



PaNOSC Closing Event

Paving the way towards the PaN FAIR Data Commons

29-30 November 2022

Grenoble - France

PaNOSC Major Achievements

Andy Götz – PaNOSC Coordinator

ESRF

29 November 2022



PaNOSC has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement no. 823852

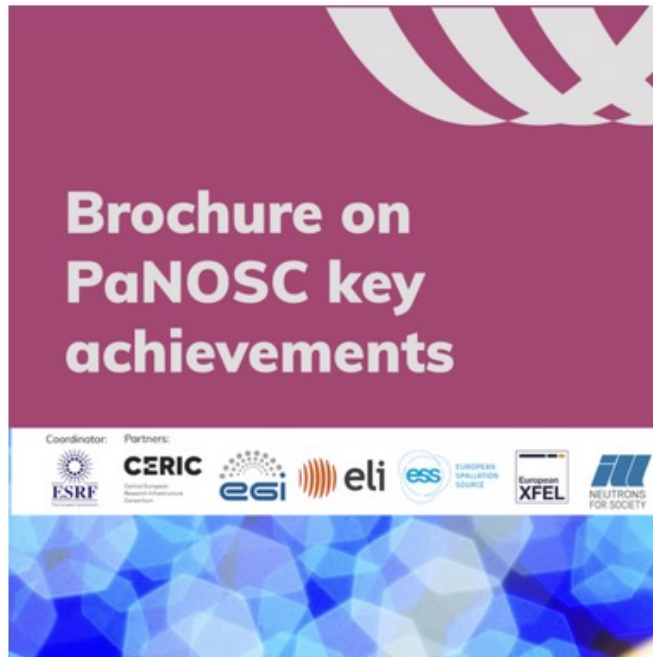
PaNOSC

Major Achievements



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

PaNOSC News - <https://www.panosc.eu/news/>



Published on 23 November 2022

Brochure on PaNOSC key achievements released

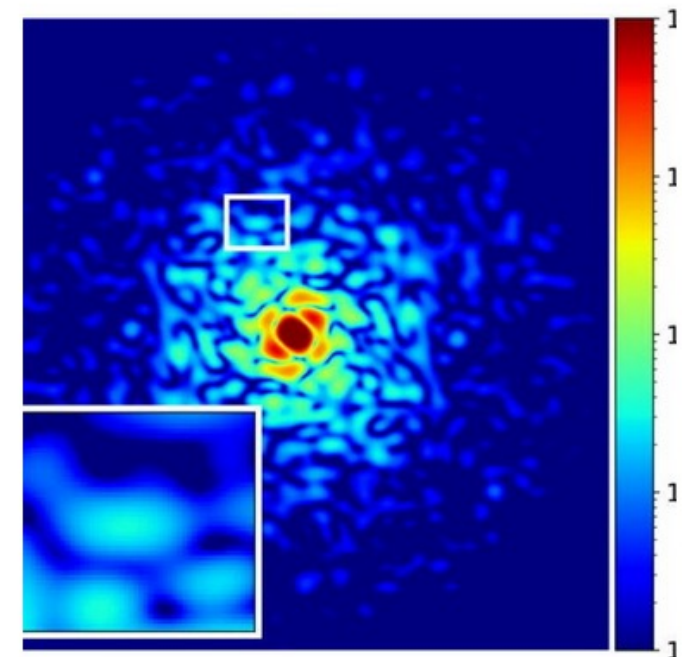
The PaNOSC final brochure has been released,



Published on 18 November 2022

3rd PaN EOSC Symposium – An Overview

18 October 2022, Brno & online PaNOSC and



Published on 18 November 2022

New SIMEX-based research paper just published in Structural Dynamics



Use Cases - <https://www.panosc.eu/all-use-cases/>



photon and neutron
open science cloud

Contacts Search PaNOSC 🔍

About ▾ Data ▾ Services ▾ Training ▾ Use Cases ▾ Materials ▾ News & Events ▾ EOSC ▾

