PaNOSC & ExPaNDS : résultats des projets EOSC

8th April 2022

Presenters: **Jean-François Perrin** (ESRF)

Contributors: Andy Götz (ESRF) , Patrick Fuhrmann (DESY)

on behalf of PaNOSC + ExPaNDS partners

PaNOSC and ExPaNDS projects have received funding from the European Union's Horizon 2020 research and innovation programme under grant agreements 823852 and 857641, respectively.

Quick reminder : The Projects Cheat Sheet

	PaNOSC	ExPaNDS
EU Call	HORIZON 2020 INFRA-EOSC-04	HORIZON 2020 INFRA-EOSC-5B
Description	Cluster of ESFRI PaN Sources	EOSC PaN Data Services
Partners	ESRF, ILL, ESS, EU-XFEL, CERIC-ERIC, ELI-DC, EGI	DESY, ALBA, DLS, ELETTRA, EGI, HZB, HZDDR, Max IV, PSI, Soleil, UKRI
Observers	GEANT EU-DAT National RI's	
Linked 3rd Party	DESY STFC CESNET	
Start – End (Duration)	2018-12-01 – 2022-11-30 [4 Years]	2019-09-01 – 2023-02-28 [3 ½ Years]
Coordinators	A. Götz, G. Bodera	P. Fuhrmann, S. Servan
Budget	12 M Euros	6 M Euros
Home Page	PaNOSC.EU	ExPaNDS.EU
Twitter	@PaNOSC_eu #PaNOSC	@ExPaNDS_eu #ExPaNDS
GitHUB	github.com/panosc-eu	Github.com/expands-eu

PaNOSC and ExPaNDS projects have received funding from the European Union's Horizon 2020 research and innovation programme under grant agreements 823852 and 857641, respectively.

PaN facilities covered by PaNOSC/ExPaNDS

PaNOSC

ExPaNDS

Photon (LEAPS)

EGI

eli delivery consortium

CERIC
Central European Research Infrastructure Consortium

Neutron (LENS)

DESY
CESNET
STFC

PaNOSC and ExPaNDS projects have received funding from the European Union's Horizon 2020 research and innovation programme under grant agreements 823852 and 857641, respectively.

panosc photon and neutron open science cloud

ExPaNDS European Open Science Cloud Photon and Neutron Data Services

PaN Open Science Commons - Concept

- **Vision** – create a common space for PaNOSC and ExPaNDS facilities where petabytes of PaN FAIR data, analysis software, notebooks, workflows, and training material can be **F**ound, **A**ccessed (downloaded and/or executed), **R**e-Used + **I**mproved i.e. **FAIR**

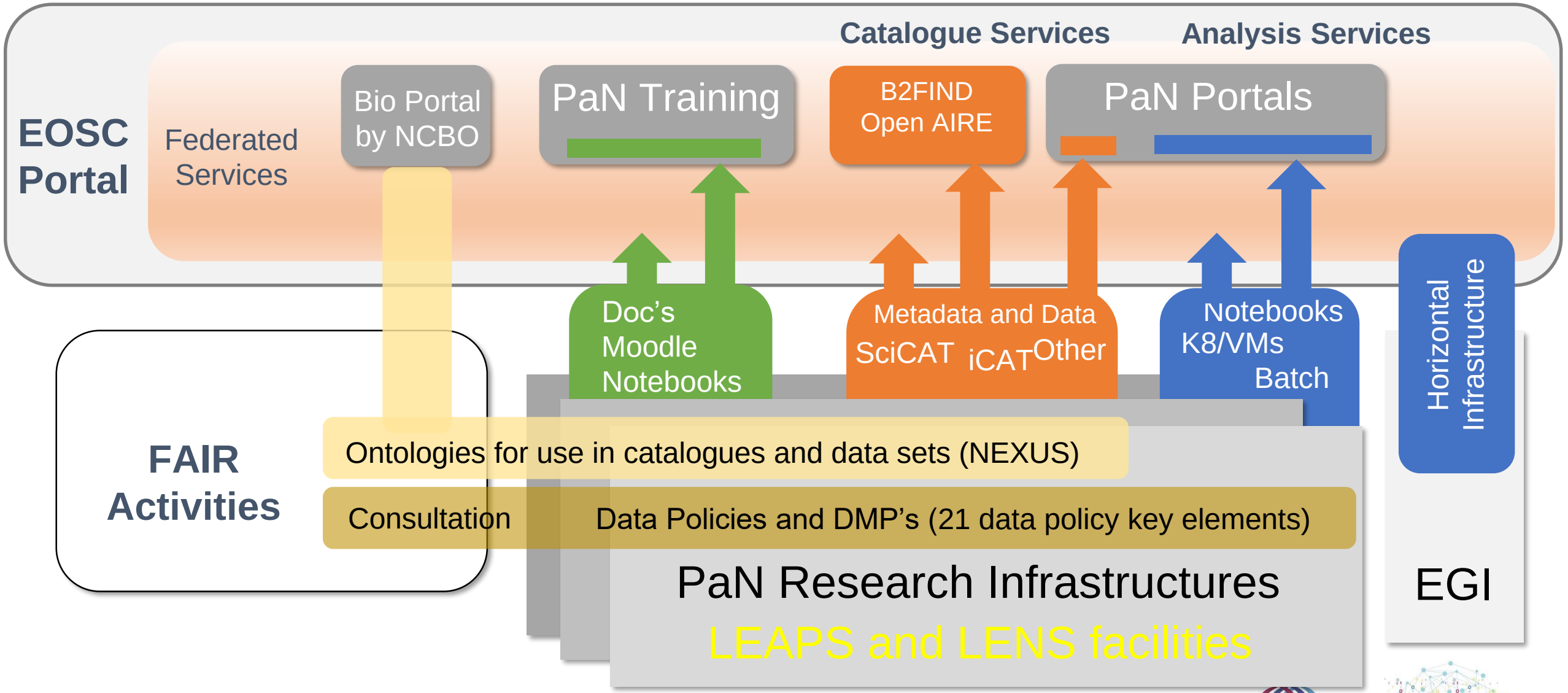
- **Remote access** – the PaN commons will be accessible remotely while being executed locally (close to the data) or via the EOSC (data needs to be moved)

- **Remote users** – the PaN commons will enable and encourage remote users and experiments (urgently required in the post-COVID-19 phase)

10 Primary Outcomes of PaNOSC and ExPaNDS

1. **FAIR data policy** and **DMPs**
2. **FAIR assessment** and common **PID** framework
3. Standardised metadata (**Nexus/HDF5**, PaN ontologies)
4. **Federated search API** for PaN data catalogues
5. **Open Data portal** for searching + downloading data
6. Community **AAI UmbrellaId**
7. **JupyterLab notebooks** and **Nexus/HDF5** files visualisation
8. **Remote data analysis** with VISA + data analysis pipelines
9. **Simulation** software for simulating experimental data (SIMEX)
10. **PaN-learning** platform (pan-learning.org + pan-training.org)

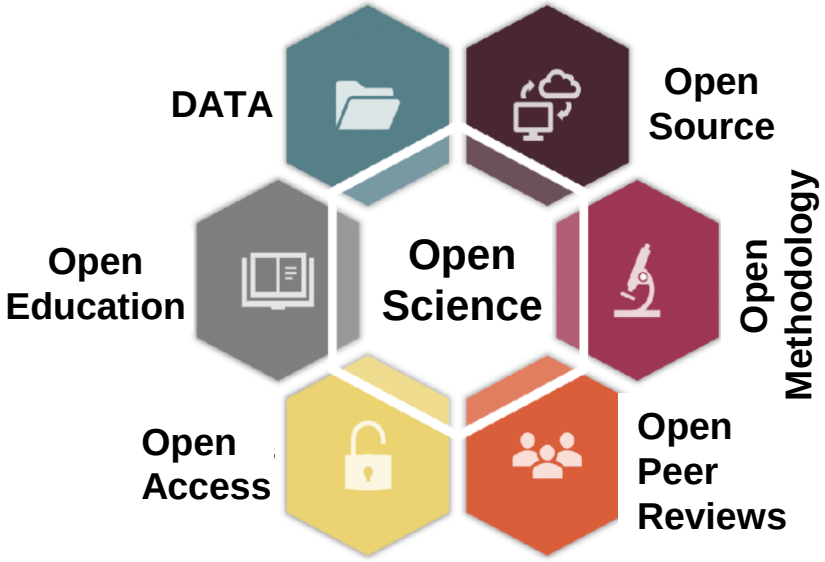
The Big Picture of EOSC in PaNOSC and ExPaNDS



PaNOSC and ExPaNDS projects have received funding from the European Union's Horizon 2020 research and innovation programme under grant agreements 823852 and 857641, respectively.

