



PaNOSC + ExPaNDS tools for scientists + science clusters

Andy Götz (ESRF, PaNOSC coordinator)



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.

Talk outline



1. Science clusters tools and platforms
2. PaN Tools for visualization, processing, remote analysis
3. PaN Training platforms
4. PaN DMPs



Science Clusters are providers + consumers of Tools + Platforms

Category	EOSC-Life	ENVRI-FAIR	ESCAPE	PaNOSC	SSHOC
AAI	LS-AAI		IAM & existing X509 infrastructure	UmbrellaId	AARC BPA implementations
Data Repositories	Covid portal FAIRsharing catalog	Certified data centers at RI level ENVRI-Hub	HEP Open Data Portal Virtual observatory Software catalogues	ICAT + Scicat Federated search, Human Organ Atlas Software catalogue,	> 40 certified data centres (CTS)
Metadata standards	Fairsharing.org Bioschemas.org	DCAT-AP	FITS	Nexus/HDF5	CMDI, DDI
Data transfer	Globus, http, rsync	https	FTS (http, gridftp), EUDAT solutions	Globus, http, rsync	http, smtp
Training platforms	TESS	ENVRI Training Catalogue	Summer schools, training material in OSSR	Moodle + TESS	SSH Training Toolkit
Tools used / developed	DS-Wizard Galaxy WorkflowHub.eu	FIP, KB, ENVRI-Hub, Jupyter, VRE	Data Lake, OSSR, VO tools, VRE, Jupyter, etc. Citizen Science platforms	DS-Wizard, Jupyter, H5Web, VISA, VINYL,	SSH Open Marketplace; Virtual Collection Registry; Jupyter





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**



 ESRF DMP


 Users


 Knowledge Models


 Projects


List

Importers

 Documents

 Settings

 Privacy Policy


 Documentation

View

Comments

TODOs

Version history



Current Phase

Before starting the project/proposal

Chapters

I. General / Topic

II. Content classification / Datas

III. Technical classification / Metadata

IV. Ethics / General legal issues

VII. Storage and long-term preservation / Selection

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.

