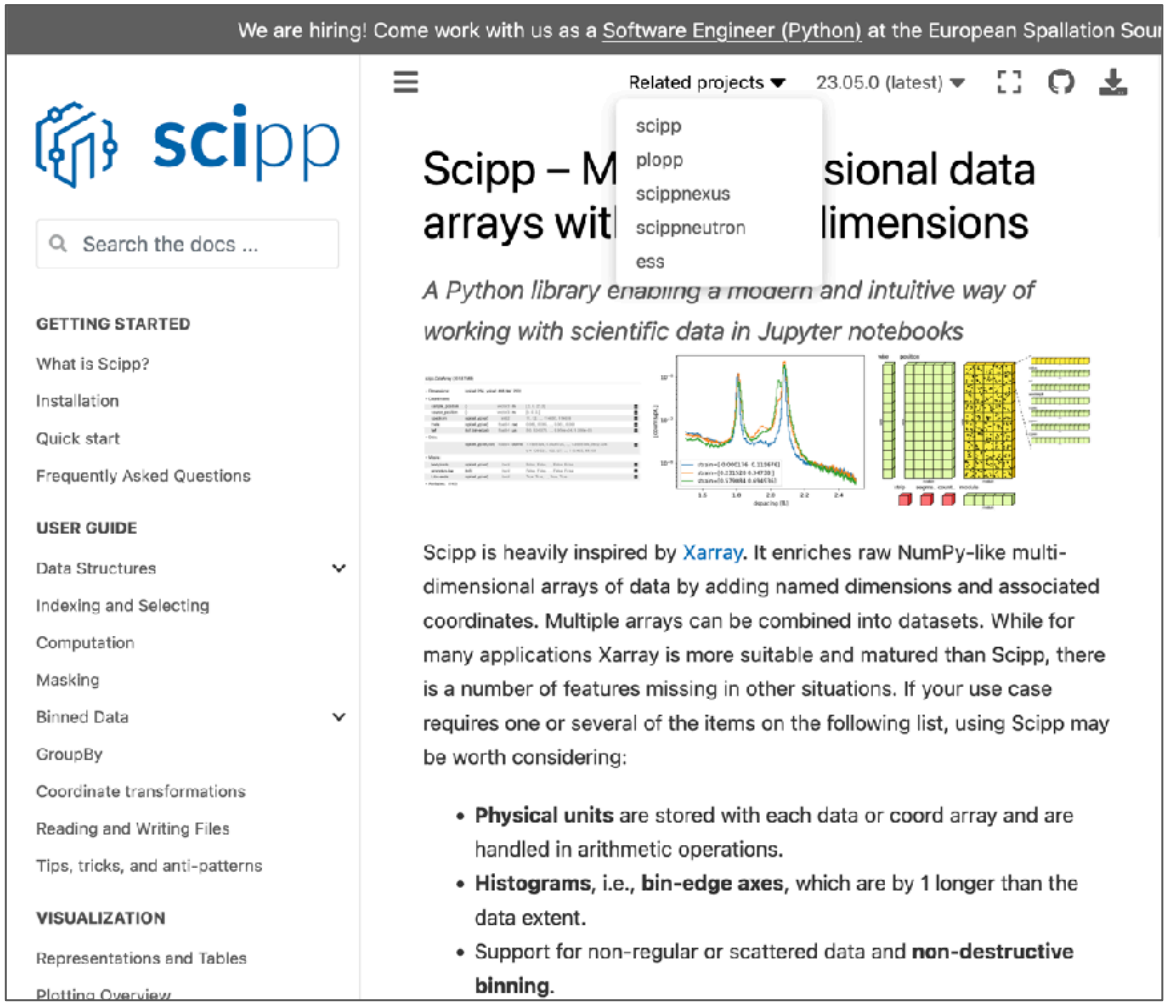
NICOS

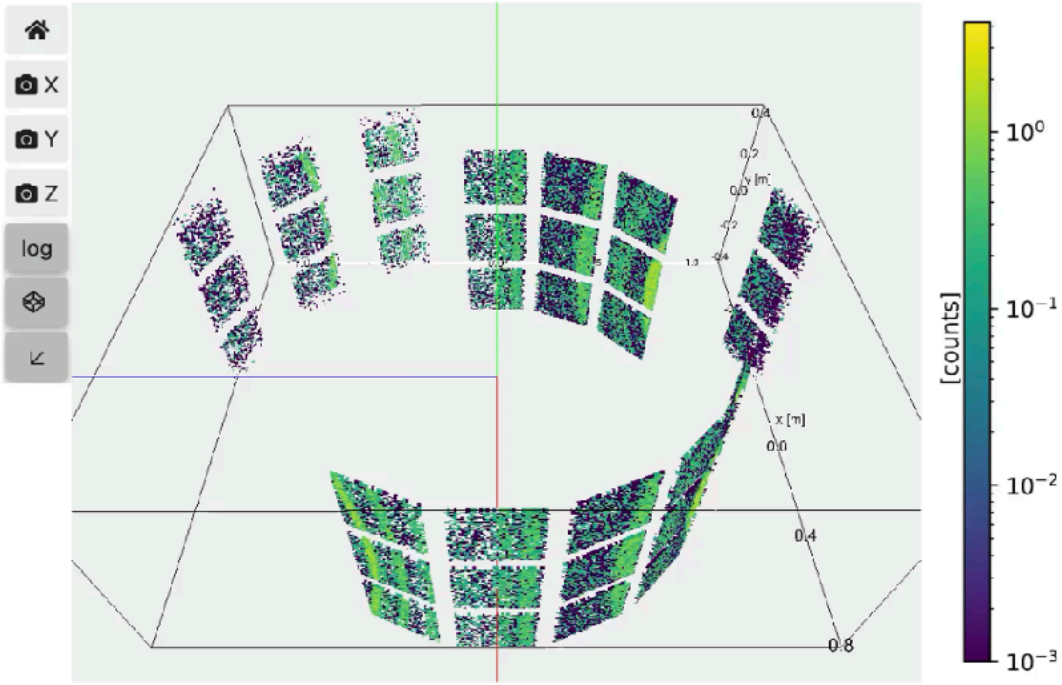
With User Office integration