All Use Cases

Submit your use case

- Use case 31 – [Seamless connection of Jupyter notebooks and GUI applications for e-learning purposes](#) (CERIC-ERIC, ESS; PaNOSC-related WP: [WP5](#), [WP8](#))
- Use Case 30 – [VISA – Data Analysis in the Cloud](#) (ILL; PaNOSC-related WP: [WP4](#))
- Use Case 29 – [Run orange-pylost as a cloud service](#) (ESRF; PaNOSC-related WP: [WP4](#))
- Use Case 28 – [Online visualisation, exploration and analysis of HDF5 files with h5nuvola](#) (CERIC-ERIC, PaNOSC-related WP: [WP4](#))
- Use Case 27 – [CCP4 cloud service for MX](#) (ESRF, PaNOSC-related WP: [WP6](#))
- Use Case 26 – [Data uploader for automatic transfer of curated experimental RAW data](#) (CERIC-ERIC, PaNOSC-related WP: [WP3](#))
- Use Case 25 – [WebKnossos, a web-based tool for 3D data viewing and annotation](#) (ESRF; PaNOSC-related WPs: [WP3](#), [WP4](#), [WP6](#))
- Use Case 24 – [View HDF5 files in ESRF Data Portal /ESRF; PaNOSC related WPs: WP3, WP4](#)

Use Cases

- Data Analysis
- Data Simulation
- Data Catalogue
- E-learning
- Software

PaN Data Portal - <https://data.panosc.eu>



European Photon and Neutron Open Data Search Portal

Type a query to search for open data from photon and neutron sources:

... or try one of these queries: *diffraction, lung*

The European Photon and Neutron sources are working together in the PaNOSC and ExPaNDS projects financed by the European Commission to build the **European Open Science Cloud**. One of the main objectives of the EOSC is to make **Open Data** from these facilities FAIR. This portal implements the F(indable) part of FAIR via a **federated search engine** from the following facilities:

- European Synchrotron Radiation Facility
- European Spallation Source
- Institut Laue Langevin
- MAX IV
- Paul Scherrer Institut
- Central European Research Infrastructure Consortium
- European XFEL

Additional facilities will be included in the federated search as their search engines come online locally. The goal is to include all photon and neutron facilities who provide open data by the end of the two projects PaNOSC and ExPaNDS.

The mission of the PaN data search portal is to contribute to the realization of a data commons for Neutron and Photon science. The search results provide a link to the landing page of the

Welcome to the Human Organ Atlas

The Human Organ Atlas uses **Hierarchical Phase-Contrast Tomography** to span a previously poorly explored scale in our understanding of human anatomy, the micron to whole intact organ scale.

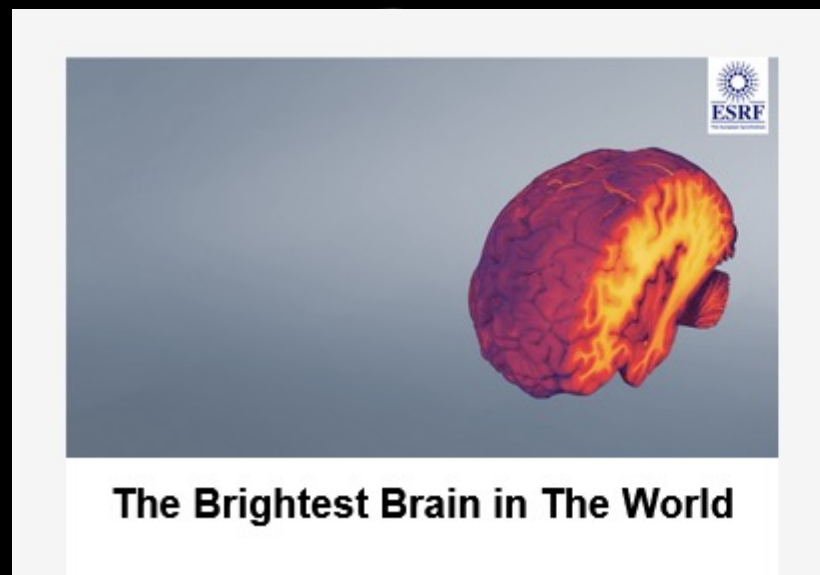
Histology using optical and electron microscopy images cells and other structures with sub-micron accuracy but only on small biopsies of tissue from an organ, while clinical CT and MRI scans can image whole organs, but with a resolution only down to just below a millimetre. HiP-CT bridges these scales in 3D, imaging intact organs with ca. 20 micron voxels, and locally down to microns.