What did we achieve and what will we have achieved until the end of our projects?

PaNOSC and ExPaNDS projects have received funding from the European Union's Horizon 2020 research and innovation programme under grant agreements 823852 and 857641, respectively.



Open Science:
Open Data, FAIR Data Policies, PIDs, catalogues, DMPs .....

PaNOSC and ExPaNDS projects have received funding from the European Union's Horizon 2020 research and innovation programme under grant agreements 823852 and 857641, respectively.

Why are we involving ourselves into those policies?

The ultimate goal is Open Science

Accelerates scientific findings, as it:

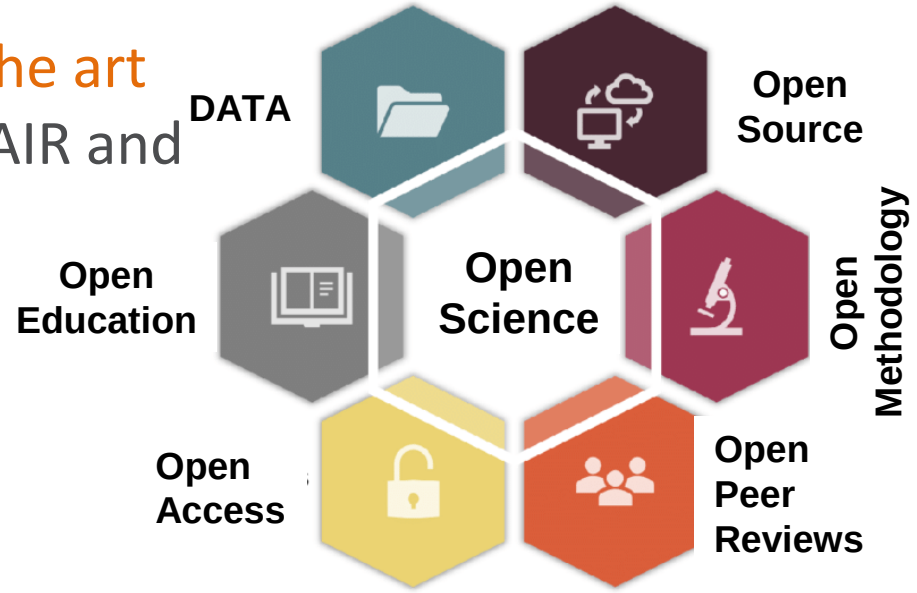
- Allows verifiable results;
- Allows to easily find relevant data;
- Reduces the download-data barrier;
- Allows to **re-use data** from other groups or to find data not yet analyzed or published by the original authors, trainings new scientists, development of software, … ;
- We can benefit from similar sciences using synergies;
- Avoid statements like:
  - "The original data can be provided by the author on request!"

PaNOSC and ExPaNDS projects have received funding from the European Union's Horizon 2020 research and innovation programme under grant agreements 823852 and 857641, respectively.

Why are we involving ourselves into those policies,

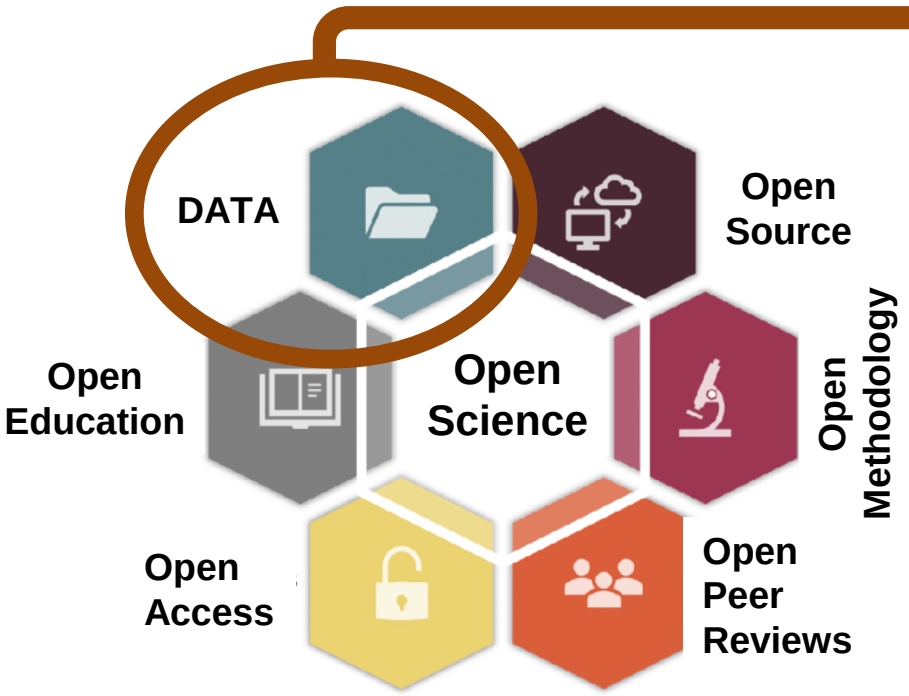
and why by means of the European Commission?

- The Commission provides us with established, state of the art frameworks to evaluate and implement Data Policies, FAIR and DMP, PID, by funding
  - EOSC with
  - specialized projects like FAIRsFAIR and
  - by supporting international organizations like RDA.
- provides specialized funding for
  - Activities to agree on common standards between the facilities
    'working together makes us stronger'
  - Activities helping us to make our services cross facility portable and EOSC
- However EC is making Open Science a condition.
  - No free lunch!



PaNOSC and ExPaNDS projects have received funding from the European Union's Horizon 2020 research and innovation programme under grant agreements 823852 and 857641, respectively.

From Open Science to FAIR data.



DOI: 10.5281/zenodo.5205825

Elements of the Data Policy Framework

- FAIR Data
- Rights of Stake holders
- Retention Policy
- Embargo Licenses
- Curation
- Discovery
- FAIR Assess-ment
- DMPs
- ***

- ✓ Global Unique PIDs
- ✓ Contextual Metadata, following community standards (Ontologies)
- ✓ Open Access Protocols (Data, AAI)
- ✓ Human and machine readable access to data and meta data.
- ✓ Standard Fileformats

PaNOSC and ExPaNDS projects have received funding from the European Union's Horizon 2020 research and innovation programme under grant agreements 823852 and 857641, respectively.

Our contribution to Data Polices, DMPs and FAIR data handling.

