



Technical Talk - ESRF

9th of October 2020



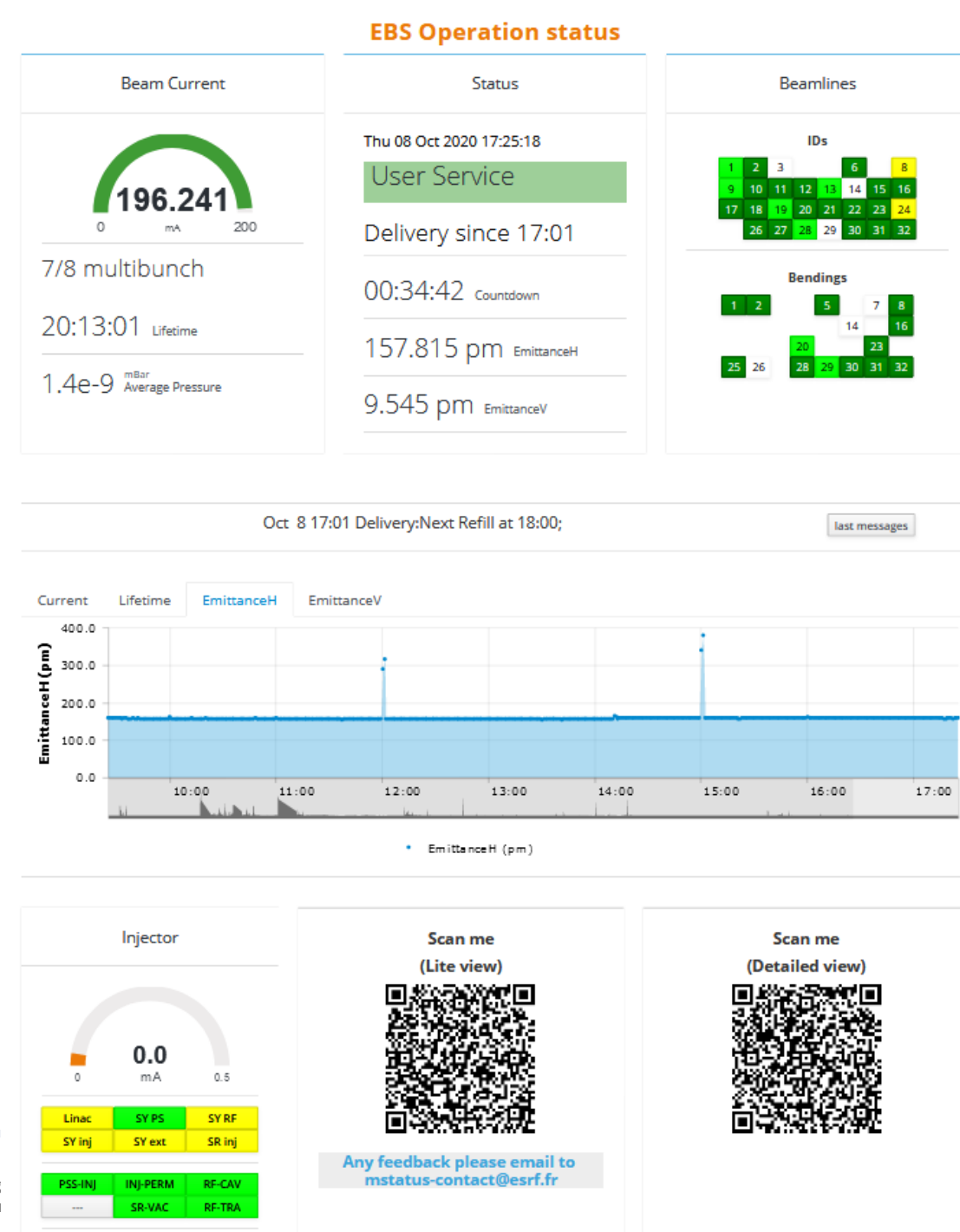
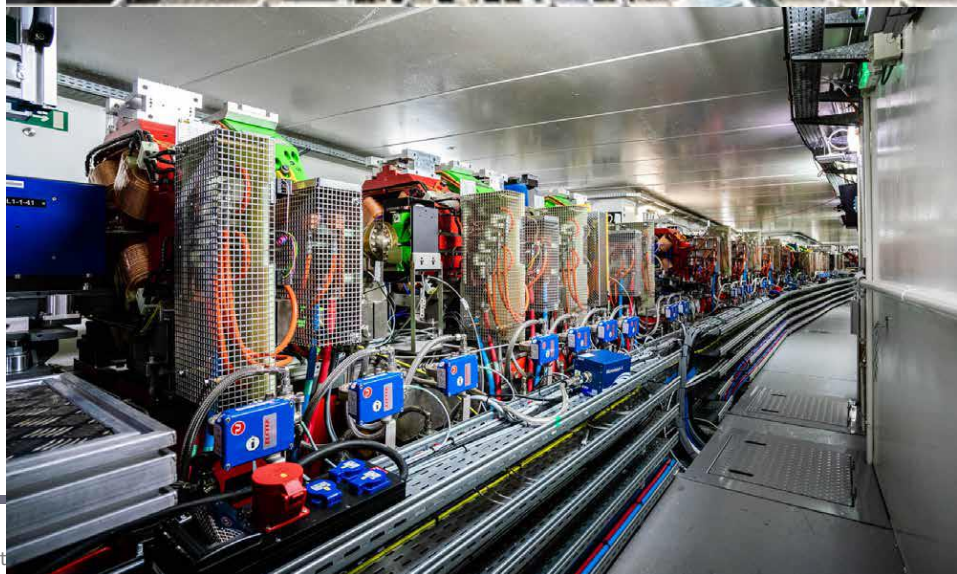
The ExPaNDS project has received funding from the *European Union's Horizon 2020 research and innovation programme* under grant agreement No 857641.
The PaNOSC project has received funding from the *European Union's Horizon 2020 research and innovation programme* under grant agreement No 823852.