We hope this open access Atlas, enabled by the ESRF-EBS, will act as a reference to provide new insights into our biological makeup in health and disease. To stay up to date, follow [@HiP-CT](#)

Funding

This project has been made possible by funding from:

- The [European Synchrotron Radiation Facility \(ESRF\)](#) — funding proposal MD-1252
- The [Chan Zuckerberg Initiative](#), a donor-advised fund of the Silicon Valley Community Foundation
- The [German Registry of COVID-19 Autopsies](#) (DeRegCOVID), supported by the German Federal Ministry of Health
- The Royal Academy of Engineering, UK
- The UK Medical Research Council



The Brightest Brain in The World

HiP-CT imaging and 3D reconstruction of a [complete brain](#) from the body donor LADAF-2020-31.

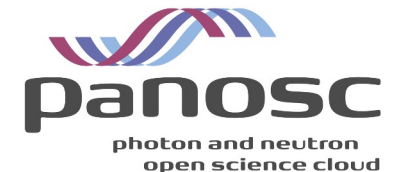
More videos can be viewed on the [HiP-CT YouTube channel](#).

Collaborators

- [UCL](#), London, England: **Peter D Lee, Claire Walsh, Simon Walker-Samuel, Rebecca Shipley, Sebastian Marussi, Joseph Jacob, David Long, Daniyal Jafree, Ryo Torii, Charlotte Hagen**
- [ESRF](#), Grenoble, France: **Paul Tafforeau, Elodie Boller**
- Medizinische Hochschule Hannover, Germany: **Danny D Jonigk, Christopher Werlein, Mark Kuehnel**
- Universitätsmedizin der Johannes Gutenberg-Universität Mainz, Germany: **M Ackermann**
- University Hospital of Heidelberg, Germany: **Willi Wagner**
- Grenoble Alpes University, Department of Anatomy, French National Center for Scientific Research: **A Bellier**

Active Data Management Plans

1. ExPaNDS and PaNOSC have adopted active DMPs
2. Active DMPs are updated at different phases of the project
3. ESS and ESRF have chosen to use DS Wizard developed by Elixir
4. Example of implementation @ ESRF
 1. Automatically generates a DMP automatically for every proposal
 2. 50 out of 82 questions are automatically filled in from DP/User/Data Portals
 3. DMPs offer a structured way to communicate information
 4. Users can use the DMP for satisfy funders requirements
 5. Next step is to use the DMPs to ensure users can manage their **data**



View

Current Phase

Before starting the project/proposal

Chapters

- I. General / Topic
- II. Content classification / Datas
- III. Technical classification
- IV. Metadata 2
- V. Ethics / General legal issues
- VII. Storage and long-term preservation / Selection**
- What are the criteria / rules for the selection of the data...
- Who selects the data to be archived?

Comments TODOs Version history

VII. Storage and long-term preservation / Selection

Here the user is asked to define their criteria for archiving data as well as the duration and accessibility of the data.

What are the criteria / rules for the selection of the data (after the end of the project)?

discussed between the instrument scientist and the user. Normally the instrument scientist can guide the user.