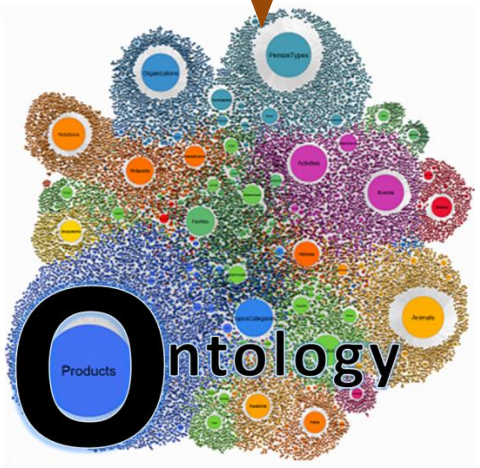
We provide

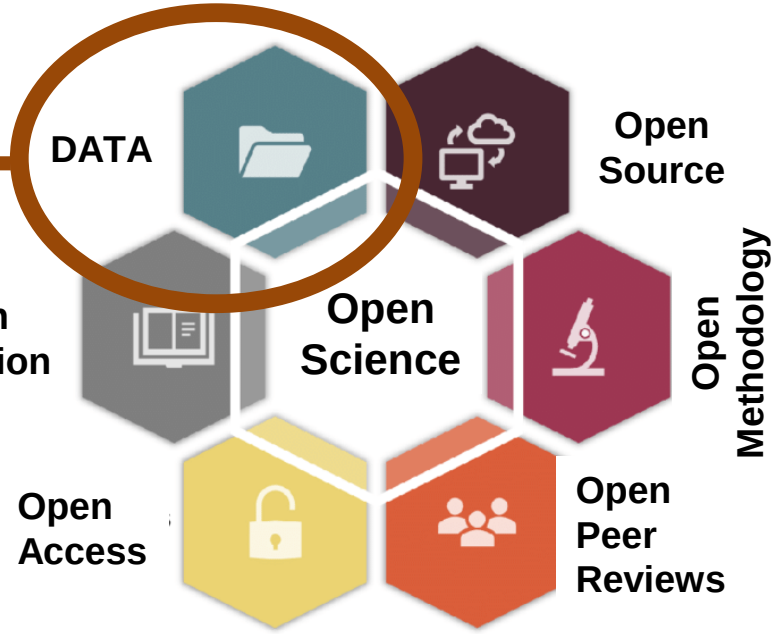
- Data Policy Framework for easy adoption in not so complex cases
- Data Policy Guidebook to compose a customized, more complex Data Policy.
- Per Facility consultation
  - We provided the Data Policy Guidebook to the facility responsible person.
  - We provided a facility by facility consultation and discussed the Guide Book.
  - We collected feedback and redesigned accordingly.
  - We help and monitor the implementation or the Policies at the facilities.
- Help on permanent FAIR assessment of the facilities
  - We provide guidelines on how to continuously verify FAIRness of the beamline data taking process, the meta data and the repositories.

PaNOSC and ExPaNDS projects have received funding from the European Union's Horizon 2020 research and innovation programme under grant agreements 823852 and 857641, respectively.

panosc
photon and neutron open science cloud

ExPaNDS
European Open Science Cloud Photon and Neutron Data Services





Meta Data Ontologies, Keywords, Catalogues and File Formats

photon

Showing 1 of 1099 Sort: Search Rank

PaN Experimental technique (PANET)

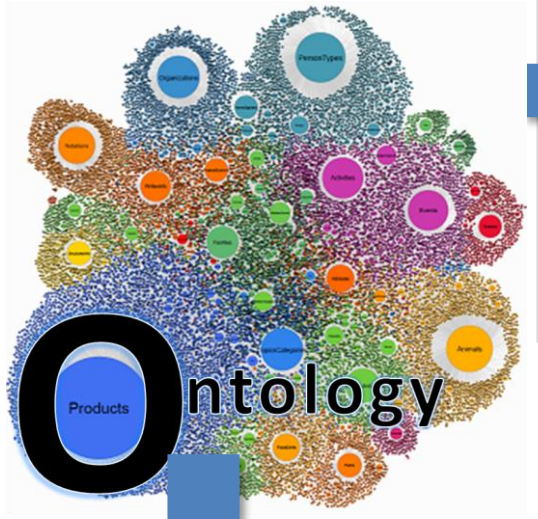
classes 379

The PaNET ontology provides a taxonomy and thesaurus of photon and neutron (PaN) experimental techniques, based mainly on accelerator-based light sources and neutron facilities

Uploaded: 6/19/21

PaNOSC and ExPaNDS projects have received funding from the European Union's Horizon 2020 research and innovation programme under grant agreements 823852 and 857641, respectively.

Meta Data Ontologies and their consequences



BioPortal

EUROPEAN OPEN SCIENCE CLOUD

Keywords

Defining NeXus Keywords in a common file format like HDF5

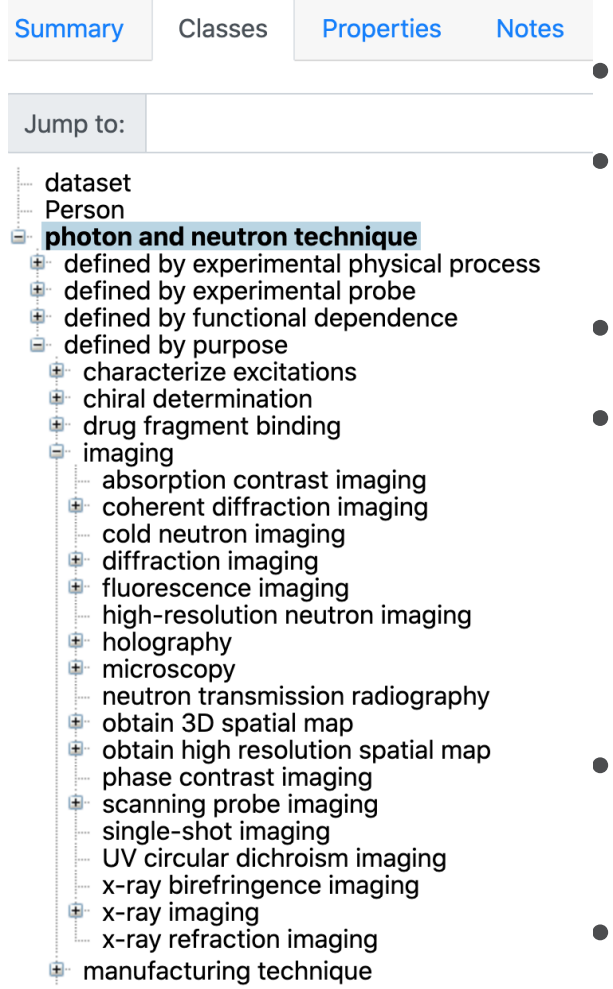
Keywords for Meta Data Catalogues and the ‘PaN Search API’

PaNOSC and ExPaNDS projects have received funding from the European Union’s Horizon 2020 research and innovation programme under grant agreements 823852 and 857641, respectively.

panosc photon and neutron open science cloud

ExPaNDS European Open Science Cloud Photon and Neutron Data Services

Our contribution to Ontologies, NeXus formats and catalogue definitions



- Worked on a common ontology for different techniques in PaN;
- Checked ‘keywords’ for duplication in other sciences and searchability;
- Modelled our Ontology into the BioPortal;
- Working on making the BioPortal (not our service) available through the EOSC Portal;

BioPortal

EUROPEAN OPEN SCIENCE CLOUD

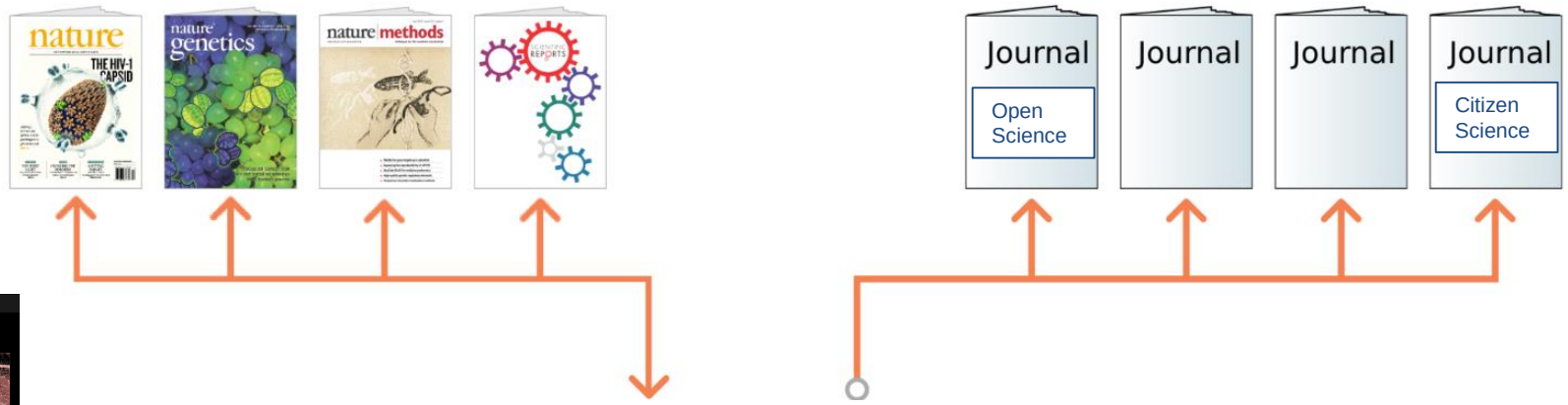
- Publishing hints for catalogues to pick up the right keywords form our ontology;
- Involved in the NeXus definitions following our ontology. (NIAC Involvement).