Overview

- The Portal Architecture test experience
- COVID and the impact on the team's roadmap
- ESRF Use Cases – what we need and what we offer



ESRF EBS



The ExPaNDS project
The PaNOSC project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement

ESRF PaNOSC Team



Axel Bocciarelli

Software Engineer – Software Group
Expertise in web technologies (React, UX,...)
Will contribute to WP4 + WP5, data portal, Nexus/HDF5 visualization, graphical workflows



Loic Huder

Scientific software Engineer – Software Group
Expertise in visualization, HDF5, Jupyter
Will contribute to WP4 Nexus/HDF5 visualization, Jupyter notebooks



Hire #1 + Hire #2

System administrator – Systems Group
Devops
Will contribute to WP6 + WP4 system infrastructure



Thomas Vincent

Software Engineer – Software Group
Visualisation + Jupyter notebooks
Will contribute to WP4 visualization + coordination



ESRF PaNOSC Team



Alex de Maria

Software Engineer – Software Group
Expertise in data management, web technologies (React, UX, mongo, camel, SQL, ...)

Will contribute to WP3 + WP4, ESRF data portal, data policy implementation



Maxime Chaillet + Hire #3

Software developer –
Software Group
Web developer
Will contribute to WP3 + WP4
e-logbook, DOIs,



Jean-Francois Perrin

Manager – Systems Group
AAI + WP6 leader



Jordi Boderà

Project management – ISDD Office
PaNOSC project manager



Rudolf Dimper

Manager – Directorate
Will contribute to WP2
EOSC Future ESRF lead



Andy Götz

Manager – Software Group
WP2 leader + PaNOSC coordinator



The ExPaNDS project has received funding from the *European Union's Horizon 2020 research and innovation programme* under grant agreement No 857641.
The PaNOSC project has received funding from the *European Union's Horizon 2020 research and innovation programme* under grant agreement No 823852.



The Portal Architecture test experience

- **What went well?**

- We were able to run containers using Docker and Kubernetes on ESRF infra.
- We also had ESRF data analysis software installed and usable.
- We could access all of our data in home directories.

- **What could go better?**

- Supporting all of the different authentication methods and different plugins like icat and scicat was difficult and leads to a configuration mess with too many settings.
- Supporting other synchrotrons on their infrastructure was borderline impossible as we didn't know their network, why they were having problems and how to integrate it into their synchrotron.
- The deployment of the entire portal is still quite difficult with quite a few parts.



Common Portal front-end

ESRF feedback on codebase

- **Modern, popular stack**
 - React + Create React App, styled-components (CSS-in-JS library), Rebass (CSS framework)
 - Prettier + ESLint for code quality
- **React best practices**
 - Small functional components, one component per file, React hooks, error boundaries, authentication via provider, etc.
- **Great code quality overall**
- **Easy to install and start**



Common Portal front-end – ESRF feedback on Features

- **Search is crucial**
 - Fully on board with idea of filtering sidebar like on e-commerce sites
 - Need to carefully consider each filter (user need + usability)
 - Filters can be added incrementally as user needs are discovered
- **Drag-and-drop UI**
 - Works great, but need to double check WCAG accessibility and discoverability (for keyboard and screen reader users)
 - May be worth adding a simpler UI (e.g. drop-down menu of available environments on each search result) to complement the drag-and-drop interface
- **Dashboard**
 - Nice to have but not necessary; search can remain the main screen for now



ESRF needs for the Portal

- What are the main use cases for deploying the portal @ ESRF
 1. **Link to the ESRF data portal** <https://data.esrf.fr>
 2. **Browse domain specific data similar to** <https://paleo.esrf.eu>
 3. **Provide an access to local Jupyter-slurm notebook service**
 4. **Provide access to standard ESRF data analysis programs**
 5. **Link software catalogue to data analysis portal**
 6. **Provide a remote desktop integrated experience**
 7. **Search for data across domains at ESRF and all PaNOSC sites**
 8. **Launch standard notebooks for standard experiments**
 9. **Expose COVID-19 and other data in a user friendly browser**
 10. **Link data portal to data management KPIs e.g. processing with AI/ML**



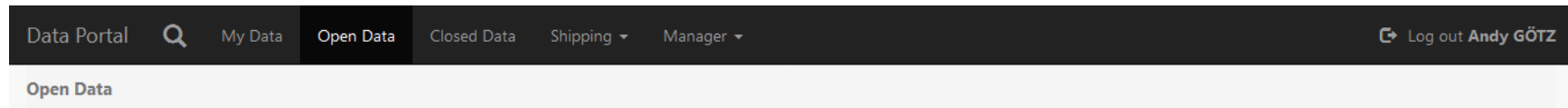
ESRF data portal – <https://data.esrf.fr>