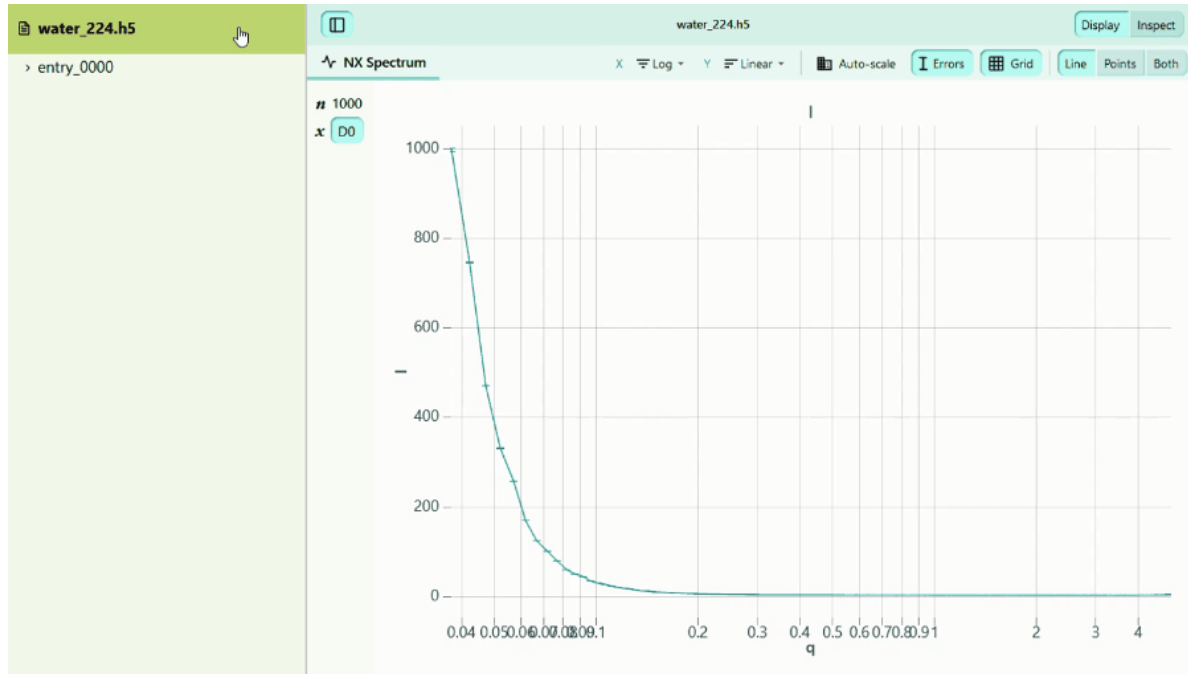
 Clear answer

Answered 21 days ago by admin admin.

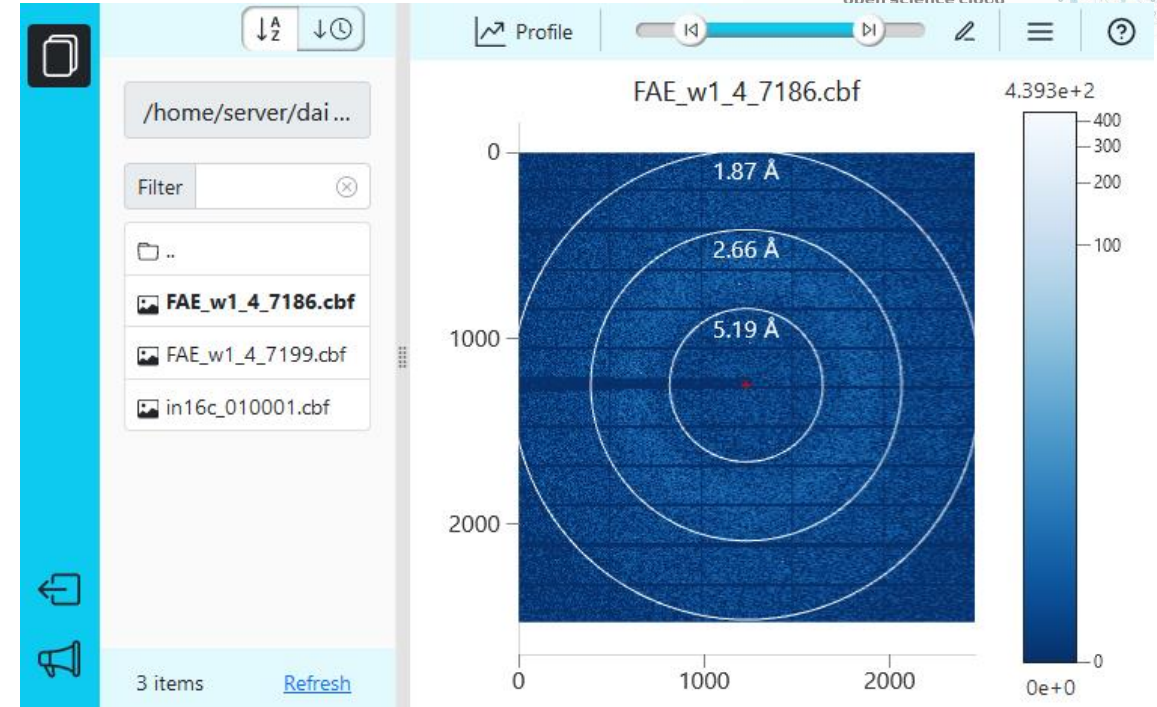
DMPs need to be more useful to convince users to use them

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.

H5Web Visualization Ecosystem



Generic HDF5 file viewer



Visualization components

- 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

Used in various web applications at ESRF including:

- **Braggy**, diffraction image viewer (screenshot above)
- **Daiquiri**, beamline control and data acquisition software

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



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.