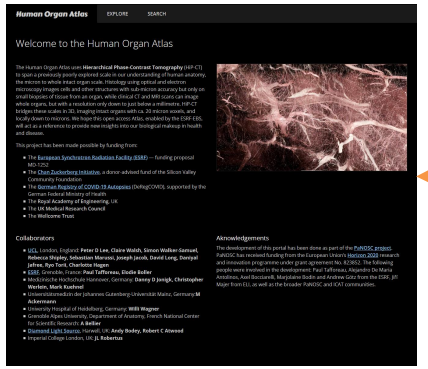
PaNOSC and ExPaNDS projects have received funding from the European Union’s Horizon 2020 research and innovation programme under grant agreements 823852 and 857641, respectively.

panosc photon and neutron open science cloud

ExPaNDS European Open Science Cloud Photon and Neutron Data Services

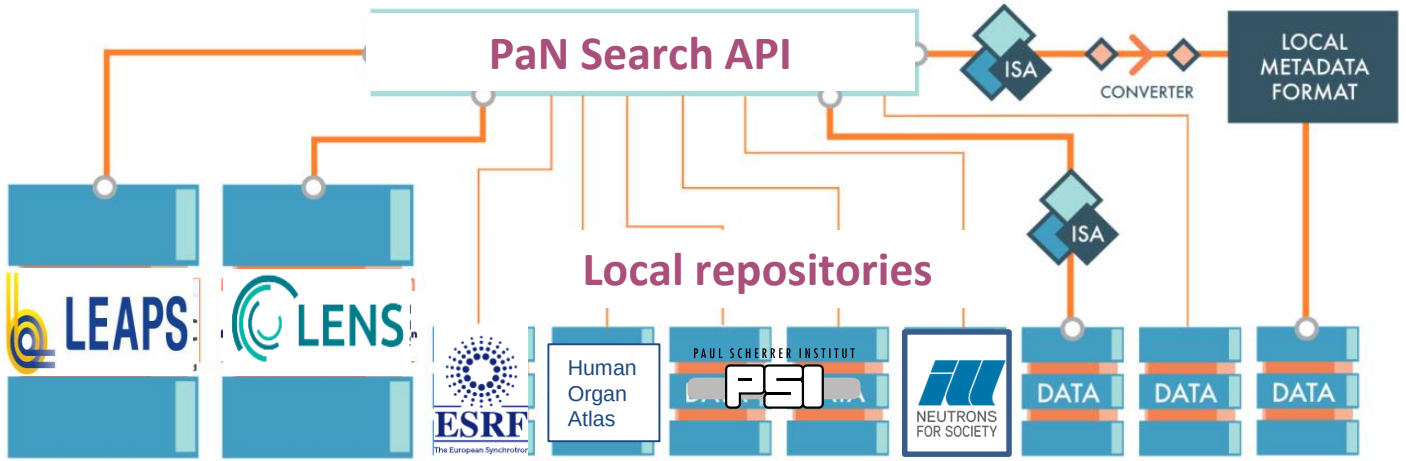
Sustain Published Data ➡ Through PaN Repositories





PaN Open Data Portal

PaN Open Data Repository



ICAT

SciCat

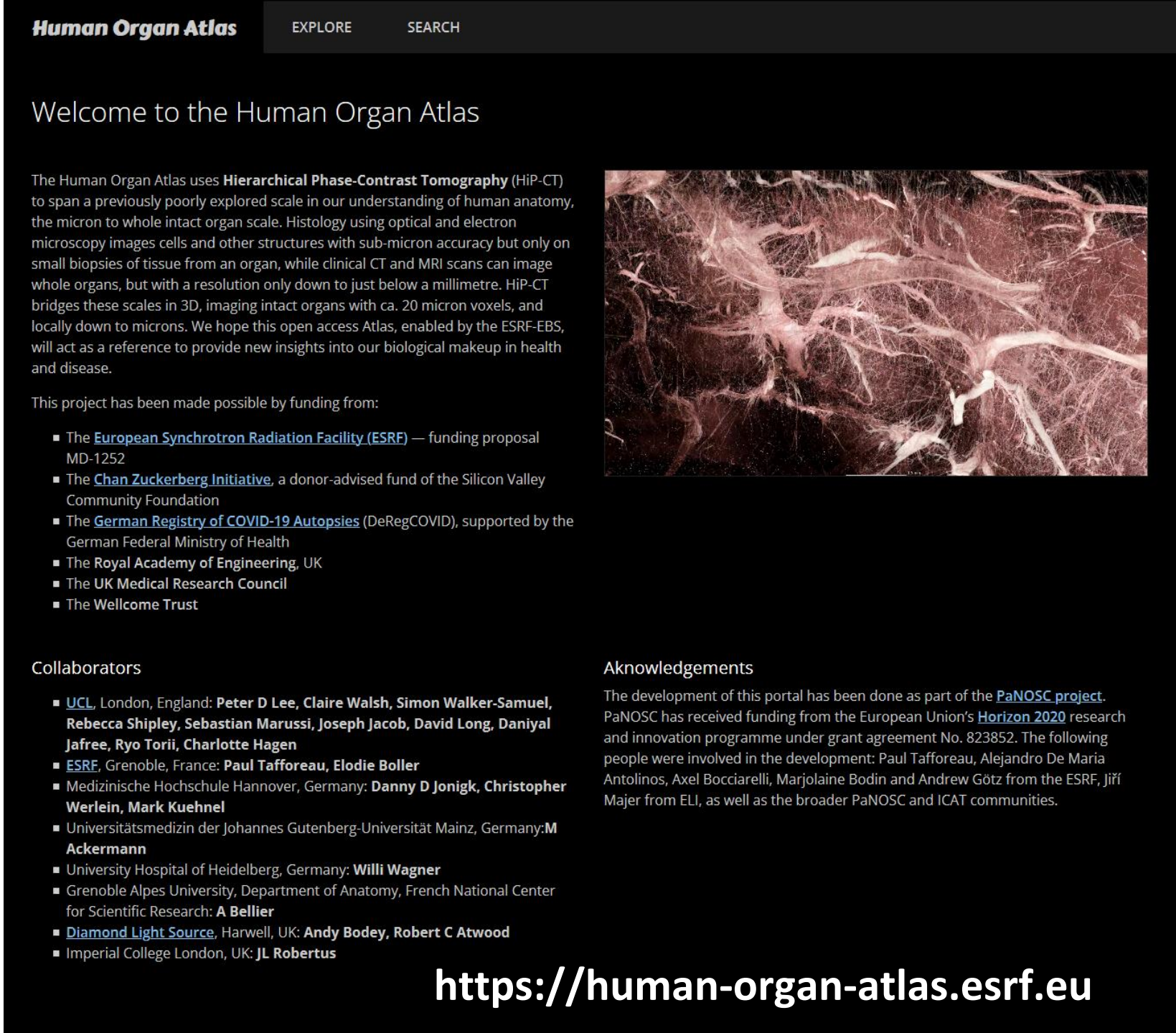
image: http://blogs.nature.com/scientificdata/2013/07/23/scientific-data-to-complement-and-promote-public-data-repositories/

Example of open data publishing:

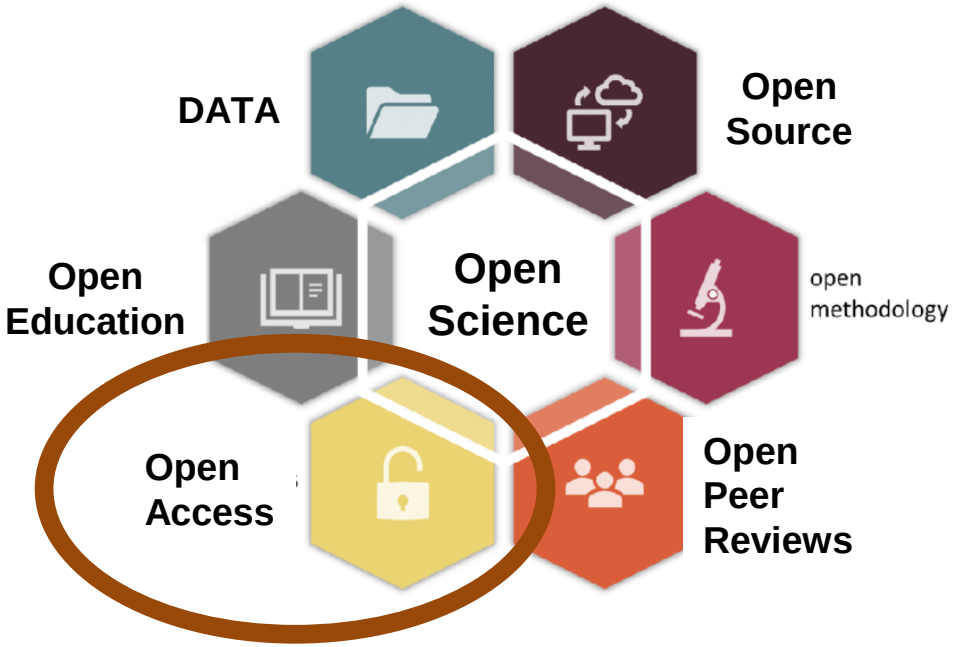
Human Organ Atlas

*Walsh, C.L., * Tafforeau, P., * Wagner, W.L., Jafree, D.J., Bellier, A., Werlein, C., Kühnel ,M.P., Boller, E., Walker-Samuel, S., Robertus, J-L., Long, D.A., Jacob, J., Marussi, S., Brown, E., Holroyd, N., Jonigk#, D.D., Ackermann#, M., Lee#, P.D. **Imaging intact human organs locally resolving cellular structures using hierarchical phase- contrast tomography.** Nat Methods (2021) Accepted

refer to PaNOSC Use Case 23 for more info



Prerequisite for open access : the common AAI



PaNOSC and ExPaNDS projects have received funding from the European Union's Horizon 2020 research and innovation programme under grant agreements 823852 and 857641, respectively.

PaN and EOSC Core service: Authentication and Identity Management (AAI)

- Originally
  - Scientist needed a different identity at each facilities to access their services.
    - Lots of passwords to remember and to loose
    - Difficult with cross facility services.

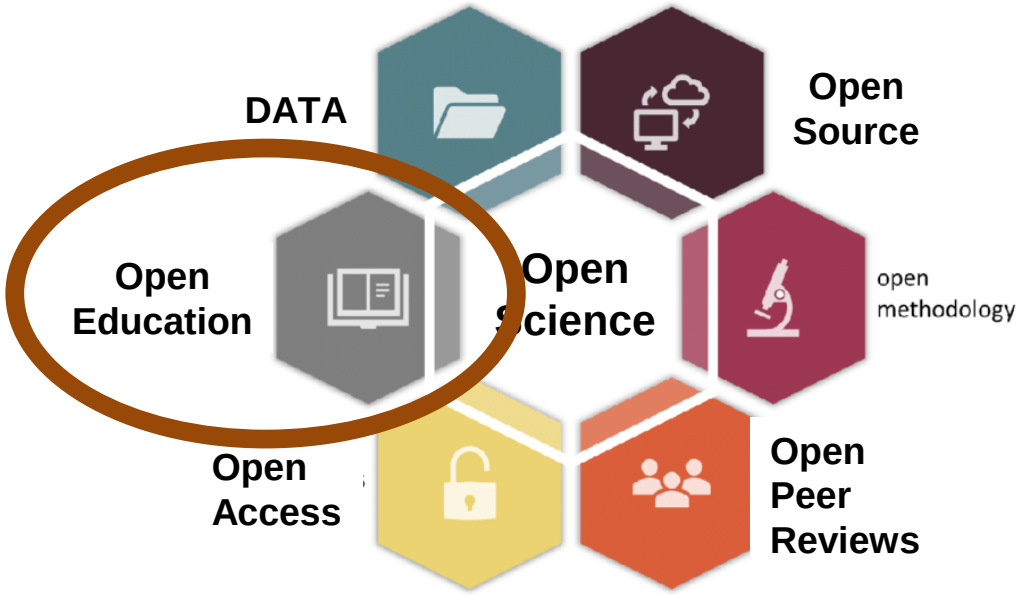
- UmbrellaID
  - Scientist only needs one identity with UmbrellaID (on top of home identity)
  - 'Catch all' identity providers are no longer state of the art and not trusted.

- Now: eduTeams [GÉANT service] (in progress)
  - Scientist only needs the one identity from his/her home facility!
  - Single Sign On: for cross facility services (data orchestration and automatic analysis)
    - One service can use another services somewhere else on your behalf!
  - Less Prison Time: Legal Issues for IdPs and Services are sorted out (AARC Blueprint)
  - Security: We are part of the European wide CERT system.

- EOSC-Future -> Science agnostic AAI
  - Scientist can use his home Identity for different services cross communities.

PaNOSC and ExPaNDS projects have received funding from the European Union's Horizon 2020 research and innovation programme under grant agreements 823852 and 857641, respectively.

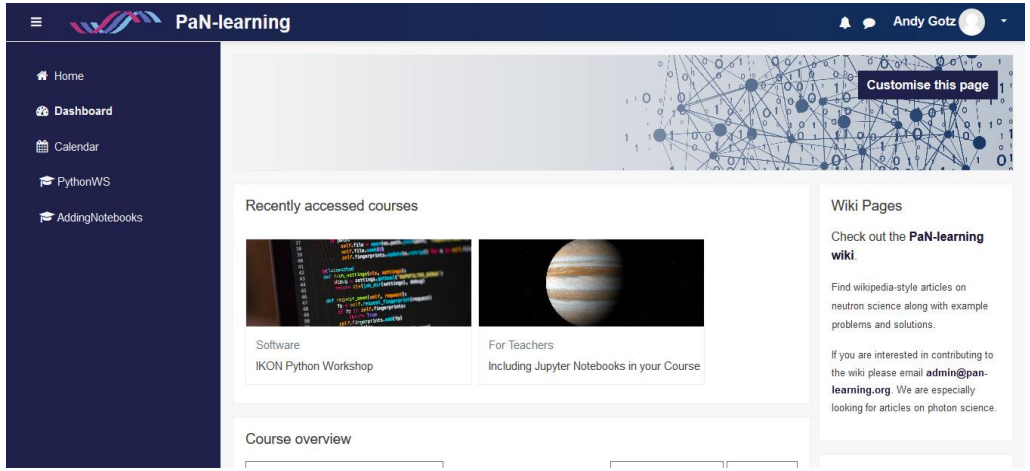
Training and Learning Platform



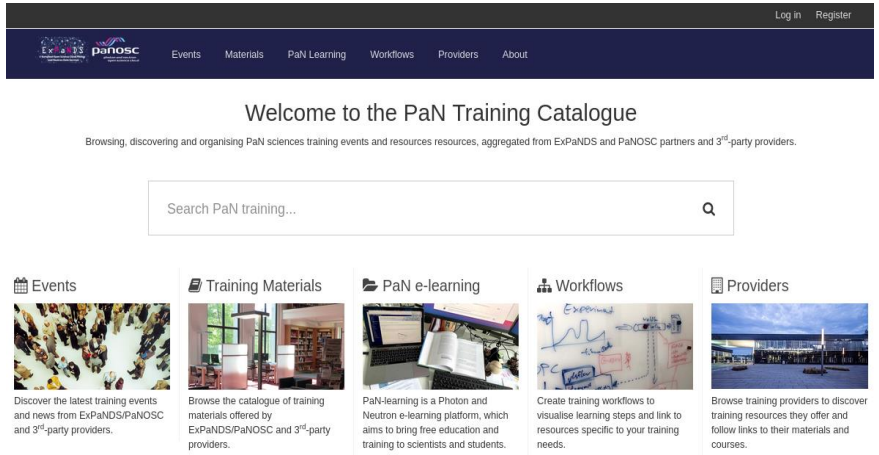
PaNOSC and ExPaNDS projects have received funding from the European Union's Horizon 2020 research and innovation programme under grant agreements 823852 and 857641, respectively.