Data policy implemented on
all ESRF beamlines in 2020

Use Case: integrate
<https://data.esrf.fr>

Users: external users

Expect: common UX



Open Data Collections				
Search				
Proposal	Beamline	Title	Datasets	DOI
OPID-1	ID30A1	Structural Evidence for a Role of the Multi-functional Human Glycoprotein Afamin in Wnt Transport	11	DOI 10.15151/ESRF-DC-142893590
OPID-1	ID30A-1	Controlled dehydration, structural flexibility, and Gadolinium MRI contrast compound binding in human plasma glycoprotein afamin	29	DOI 10.15151/ESRF-DC-142915526
OPID-1	ID30A1	Gadolinium MRI contrast compound binding in human plasma glycoprotein afamin	11	DOI 10.15151/ESRF-DC-186857652
MX-2005	CM01	Biogenesis and structure of a type VI secretion baseplate	8	DOI 10.15151/ESRF-DC-186877747
OPID-1	ID30A1	Raw diffraction images for apo murine Fumarylacetoacetate Domain Containing Protein 1 (FAHD1)	9	DOI 10.15151/ESRF-DC-186917546
OPID-1	ID30A1	Raw diffraction images for oxalate bound murine Fumarylacetoacetate Domain Containing Protein 1 (FAHD1)	6	DOI 10.15151/ESRF-DC-186917546

Data Portal

Sign in

ESRF, CRG staff or long term visitors
Please sign in with ESRF SSO

I am new to the portal
[Create a new account](#)

I need further assistance

Important note

During 2019 and according to the General Data Protection Regulation, all portal users who did not consent to the [User Portal Privacy Statement](#) have had their account deactivated. Please contact the [User Office](#) if you wish to reactivate it.

SSO Database

Sign in with ESRF SSO

or sign in as anonymous

ESRF

Data collection

DOI > 10.15151/ESRF-DC-142893590

STRUCTURAL EVIDENCE FOR A ROLE OF THE MULTI-FUNCTIONAL HUMAN GLYCOPROTEIN AFAMIN IN WNT TRANSPORT

Andreas Naschberger ; Matthew W. Bowler ; Bernhard Rupp.

Abstract

Afamin, a human plasma glycoprotein and putative transporter of hydrophobic molecules, has been shown to act as extracellular chaperone for poorly soluble, acylated Wnt proteins, forming a stable, soluble complex with functioning Wnt proteins. The 2.1-Å crystal structure of glycosylated human afamin reveals an almost exclusively hydrophobic binding cleft capable of harboring large hydrophobic moieties. Lipid analysis confirms the presence of lipids, and density in the primary binding pocket of afamin was modeled as palmitoleic acid, presenting the native O-acylation on serine 209 in human Wnt3a. The modeled complex between the experimental afamin structure and a Wnt3a homology model based on the XWnt3-Fz8-CRD fragment complex crystal structure is compelling, with favorable interactions comparable with the crystal structure complex. Afamin readily accommodates the conserved palmitoylated serine 209 of Wnt3a, providing a structural basis how afamin solubilizes hydrophobic and poorly soluble Wnt proteins.

Proposals

Beamlines

Publication year

OPID-1

ID30A1

2018

Experimental report

There is currently no experimental report.

Experimental data

The data can be accessed by clicking on the link below

The ExPaNDS project has received funding from the *European Union's Horizon 2020 research and innovation* project.

The PaNOSC project has received funding from the *European Union's Horizon 2020 research and innovation* project.

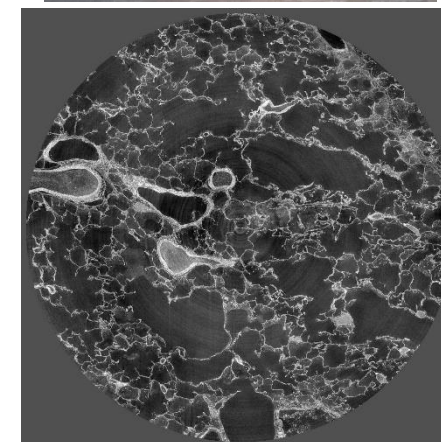
ESRF domain specific data portal(s)

ESRF heritage database for palaeontology, evolutionary biology and archaeology

By ESRF

Please cite the original articles linked to the data you are using, as well as the repository institutions. CC BY-NC-SA Attribution-NonCommercial-ShareAlike

The screenshot shows the 'Home' page of the ESRF domain specific data portal. On the left, there is a sidebar with an 'Albums' section listing categories like paleoanthropology, invertebrate paleontology, vertebrate paleontology, vertebrate biology, paleobotany, ichnology, and Archaeology. Below this is an 'Identification' section with 'Register' and 'Login' links, and a 'Quick connect' section with fields for 'Username' and 'Password', an 'Auto login' checkbox, and a 'Submit' button. The main content area displays a grid of category tiles, each with a representative image and text indicating the number of photos and sub-albums. The categories shown are paleoanthropology (25 photos in 32 sub-albums), invertebrate paleontology (38 photos in 49 sub-albums), vertebrate paleontology (48 photos in 57 sub-albums), vertebrate biology (67 photos in 82 sub-albums), paleobotany (2 photos in 2 sub-albums), ichnology (2 photos in 3 sub-albums), and Archaeology.