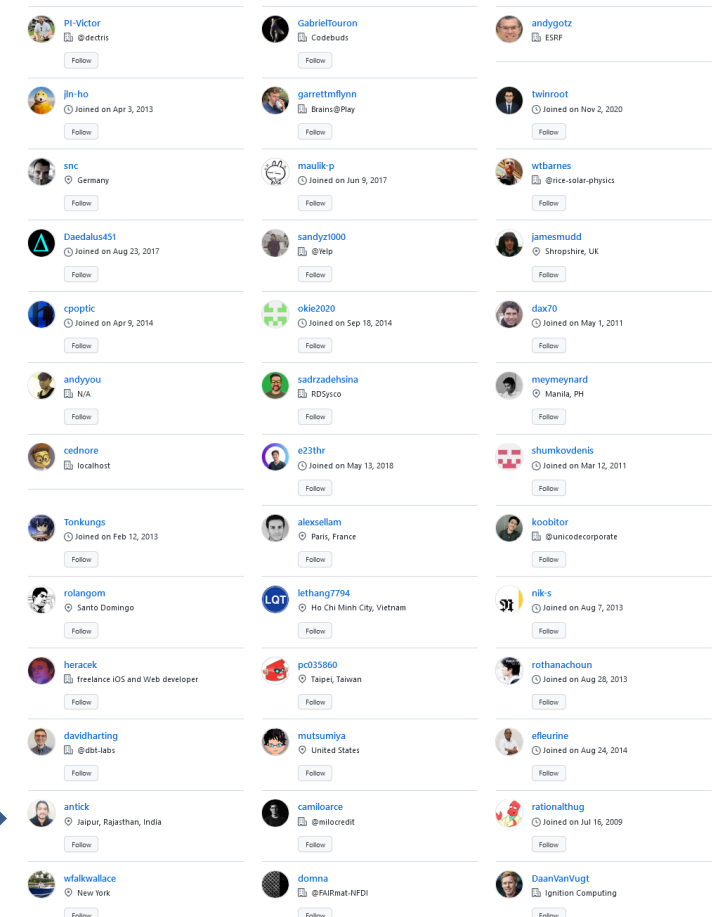
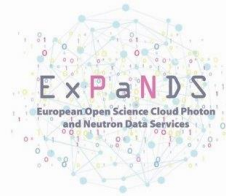
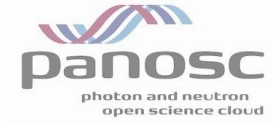
H5Web

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



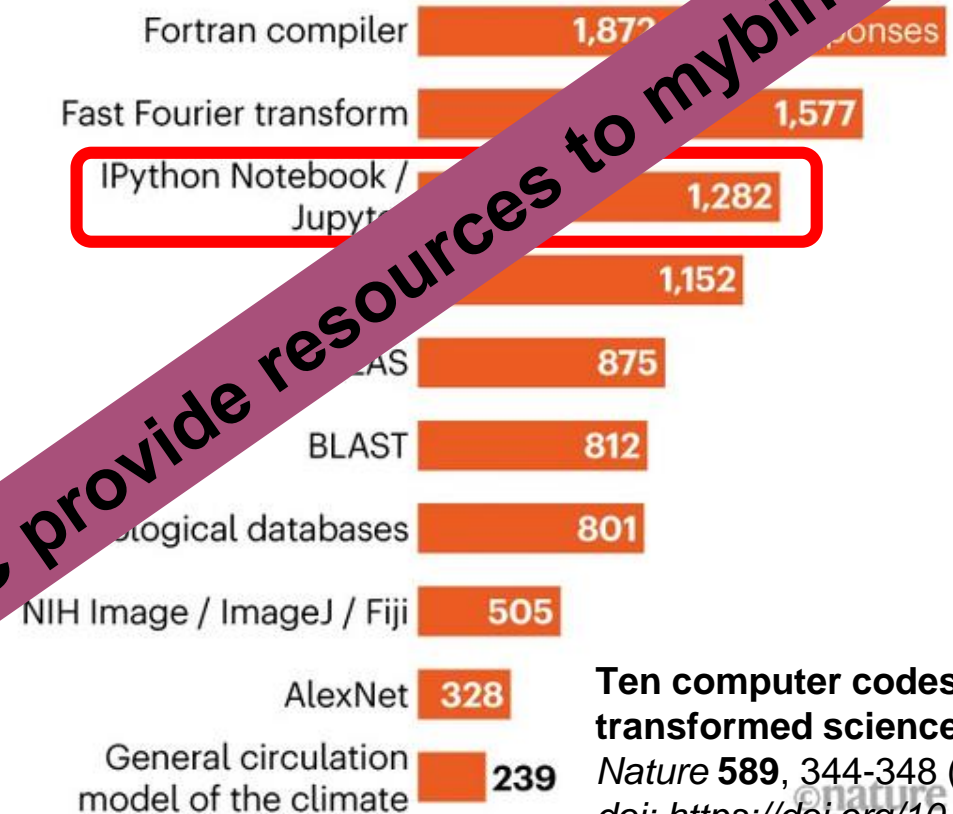
PaNOSC and ExPaNDs projects have received funding from the European Union's Horizon 2020 research and innovation programme under grant agreements 823852 and 857641, respectively.