Training and Learning Platforms

- A centralized platform for collecting our highly distributed knowledge, documents and teaching material
  - which was identified at the facilities as important and missing.
- One can see golden thread in the current disparity of content.



https://pan-training.hzdr.de/

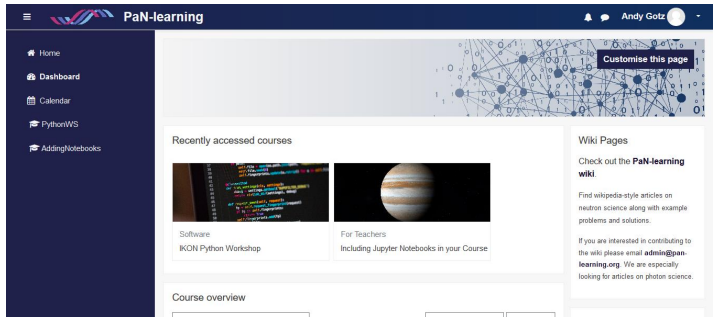


https://pan-learning.org

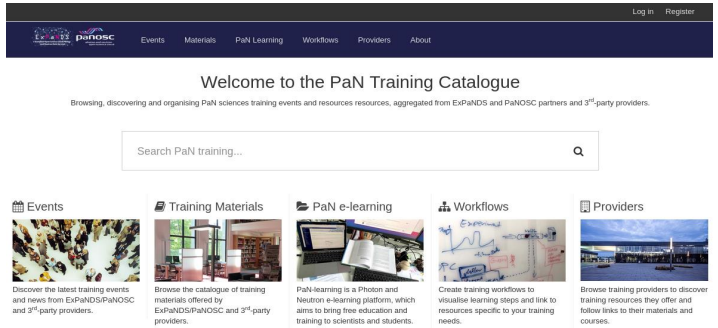
PaNOSC and ExPaNDS projects have received funding from the European Union's Horizon 2020 research and innovation programme under grant agreements 823852 and 857641, respectively.

Our contribution to the teaching and learning platform.

- We evaluated available, state of the art technologies for teaching and data collection platforms.
- We introduced a PaN training platform to
  - **create**/store courses and to
  - **collect** existing material.
- We are re-using successful projects developed by
  - **Elixir** (TeSS) and
  - **SINE2020** e-neutrons.



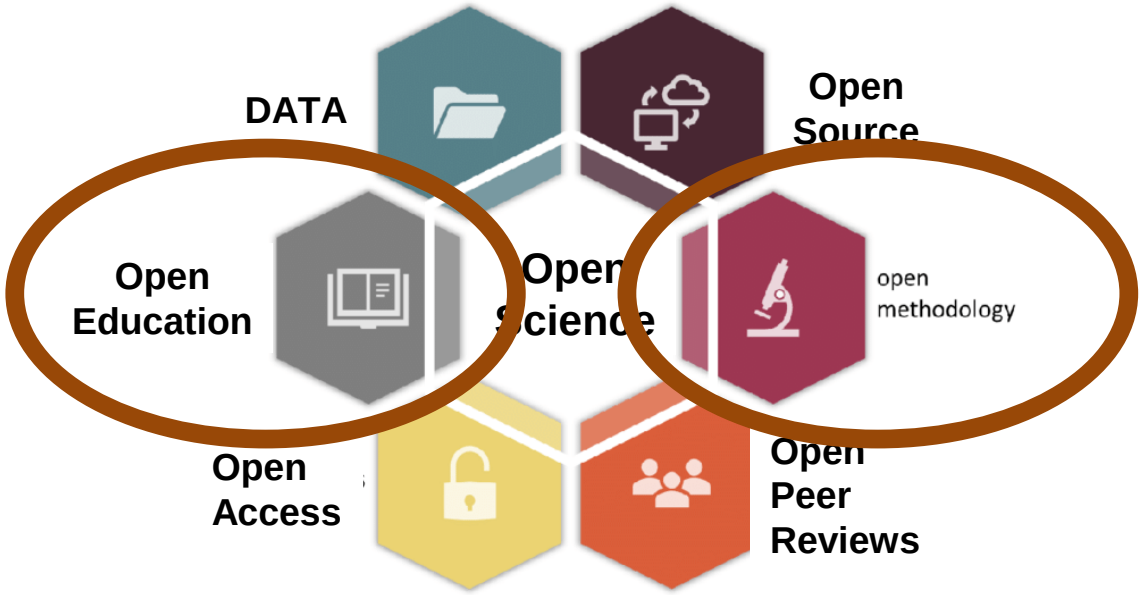
https://pan-training.hzdr.de/



https://pan-learning.org

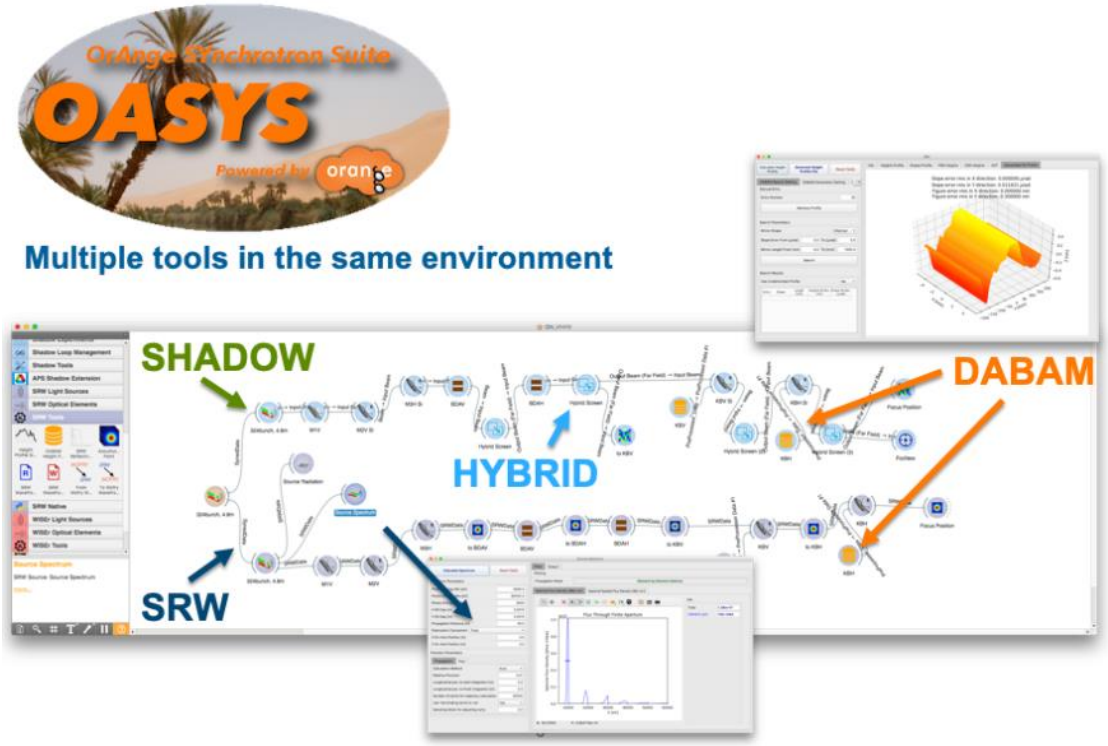
PaNOSC and ExPaNDS projects have received funding from the European Union's Horizon 2020 research and innovation programme under grant agreements 823852 and 857641, respectively.