Use Case: provide user friendly way to browse and search for paleontology + COVID-19 data

Users: paleontologists, biologists, medical

Expect: user friendly browser for domain specific categories e.g. paleontology linked to ESRF DOIs



The ExPaNDS project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No 857641.
The PaNOSC project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No 823852.

Page 11

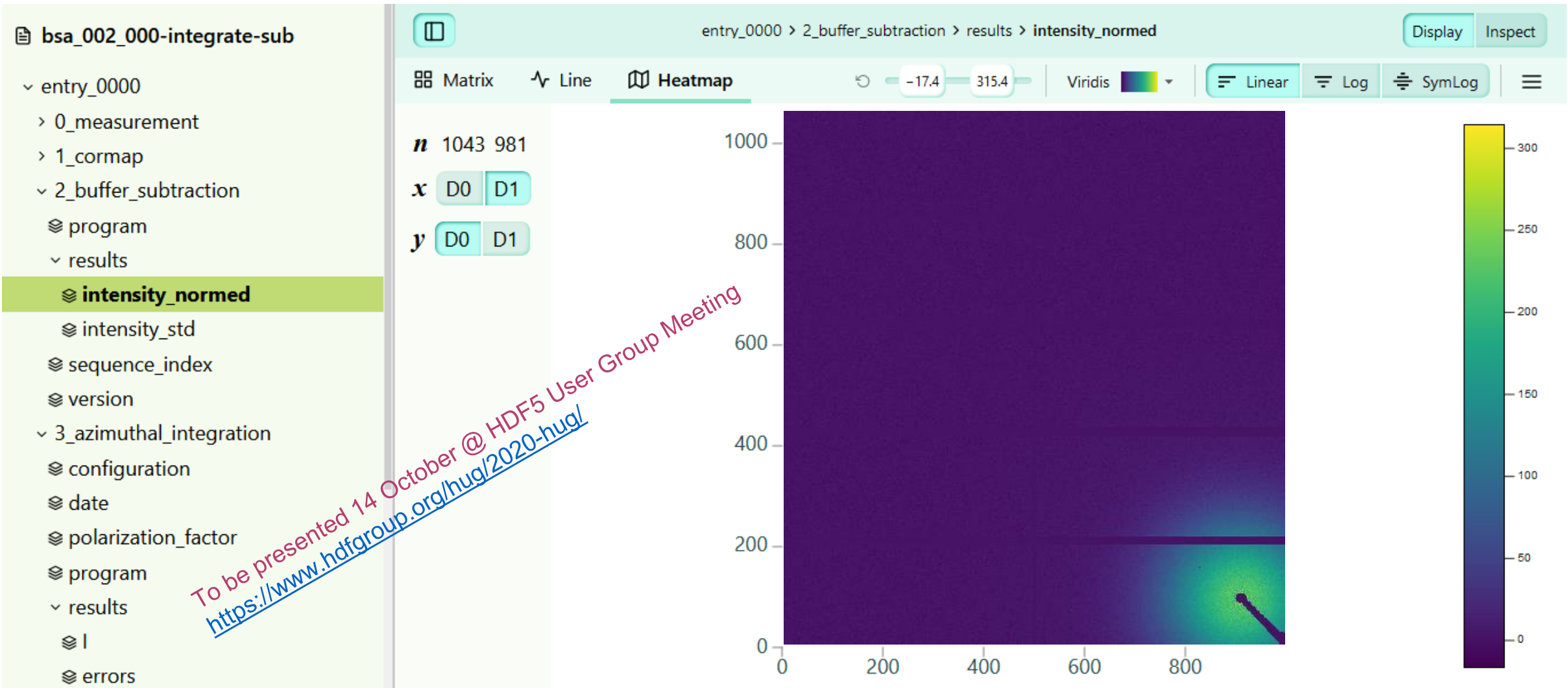


What ESRF contributes

- High performance **web visualization** of **Nexus/HDF5 (h5web)**
- Integration of visualization of **HDF5** in **Jupyterlab** as 1st class citizen
- **Graphical** visualization of **workflows** in the **web**
- **ICAT integration** and **extensions**
 - Metadata definitions
 - E-logbook
 - Search



H5web visualisation - <https://h5web.panosc.eu>




The ExPaNDS project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No 857641.
The PaNOSC project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No 823852.

Contacts: Loic + Axel









H5web storybook - <https://h5web-docs.panosc.eu>

 Storybook

...

Canvas Docs

Q Press "/" to search...

Home

VISUALIZATIONS

HeatmapVis

Default

Domain

Color Map

Log Scale

Sym Log Scale

Ignore Aspect Ratio

No Grid

Live Data Without Loader

LineVis

ScalarVis

BUILDING BLOCKS

AxisSystem

ColorBar

H5Web Component Library

H5Web is a web-based viewer to explore HDF5 files. It is built with React and uses [react-three-fiber](#) for visualizations.

This library provides data visualization components and utilities from H5Web for use in other front-end web applications.

The components are organised in two categories:

- 🧠 **Visualizations**: the top-level visualization components — i.e. the components you will most likely need.
- 🧱 **Building Blocks**: the low-level components used by the visualization components — for advanced uses only.