Data reduction & visualization software

https://scipp.github.io

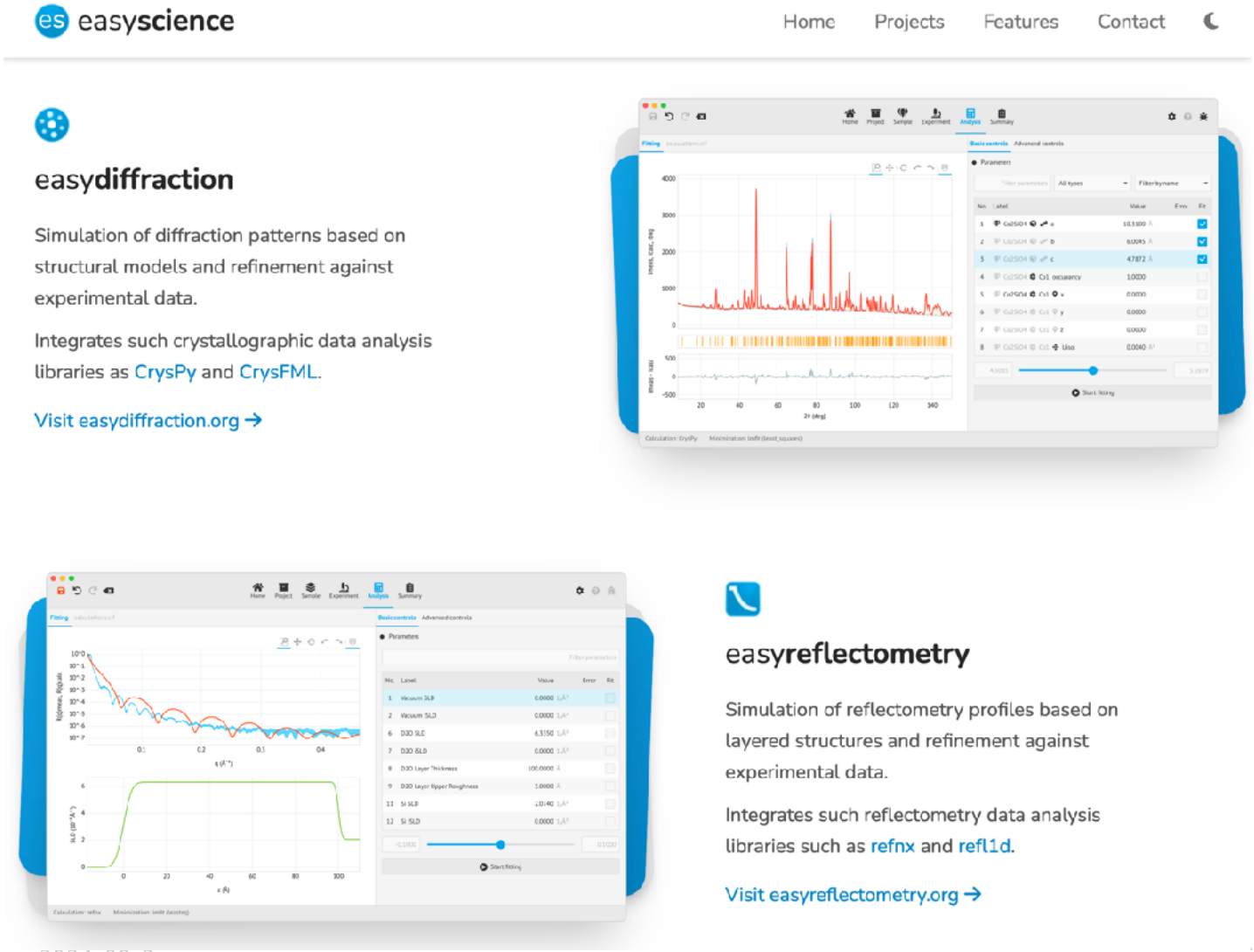




python™ jupyter

Data analysis software

https://easyscience.software



User interfaces:

- Python
- Jupyter
- GUI

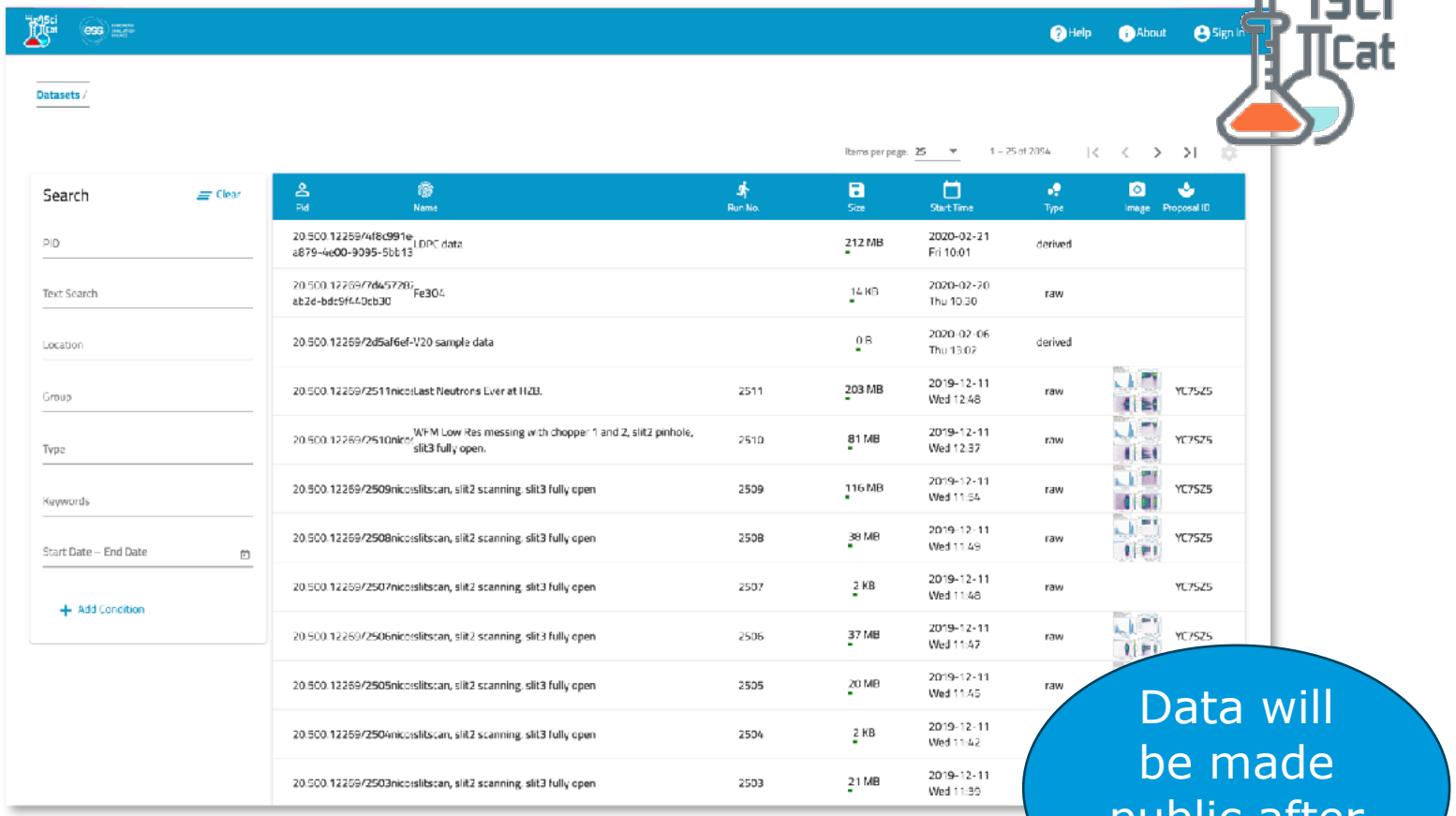
Multiple libraries (Calculators)

Other requests are:

- Data analysis for QENS
- Data analysis for TOF imaging

FAIR data management

SciCat is used as data catalogue by several facilities

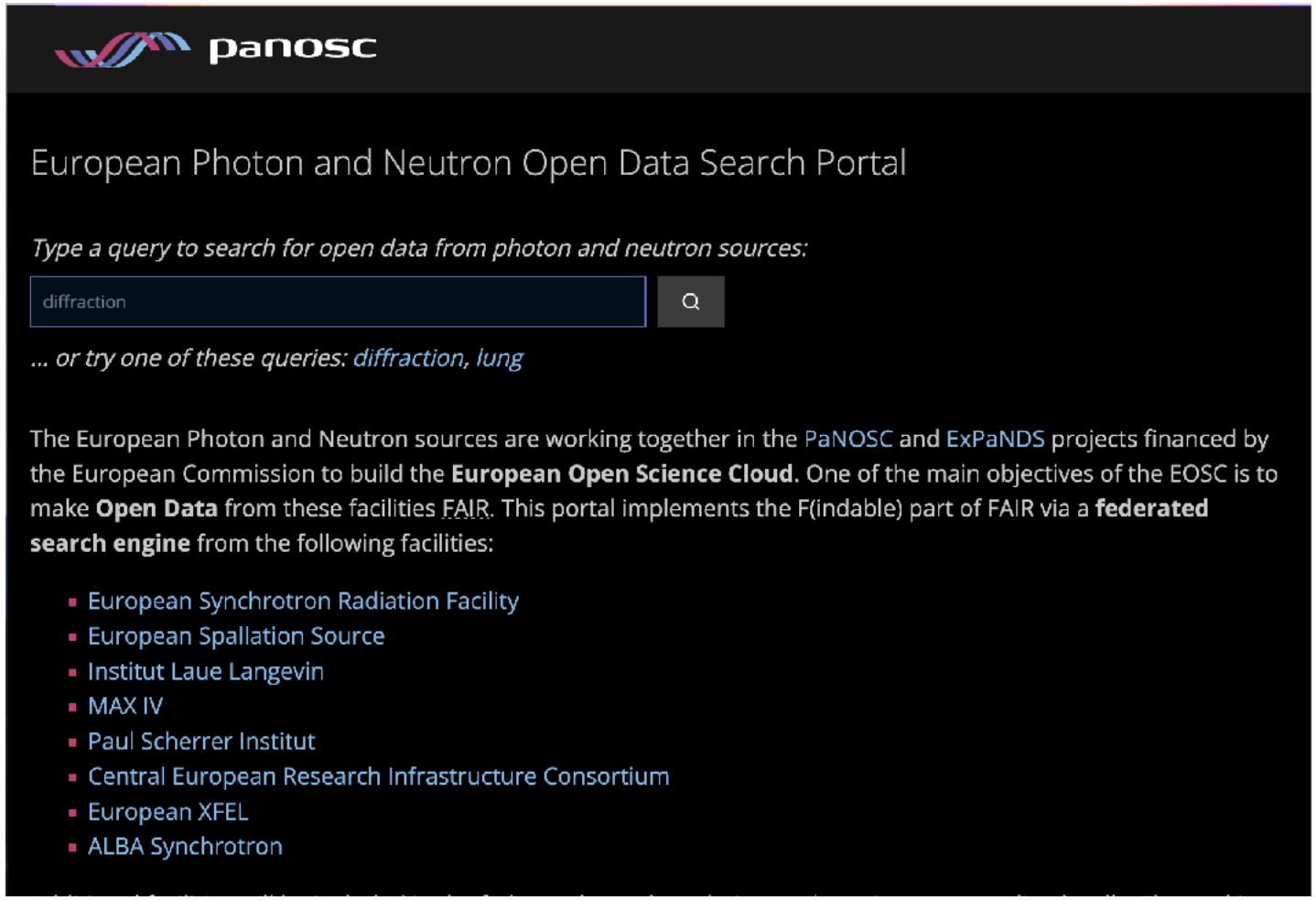


Data will be made public after 3 years

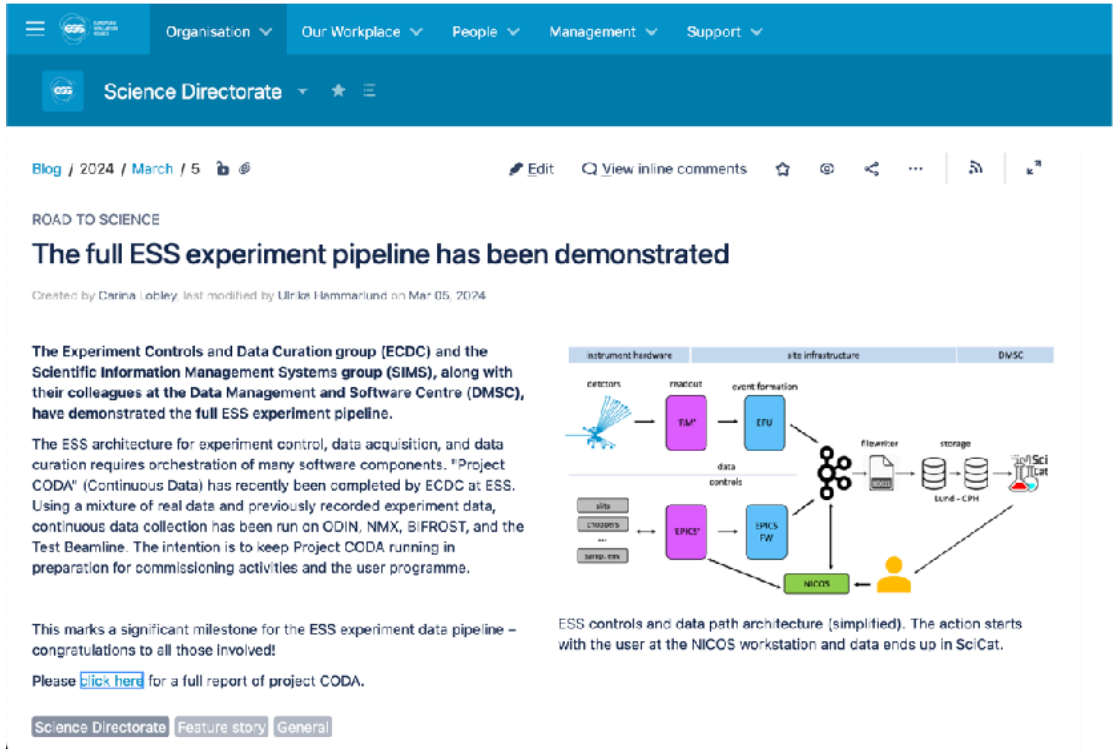
Federated data portal

You can find and access data from multiple photon and neutron sources from one place

https://data.panosc.eu



Current status



Organisation | Our Workplace | People | Management | Support

Science Directorate

Blog / 2024 / March / 5 — Edit — View inline comments

ROAD TO SCIENCE

The full ESS experiment pipeline has been demonstrated

Created by Carina Lobley, last modified by Ulrika Hammarlund on Mar 05, 2024

The Experiment Controls and Data Curation group (ECDC) and the Scientific Information Management Systems group (SIMS), along with their colleagues at the Data Management and Software Centre (DMSC), have demonstrated the full ESS experiment pipeline.

The ESS architecture for experiment control, data acquisition, and data curation requires orchestration of many software components. "Project CODA" (Continuous Data) has recently been completed by ECDC at ESS. Using a mixture of real data and previously recorded experiment