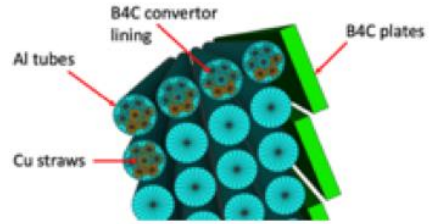
Data analyse services

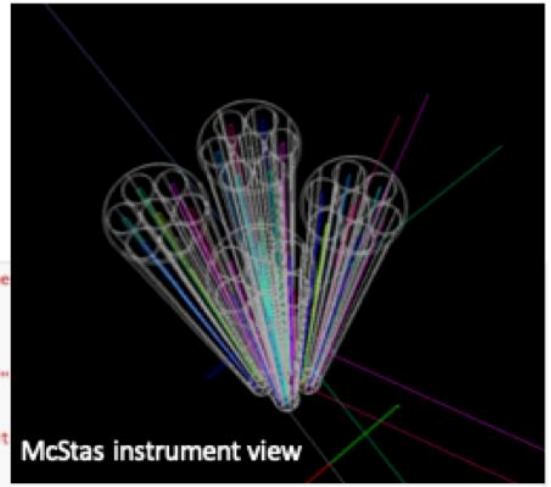


PaNOSC and ExPaNDS projects have received funding from the European Union's Horizon 2020 research and innovation programme under grant agreements 823852 and 857641, respectively.

Simulation in PaNOSC provides support to OASYS + McStas







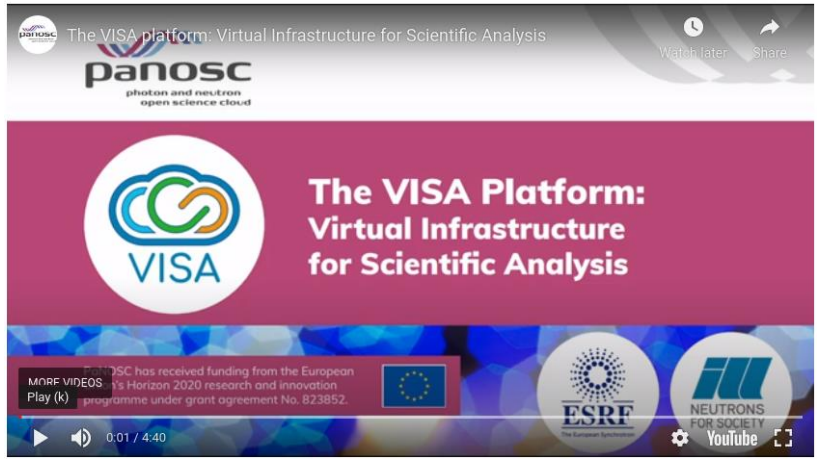

```
In [4]: Al_incoherent = ISIS_SANS2d_Mantid.add_component("Al_incohe
        Al_incoherent.sigma = "4+0.0082"
        Al_incoherent.packing_factor = 1.0
        Al_incoherent.unit_cell_volume = 66.4

        Al_powder = ISIS_SANS2d_Mantid.add_component("Al_powder", "
        Al_powder.reflections = "\"Al.laz\""

        Al = ISIS_SANS2d_Mantid.add_component("Al", "Union_make_mat
        Al.process_string = "\"Al_incoherent,Al_powder\""
        Al.my_absorption = "100*4+0.231/66.4"

        Cu_incoherent = ISIS_SANS2d_Mantid.add_component("Cu_incoherent", "Incoherent_process")
        Cu_incoherent.sigma = "4+0.55"
        Cu_incoherent.packing_factor = 1.0
        Cu_incoherent.unit_cell_volume = 47.24

        Cu_powder = ISIS_SANS2d_Mantid.add_component("Cu_powder", "Powder_process")
        Cu_powder.reflections = "\"Cu.laz\""

        Cu = ISIS_SANS2d_Mantid.add_component("Cu", "Union_make_material")
        Cu.process_string = "\"Cu_incoherent,Cu_powder\""
        Cu.my_absorption = "100*4+3.78/47.24"
executed in 9ms, finished 21:52:02 2020-12-15
```

PaNOSC continued simulation activities started in EUCALL with SIMEX, lasers + neutrons

PaNOSC and ExPaNDS projects have received funding from the European Union's Horizon 2020 research and innovation programme under grant agreements 823852 and 857641, respectively.

Users need remote data analysis

- Why it’s important
  - next generation of data analysis in globalised research
  - **remote access** to facilities
- What we do
  - develop and deploy **VISA** platform
  - make **Jupyter** notebooks available at all sites
  - make data analysis pipelines **interoperable**
- Provides remote data analysis services in a web browser with access to
  - Experimental data
  - Scientific software
  - Compute ressources
  - Support (IT and Scientific)



https://bit.ly/VISA-video

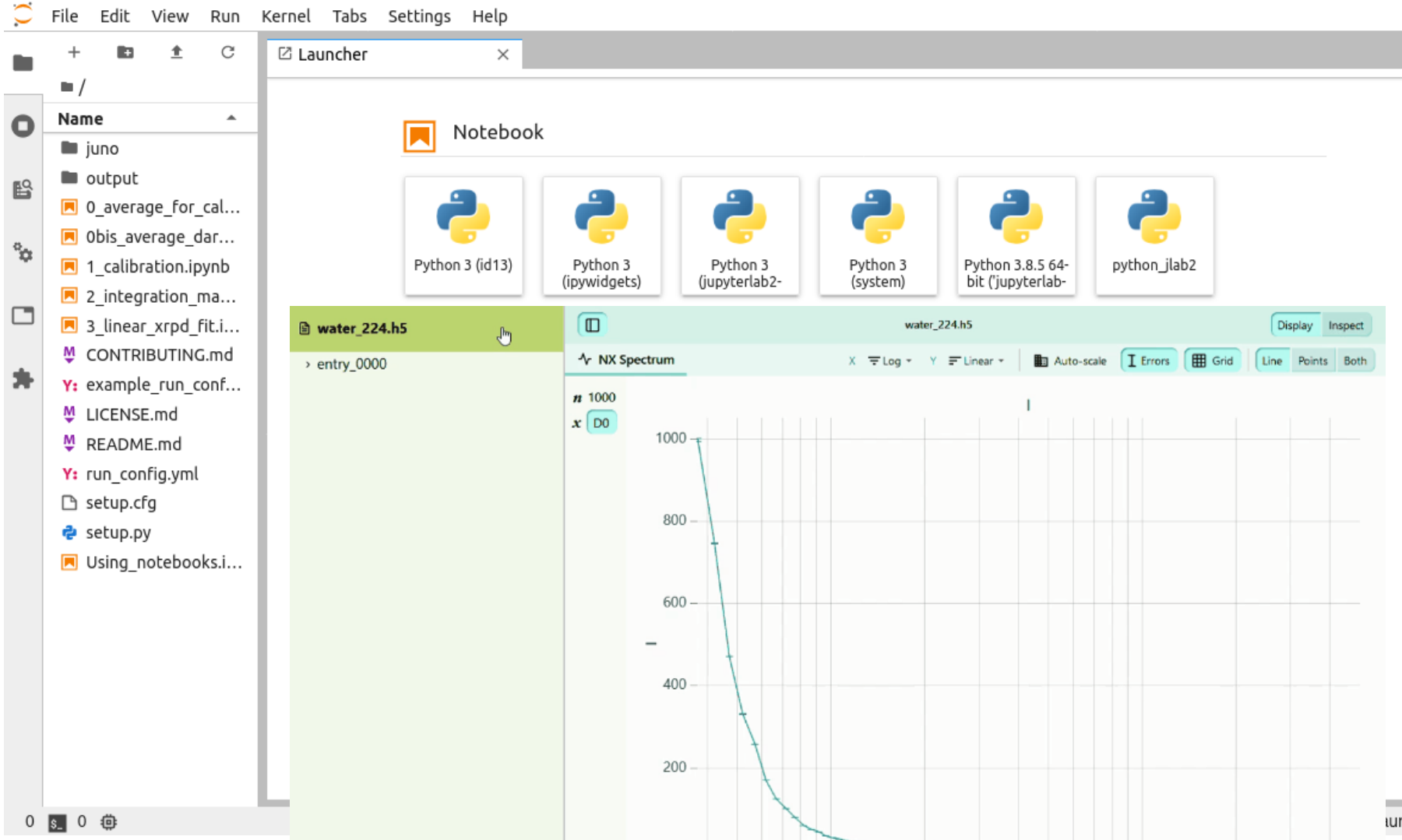
panosc
photon and neutron open science cloud

ExPaNDS
European Open Science Cloud Photon and Neutron Data Services

Remote analysis + visualization with Jupyter notebooks

- **JupyterLab has been widely adopted as remote analysis tool**

- **PaNOSC provides**
  - Jupyter-slurm
  - Nexus/HDF5 visualization
- **PaNOSC + ExPaNDS**
  - Developed for Use Cases