Desirable: *Before starting the project/proposal*

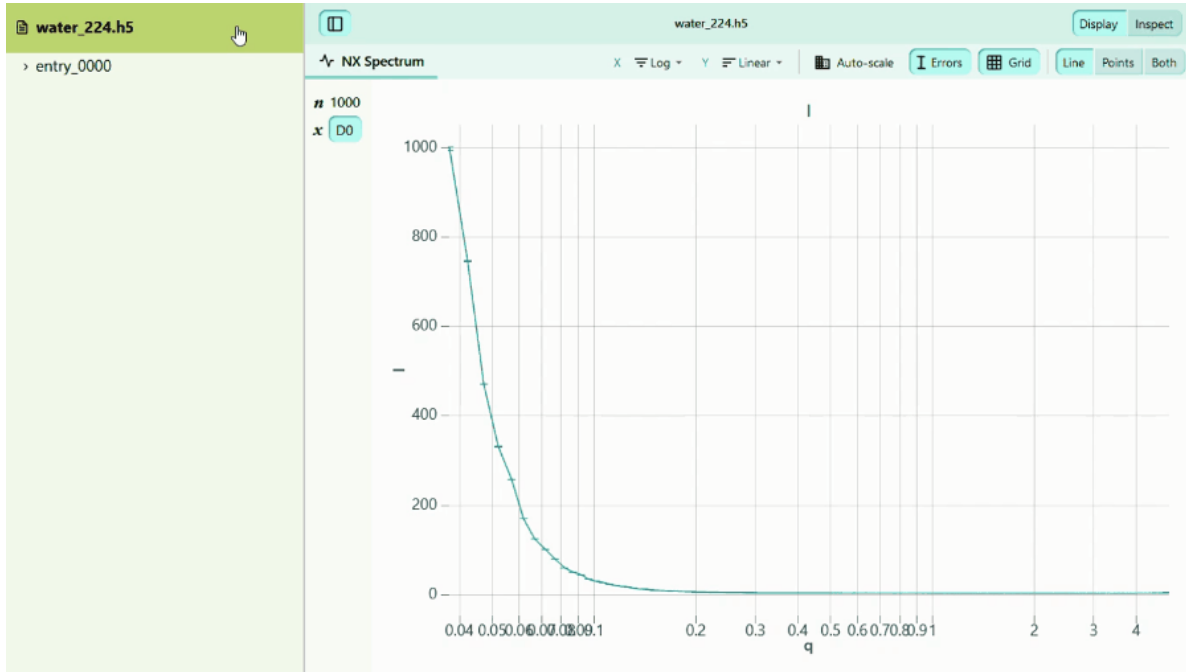
Only raw data with metadata generated by ESRF software will be archived. All raw data necessary to do a complete analysis without redoing the experiment are archived. Processed data can be stored on request.

Clear answer

Answered 21 days ago by admin admin.