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's

1 core 2 cores 4 cores 10 cores

Options

Jupyter environment: Operating system (default)

Launch JupyterLab: ☒

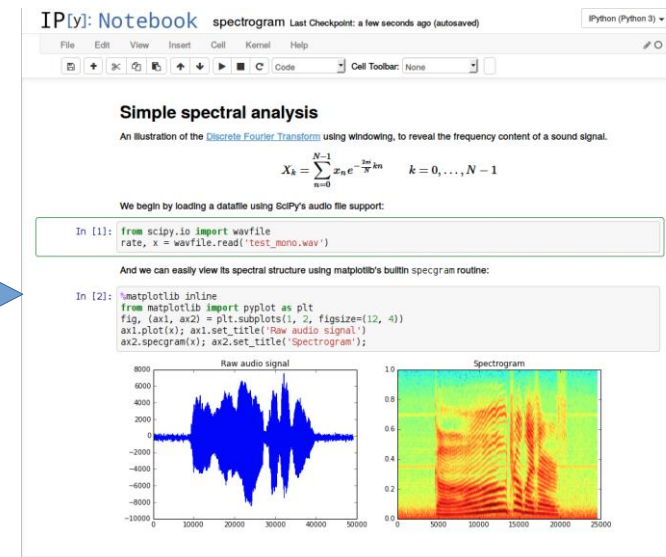
Job duration: 1 hour

List of available resources:

Partition	# 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



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.