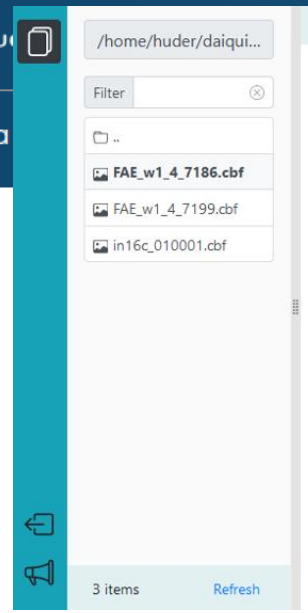


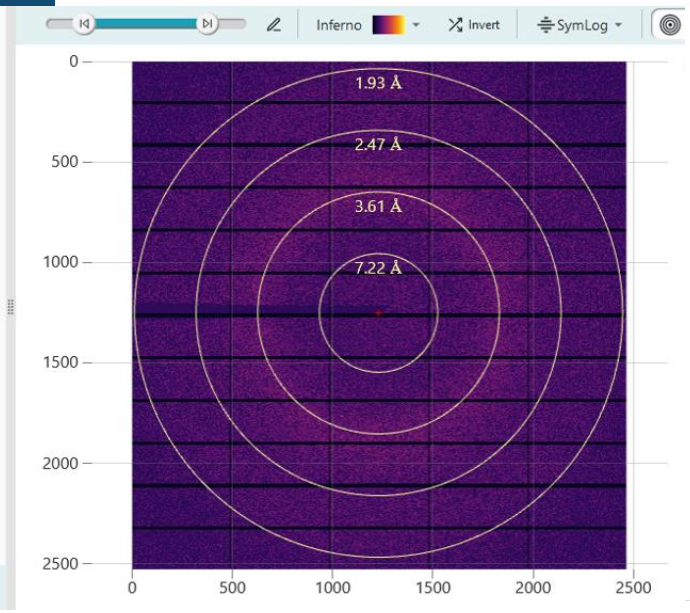
PaNOSC and ExPaNDS projects have received funding from the Europe
and innovation programme under grant agreements 823852 and 857641,

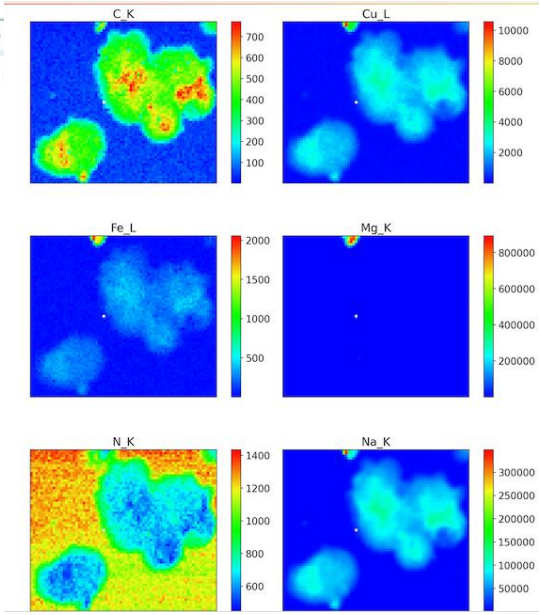
https://www.panosc.eu/all-use-cases/

All Use Cases

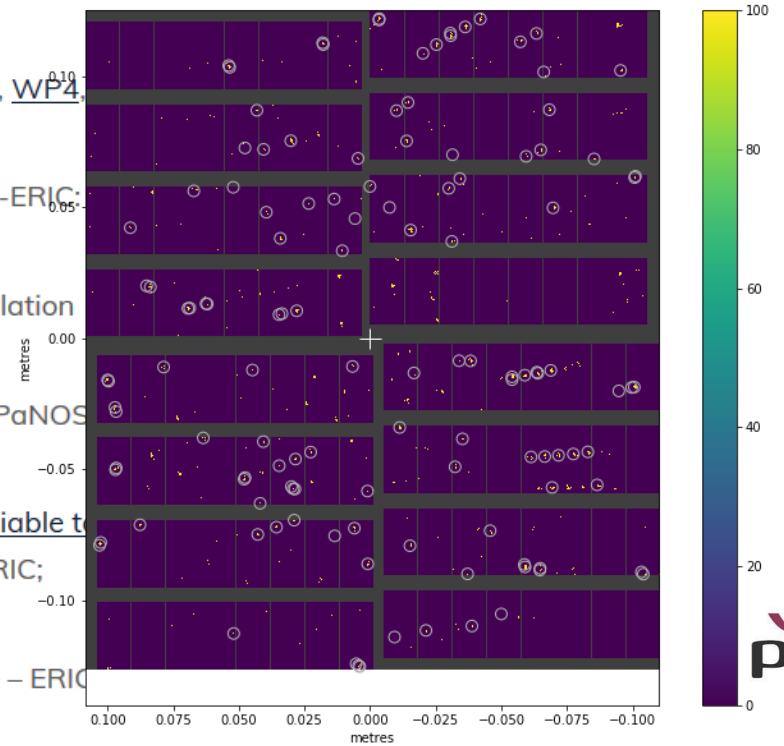
Submit your use case

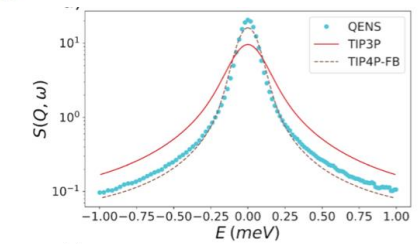


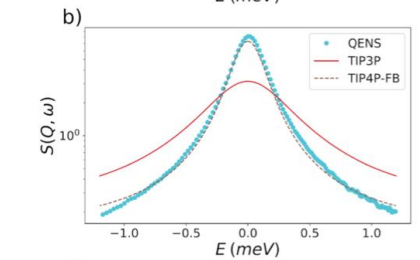


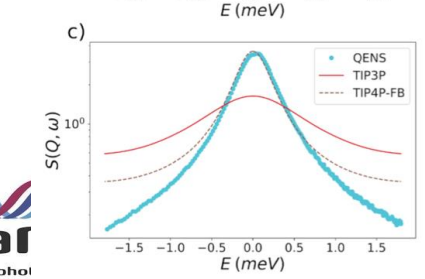


- **Use Case 22** – BRAGGY diffraction image viewer (ESRF; PaNOSC-related WPs: WP3, WP4, WP6)
- **Use Case 21** – Online visualization and analysis of HDF5 PyMCA Output Files (CERIC-ERIC; PaNOSC-related WP: WP4)
- **Use Case 20** – Using an e-learning platform to support presentations (European Spallation Source – ERIC, ESS-ERIC; PaNOSC-related WP: WP8)
- **Use Case 19** – Online visualization and analysis of Bruker NMR output (CERIC-ERIC; PaNOS related WP: WP4)
- **Use Case 18** – In-silico Neutron diffraction from Boro-carbon systems: precise and reliable t for exact structural analysis and defect detection (Extreme Light Infrastructure, ELI-ERIC; PaNOSC-related WP: WP5)
- **Use Case 17** – Reusing data for validation of force fields (European Spallation Source – ERIC ESS-ERIC; PaNOSC-related WPs: WP3, WP4, WP5)









Examples from other clusters in EOSC

- **EOSC-Life**
  - EU COVID-19 portal
- **ESCAPE**
  - High Energy Physics – CERN Open Data
  - Virtual Observatory
- **ENVRI-FAIR**
  - ICOS, DISSCO, …
- **SSHOC**
  - CLARIN, DARIAH, …
- **PaNOSC+ExPaNDS**
  - ISIS, ILL, ESRF, PSI, …

PaNOSC and ExPaNDS projects have received funding from the European Union's Horizon 2020 research and innovation programme under grant agreements 823852 and 857641, respectively.

Thank you to all PaNOSC + ExPaNDS contributors!

PaNOSC and ExPaNDS projects have received funding from the European Union's Horizon 2020 research and innovation programme under grant agreements 823852 and 857641, respectively.