DMPs need to be more useful to convince users to use them

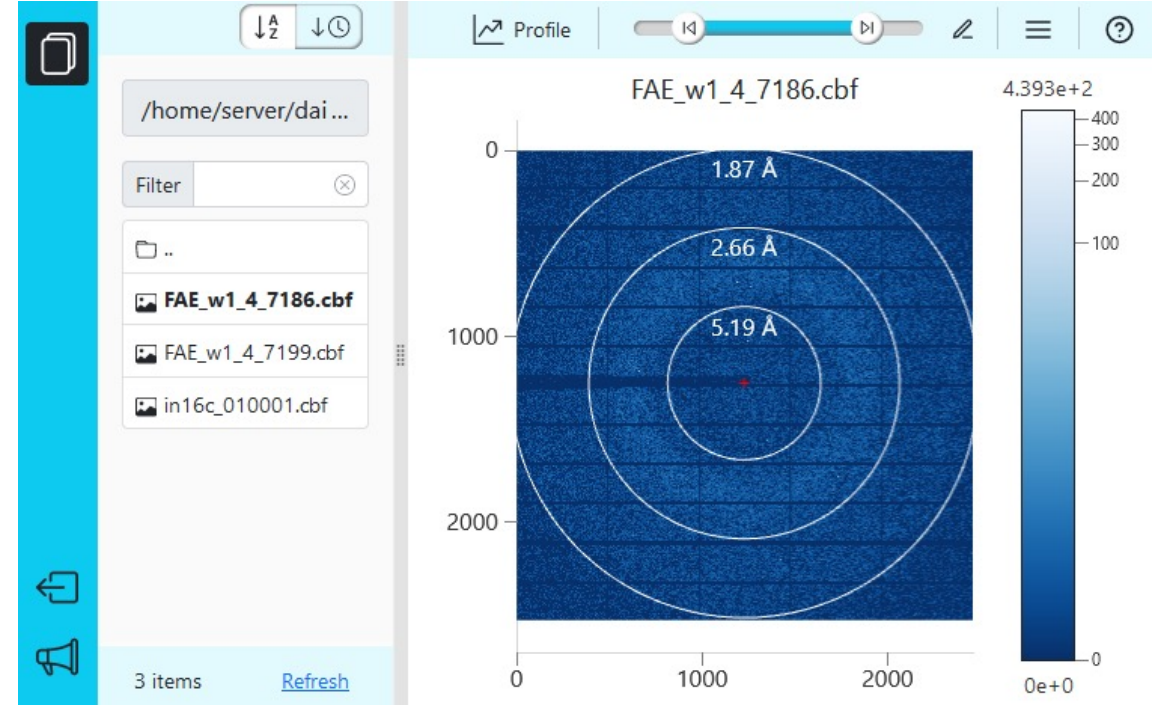
H5Web Visualization Ecosystem



Generic HDF5 file viewer

- Integrated into **ESRF data portal**, for viewing files generated during experiments
- Available as **JupyterLab** and **VS Code extensions**, and soon as part of stand-alone web service, **myHDF5**, for viewing local and hosted HDF5 files

<https://github.com/silx-kit/h5web>



Visualization components

- Used in various web applications at ESRF including:
- **Braggy**, diffraction image viewer (screenshot above)
 - **Daiquiri**, beamline control and data acquisition software

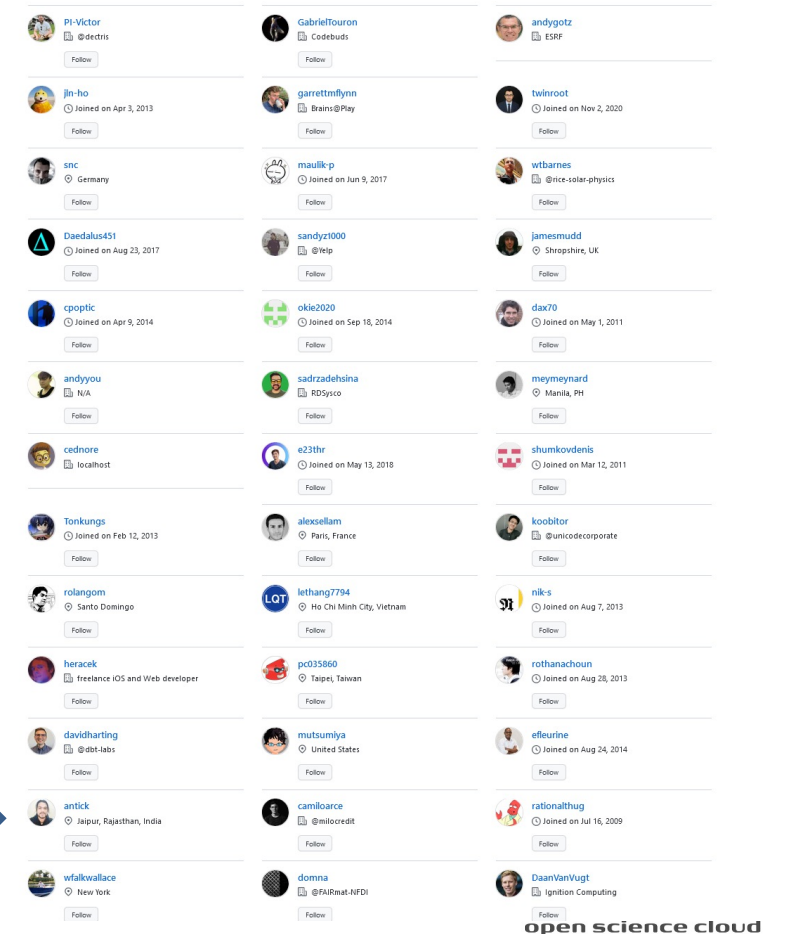
H5Web

H5Web | 1,238 installs | ★★★★★ (3) | Free

H5Web Visualization Ecosystem

- Visualisation in the web is a common requirement for many tools
- H5Web provides a modular solution for plotting in ReactJS applications
- H5Web has been welcomed by many communities e.g. photon + neutron science, neuroscience, astronomy, space, microscopy, materials science, environment, commercial companies ...