Getting started

- Make you sure you have `react` and `react-dom` v16 or greater installed.
- Install the library:

```
npm install @h5web/lib
```

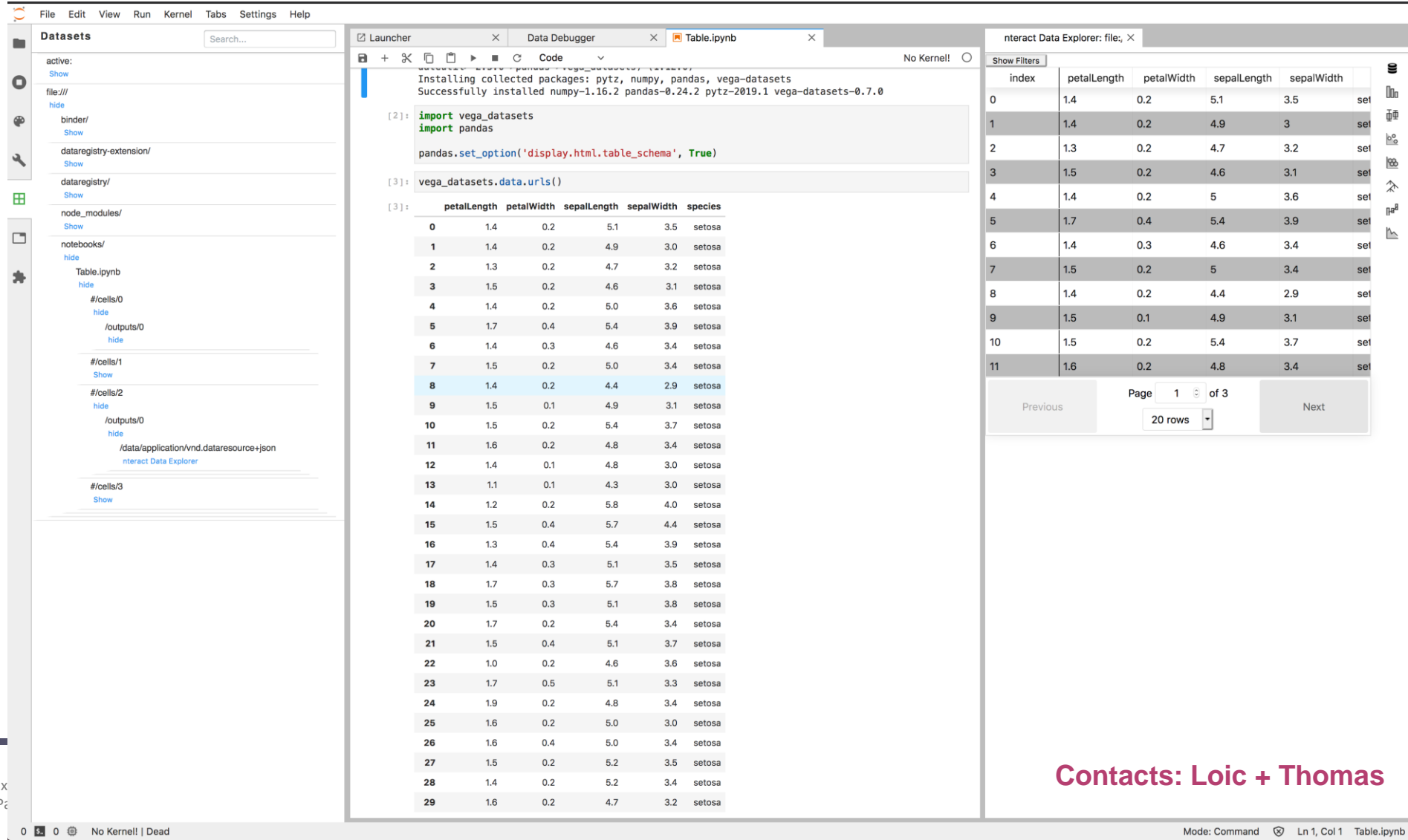
Copy
- Import and use a visualization component:

```
import { ScalarVis } from '@h5web/lib';  
// ...
```

Contacts: Loic + Axel

H5web integration in Jupyterlab – Nexus/HDF5 as 1st class citizen

<https://github.com/jupyterlab/jupyterlab-data-explorer>



The screenshot shows the JupyterLab interface with the following components:

- Left Sidebar (Datasets):** A search bar and a list of datasets including 'active', 'file:///binder/', 'dataregistry-extension/', 'dataregistry/', 'node_modules/', and 'notebooks/'.
- Main Area (Code Editor):** A code editor showing the following Python code:

```
Installing collected packages: pytz, numpy, pandas, vega-datasets
Successfully installed numpy-1.16.2 pandas-0.24.2 pytz-2019.1 vega-datasets-0.7.0

[2]: import vega_datasets
import pandas

pandas.set_option('display.html.table_schema', True)

[3]: vega_datasets.data.urls()

[3]:
```