data, continuous data collection has been run on ODIN, NMX, BIFROST, and the Test Beamline. The intention is to keep Project CODA running in preparation for commissioning activities and the user programme.

ESS controls and data path architecture (simplified). The action starts with the user at the NICOS workstation and data ends up in SciCat.

This marks a significant milestone for the ESS experiment data pipeline – congratulations to all those involved!

Please click here for a full report of project CODA.

Science Directorate | Feature story | General

BIFROST	1	31 Oct 2024	30 Jun 2025	09 Oct 2025	YES
DREAM	1	31 Oct 2024	30 Jun 2025	09 Oct 2025	YES
LOKI	1	31 Oct 2024	30 Jun 2025	09 Oct 2025	YES
ODIN	1	01 Nov 2024	30 Sep 2025	09 Oct 2025	YES
NMX	1	16 Oct 2025	30 Sep 2025	09 Oct 2025	YES
ESTIA	2	31 Mar 2025	31 Dec 2025	same	YES
FREIA	2	01 Jul 2027	30 Jun 2027	same	NO
MAGIC*	2	30 Jun 2025	31 Mar 2027	same	NO
BEER	2	30 Sep 2025	30 Sep 2026	same	NO
SKADI	2	23 Dec 2026	30 Sep 2026	same	NO
HEIMDAL*	3	01 Jul 2027	30 Jun 2027	same	NO
CSPEC*	3	20 Dec 2024	30 Jun 2027	same	NO
MIRACLES*	3	02 Jul 2027	30 Jun 2027	same	NO
TREX	3	02 Jul 2027	30 Sep 2027	same	NO
VESPA	3	01 Jul 2027	30 Sep 2027	same	NO

Questions?