GitHub STARGAZERS

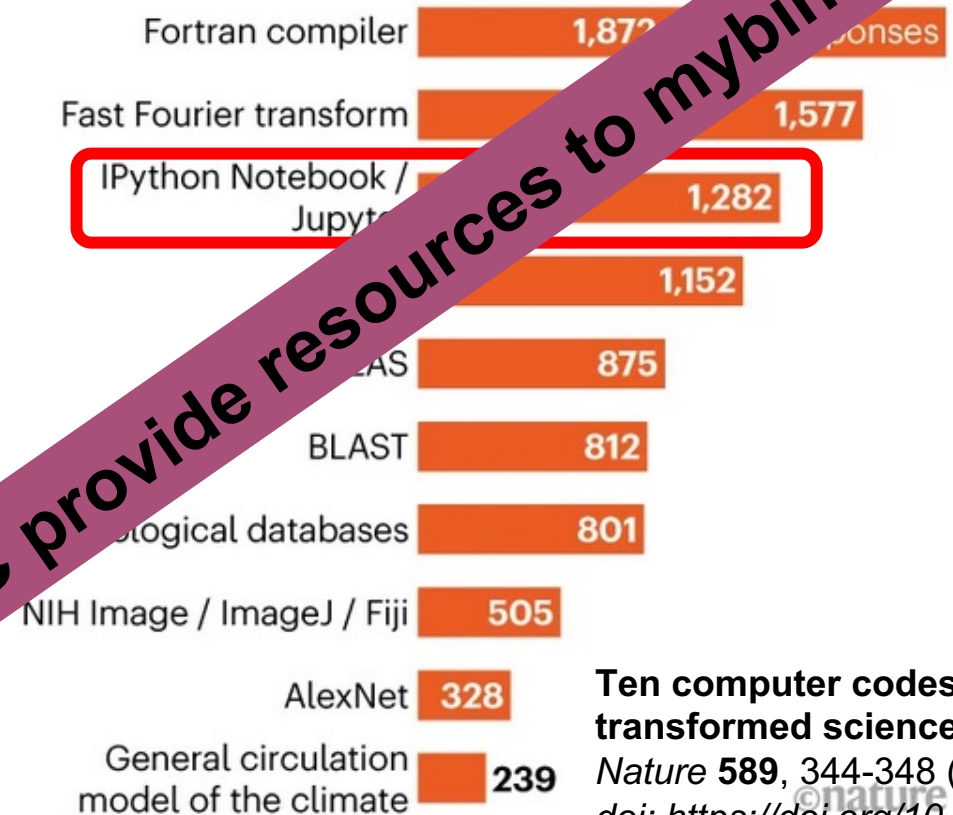


Support for Jupyter notebooks

- **Jupyter service** now available at all PaNOSC and most ExPaNDS sites
- **Jupyter on Slurm** service developed: https://github.com/silx-kit/jupyterhub_moss/
- **H5Web** Jupyterlab plugin developed
- **VISA** provides Jupyter service
- **PaN e-learning** platform provides Jupyter as a service
- **PaNOSC summer school** trained participants to program in Python using Jupyter
- **EGI** provided Jupyter and Binder as a service

TOP CHOICES FOR SCIENCE CODE

Readers voted on which of the ten software codes in this article had the biggest impact on their work. They could choose up to three. Here are the results.



Ten computer codes that transformed science

Nature **589**, 344-348 (2021)

doi: <https://doi.org/10.1038/d41586-021-00075-2>

Example Jupyter service @ <https://jupyter-slurm.esrf.fr>

- Enables users to run Jupyter Notebook on ESRF SLURM cluster

Simple Advanced

Partition

Intel Xeon (x86_64)
Partition: jupyter-nice

IBM Power9 (ppc64le)
Partition: jupyter-p9gpu

CPU

1 core | 2 cores | 4 cores | 10 cores

Options

Jupyter environment: Operating system (default)

Launch JupyterLab:

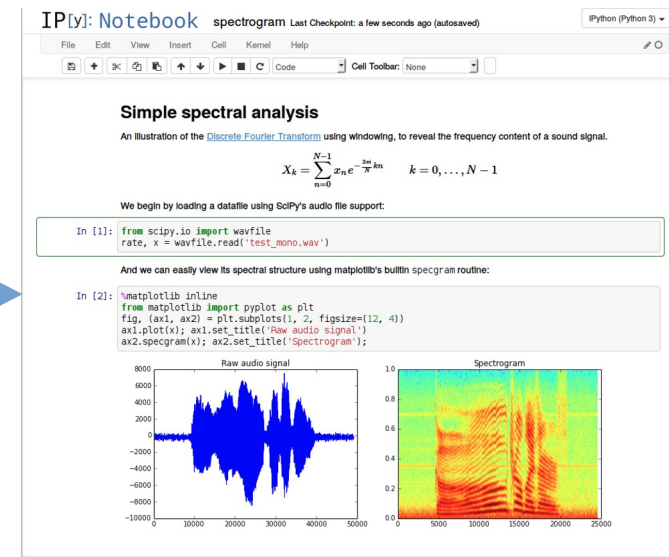
Job duration: 1 hour

List of available resources:

Partition	Current Status	# nodes	# avail
jupyter-nice		28	7
jupyter-p9gpu		8	6

Start

For information, see the [Jupyter @ ESRF documentation](#).



Unique users@ESRF: 156 (monthly average), 276 (total) over 4 months

Open Science with Jupyter notebooks

- Notebooks document
- If used app
- For example
- Notebooks
- Currently, I before they



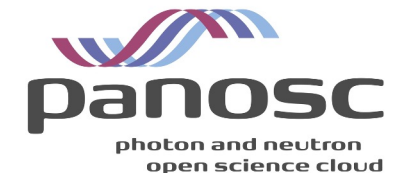
n in one

re-usable

work of others,

EOSC could provide training on making reproducible publications for FAIR data

https://youtu.be/vStbMUDI_jU





VISA - Remote Data Processing/analyses platform



New compute instance

Please fill in the details below to create a new compute instance

Experiments

Select the experiments you wish to associate with your compute instance

Instance not associated to any specific experiments

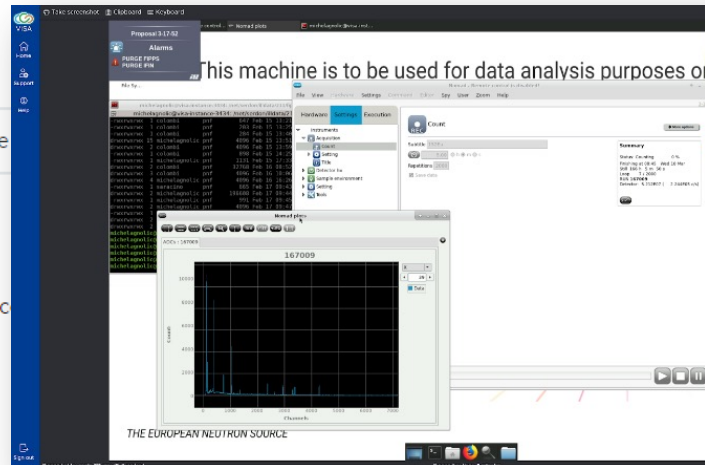
Computing Environment

Choose an environment

 Desktop staging	 Desktop	 Bliss
---------------------	-------------	-----------

Choose hardware requirements

4 Cores 4GB memory esrf.medium	8 Cores 16GB memory esrf.large	16 Cores 32GB memory esrf.gpu.a40	32 Cores 128GB memory esrf.gpu.a40.xlarge
--------------------------------------	--------------------------------------	---	---



Infrastructure for remote data processing / analysis

Users dedicated VM

Access to data

Access to Provisioning of scientific SW using CVMFS and Containers

Access to the GPUs, HPC cluster

Infrastructure based on OpenStack

Development led by ILL in the scope of the PaNOSC project

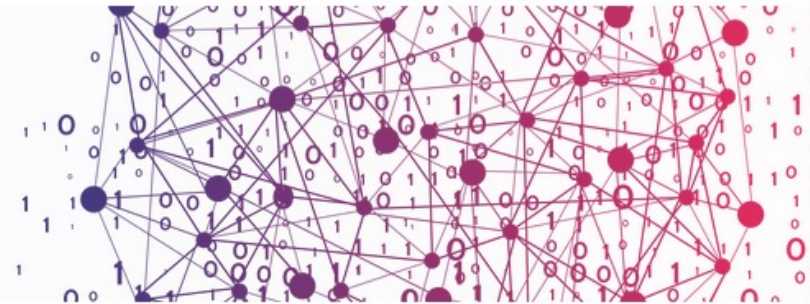


Achieving 100% Open Educational Resources:

1. *Publish training material on pan-training.org*
2. *Develop learning material on pan-learning.org*



Photon and Neutron
eLearninⁿ



Welcome to the e-Learning platform

*This e-Learning platform hosts free education and training for scientists and students.
Below you will find courses on both the theory of photon and neutron scattering
and how to use python code or software for data reduction and modelling.*

<https://e-learning.pan-training.eu/moodle/>

Login

Username



Adoption (in October 2022) of PaNOSC+ExPaNDS outcomes

FACILITY	FAIR data policy	DMPs	DOIs	Nexus HDF5	Search API	Open Data Portal	AAI	Jupyter Lab	VISA	VINYL/OASYS/McStas	Pan-learning/training
ALBA	P	P	WIP	WIP	WIP	WIP	P	Y	WIP	N	U
DESY	WIP	WIP	WIP	Y	WIP	P	WIP	Y	U	Y	WIP
CERIC-ERIC	Y	WIP	Y	WIP	Y	Y	Y	Y	Y	Y	Y
DIAMOND											
ELETTRA	Y	WIP	Y	Y	Y	Y	Y	Y	Y	Y	Y
ESRF	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y
ELI-ERIC	Y	Y	P	Y	Y	Y	WIP	Y	Y	Y	Y
ESS	Y	Y	Y	Y	Y	Y	Y	WIP	WIP	Y	Y
EuXFEL	Y	WIP	Y	WIP	Y	Y	WIP	Y	WIP	Y	Y
FELIX	Y	P	WIP	U	U	WIP	U	U	N	N	U
HZB	Y	P	WIP	Y	P	Y	P	U	U	U	U
HZDR	Y	WIP	Y	N	U	Y	Y	Y	P	WIP	Y
ILL	Y	WIP	Y	Y	WIP	Y	Y	Y	Y	Y	WIP
MAX-IV	WIP	U	Y	Y	Y	Y	Y	Y	U	U	U
PSI	Y	WIP	Y	WIP	Y	Y	WIP	WIP	N	N	N
PTB	Y	WIP	Y	WIP	N	Y	N	N	N	N	N
SOLARIS											
SOLEIL	Y	WIP	WIP	Y	WIP	WIP	Y	WIP	WIP	U	Y
SESAME	Y	U	P	Y	P	WIP	P	P	N	Y	N

Not Planning to be adopted (N)

Under evaluation (U)

open science cloud

In progress of being adopted (WIP)

PaNOSC + ExPaNDS made a major contribution to the latest LEAPS Data Strategy

The screenshot shows a Microsoft Word document with the following content:

LEAPS Data Strategy for 2023- 2030

LEAPS IT Working Group

15 November 2022

1. Introduction

The LEAPS facilities are producers of large volumes of data for science, often referred to as the “data deluge” [1]. The planned upgrades at most of the facilities will see on average a 100 times more photons focussed on the sample. More photons on the sample means shorter data acquisition times and faster experiments. Faster experiments means more data to be acquired, managed and processed so that users

Page 1 of 13 4904 words English (United Kingdom) 11:49 29/11/2022

PaNOSC Project conclusion

Four years of PaNOSC

- Team effort has been essential
- WP leaders have been the driving force
- WP members have been the working force
- Project has achieved all of its KPIs
- Description of Work for the project steered the project
- COVID-19 was a challenge but also an opportunity e.g. for VISA, remote experiments, data portal, remote collaboration etc.
- Collaboration with ExPaNDS was essential, without it the PaN community would not be in a position to continue as a community

PaNOSC most important achievements

1. FAIR data is now a standard to aim for
2. Research Data Management is now considered a part of PaN facilities
3. PaNOSC+ExPaNDS will join forces as the PaNCluster in the EOSC in the future



PaNOSC Closing Event

Paving the way towards the PaN FAIR Data Commons

29-30 November 2022

Grenoble - France

Thank you

andy.gotz@esrf.fr



PaNOSC has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement no. 823852