	petalLength	petalWidth	sepalLength	sepalWidth	species
0	1.4	0.2	5.1	3.5	setosa
1	1.4	0.2	4.9	3.0	setosa
2	1.3	0.2	4.7	3.2	setosa
3	1.5	0.2	4.6	3.1	setosa
4	1.4	0.2	5.0	3.6	setosa
5	1.7	0.4	5.4	3.9	setosa
6	1.4	0.3	4.6	3.4	setosa
7	1.5	0.2	5.0	3.4	setosa
8	1.4	0.2	4.4	2.9	setosa
9	1.5	0.1	4.9	3.1	setosa
10	1.5	0.2	5.4	3.7	setosa
11	1.6	0.2	4.8	3.4	setosa
12	1.4	0.1	4.8	3.0	setosa
13	1.1	0.1	4.3	3.0	setosa
14	1.2	0.2	5.8	4.0	setosa
15	1.5	0.4	5.7	4.4	setosa
16	1.3	0.4	5.4	3.9	setosa
17	1.4	0.3	5.1	3.5	setosa
18	1.7	0.3	5.7	3.8	setosa
19	1.5	0.3	5.1	3.8	setosa
20	1.7	0.2	5.4	3.4	setosa
21	1.5	0.4	5.1	3.7	setosa
22	1.0	0.2	4.6	3.6	setosa
23	1.7	0.5	5.1	3.3	setosa
24	1.9	0.2	4.8	3.4	setosa
25	1.6	0.2	5.0	3.0	setosa
26	1.6	0.4	5.0	3.4	setosa
27	1.5	0.2	5.2	3.5	setosa
28	1.4	0.2	5.2	3.4	setosa
29	1.6	0.2	4.7	3.2	setosa
- Right Sidebar (Interact Data Explorer):** A panel titled 'Interact Data Explorer: file, X' showing a table of data with columns 'index', 'petalLength', 'petalWidth', 'sepalLength', 'sepalWidth', and 'species'. The table contains 11 rows of data. Below the table, there are navigation buttons: 'Previous', 'Page 1 of 3', 'Next', and a dropdown menu showing '20 rows'.

Contacts: Loic + Thomas

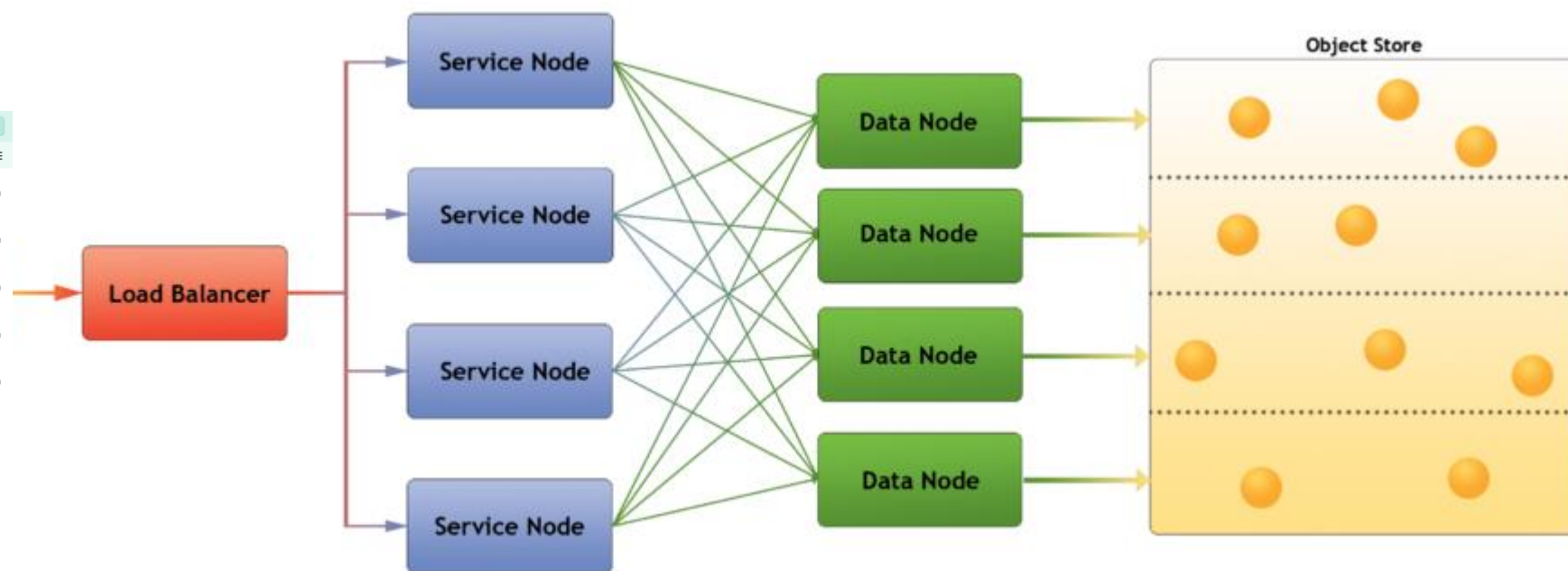
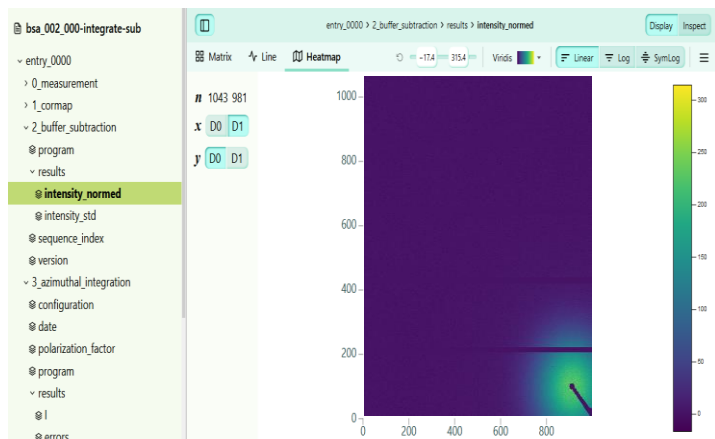
<https://github.com/jupyterlab/jupyterlab-hdf5>