VISA - Remote Data Processing/analyses



New compute instance

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

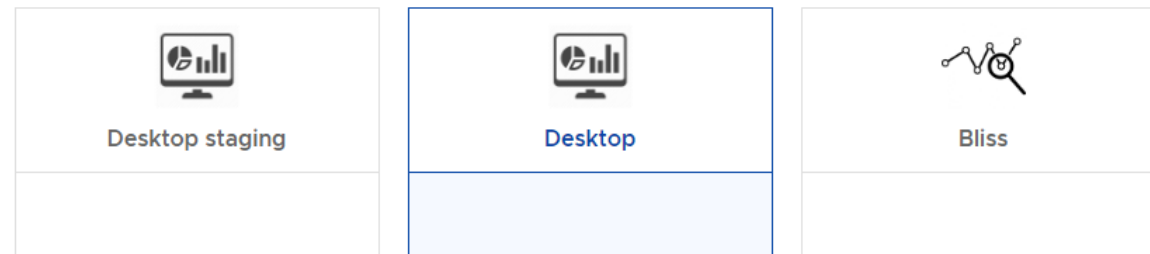
Experiments

Select the experiments you wish to associate with your compute instance

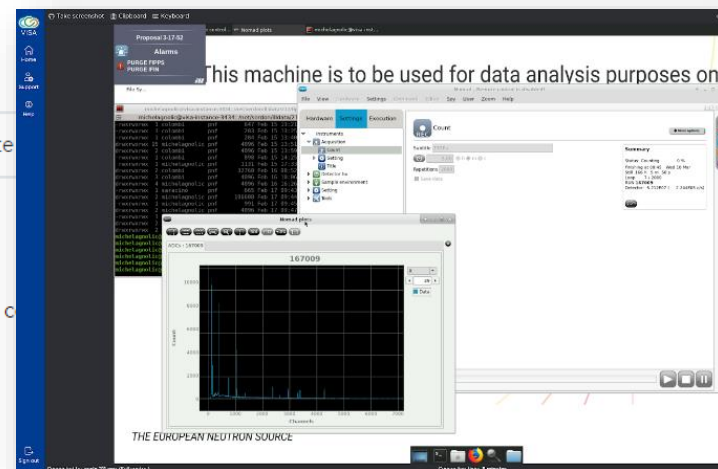
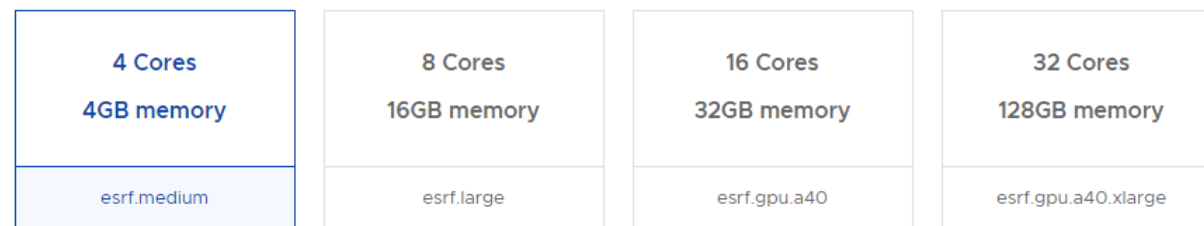
☒ Instance not associated to any specific experiments

Computing Environment

Choose an environment



Choose hardware requirements



Infrastructure for remote data processing / analysis

Users dedicated VM

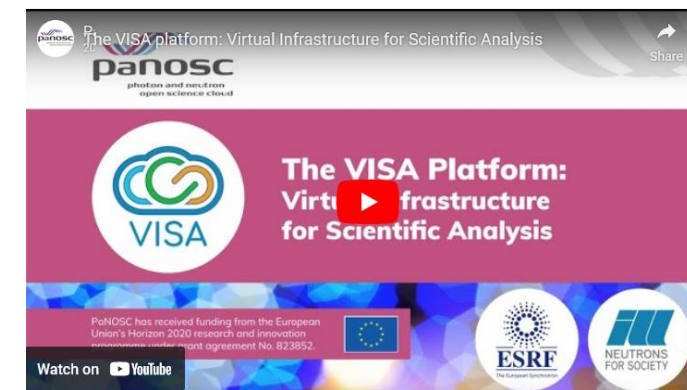
Access to data

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



PaNOSC and ExPaNDS projects have received funding from the European Union's Horizon 2020 research and innovation programme under grant agreements No 823852 and 857641, respectively.



Achieving 100% Open Educational Resources:

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



Photon and Neutron
eLearnin^{n γ}



Welcome to the e-Learning platform

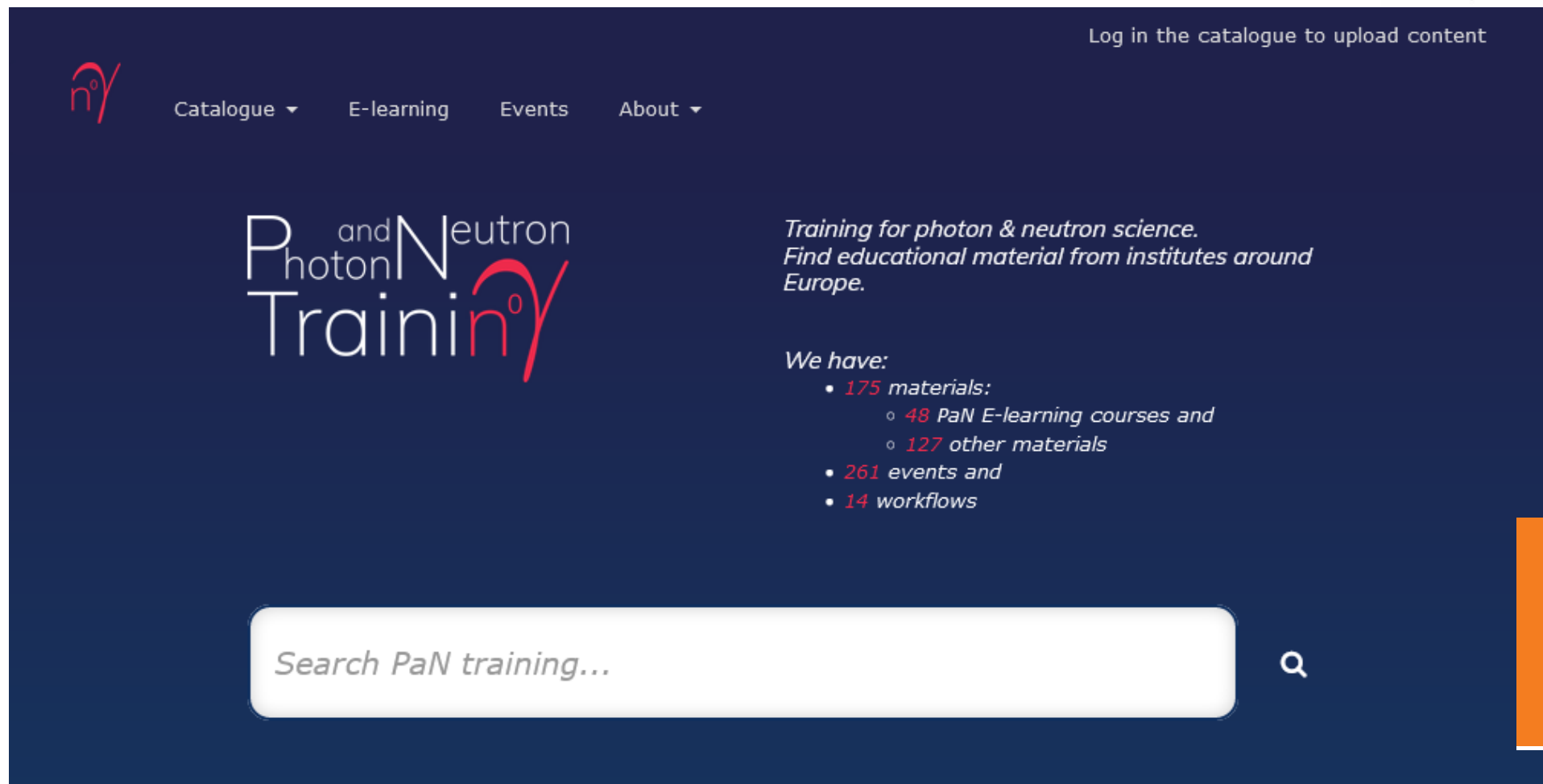
*This e-Learning platform hosts free education and training for scientists and students.
Below you will find courses on both the theory of photon and neutron scattering
and how to use python code or software for data reduction and modelling.*

Login

Username



Share training material, videos, events, etc. on Pan-training.org



The screenshot shows the Pan-training.org website. At the top right, it says "Log in the catalogue to upload content". The navigation bar includes "Catalogue", "E-learning", "Events", and "About". The main header features the "Photon and Neutron Training" logo. Below the header, it states: "Training for photon & neutron science. Find educational material from institutes around Europe." It then lists the available resources: "We have:" followed by a bulleted list: "175 materials:" (with sub-points "48 PaN E-learning courses and" and "127 other materials"), "261 events and", and "14 workflows". At the bottom, there is a search bar with the placeholder text "Search PaN training..." and a magnifying glass icon.

<https://pan-training.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.



Conclusion



- 1. Science clusters are providers + consumers of tools and platforms which need to be supported by the EOSC**
- 2. Software tools make a big difference for scientists**
- 3. Science clusters will actively collaborate around tools in the future**