The Ex
The Pa



H5web integration in HSDS



<https://www.hdfgroup.org/solutions/highly-scalable-data-service-hsds/>

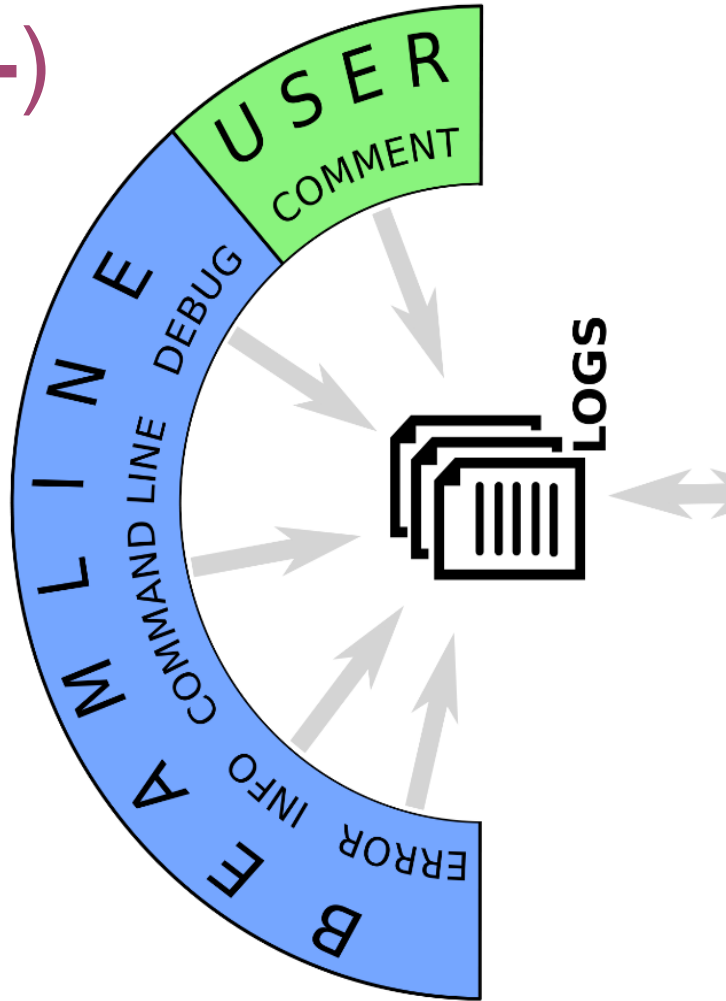


The ExPaNDS project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No 857641. The PaNOSC project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No 823852.

Contacts: Loic



E-logbook makes experiments FAIRer (implemented in ICAT+)



Web interface

Contacts: Maxime + Alex

Web interface screenshot showing a log of commands and data collection results. The interface includes a top navigation bar with buttons: + New, Take a photo, View, PDF, and search filters (Everywhere). The log entries are as follows:

Time	Command/Event	Annotation
08:03:00	OPTICS> # io/optics/figa def measure1 '_ccd_set_concat(1)'	
08:01:40	OPTICS> snap flux	→ COMMENT
08:01:25	OPTICS> dt	→ COMMAND LINE
03:46:39	OPTICS> New dataset: cchof_root	→ INFO
01:20:21	OPTICS> zapxiainage thg 6.65 10.69 10 vbg 54.8 27.6 27 0 10 (zapug: #6, spec: #3)	
01:20:20	OPTICS> New dataset: cFeo42-_root2	
01:19:51	OPTICS> prdef Maps_554	
00:54:47	no new data collected	→ ERROR
00:51:57	OPTICS> dt	



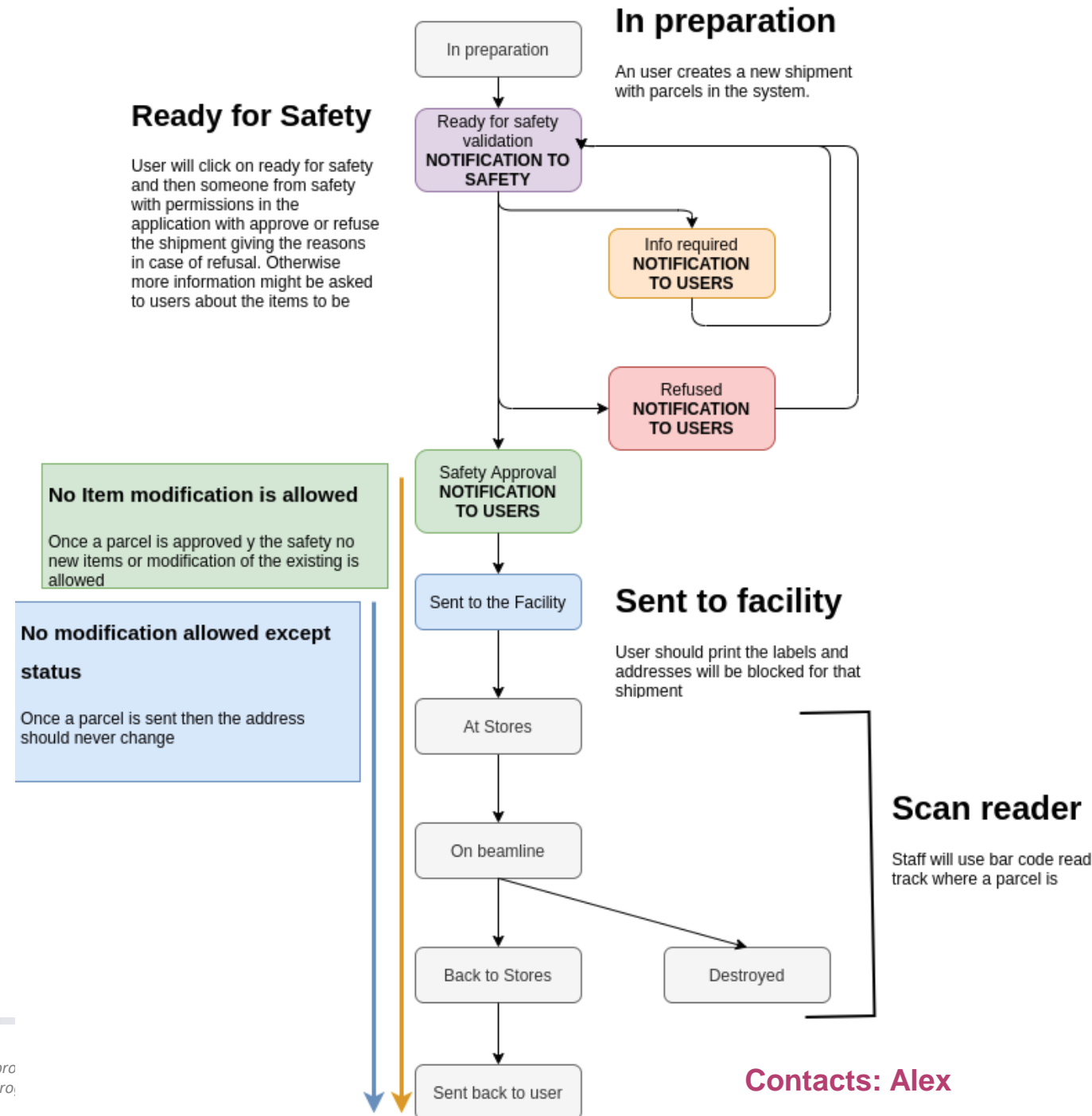
Impact of COVID-19 @ ESRF

- All **experiments** are **remote** now
- **Sample tracking** for all beamlines was developed
- Remote analysis via **Jupyter-slurm**
- **Remote access** of experiments via **Guacamole**
- **Downloading of data** has become a **critical** issue → **globus+ACLs**




Sample tracking for all Experiments

- ~10 people involved
- 3 months development
- Used for all remote experiments since 25/8/2020
- **121 Shipments + 672 Samples in 1 month**
- Implemented in ICAT+



ESRF remote experiment access – Guacamole remote desktop

RECENT CONNECTIONSgoetz ▾


rnice9-0308

ALL CONNECTIONSFilter

- Compute cluster
 - gp1d11-nice
 - rnice
 - rnice9-0308
- Teleworking

Use Case: provide remote desktop access for remote experiments

Users: beamline scientists and external users

Expect: VISA would be nice to have



Scientific Use Case

- Motivation

“... improve the efficiency of the experiment (many different users, a possible regular beamtime, remote experiment),

I think it would help a lot to have a dedicated Jupyter notebook for data processing. It would be a good way i) to save time, ii) to record which map has already been processed and how, iii) and have a coherent data processing between users and between sessions, and also for possible users who would like to reuse data after the embargo period.” Scientist

- **Use Case:** provide users with post azimuthal integration maps in hdf5, with Jupyter notebooks taking care of PyFAI integration of 2D maps.
- **Users:** Scientists doing experiments in cultural heritage
- **Contacts:** Loic + Scientist

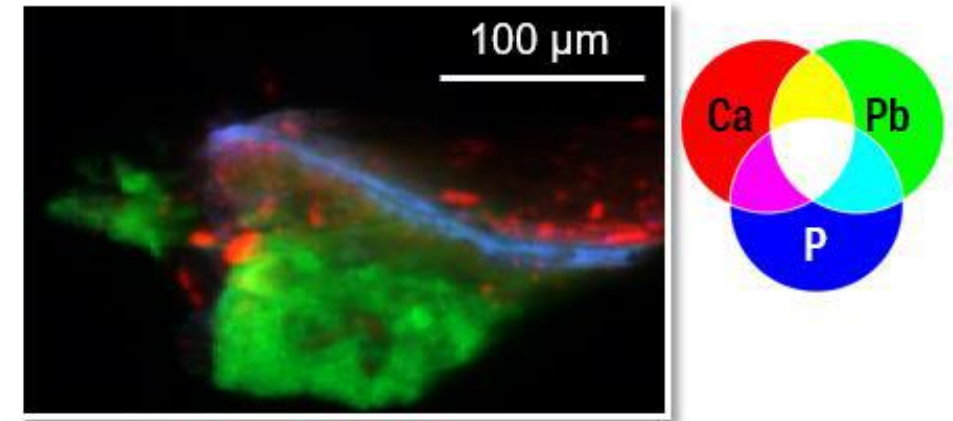
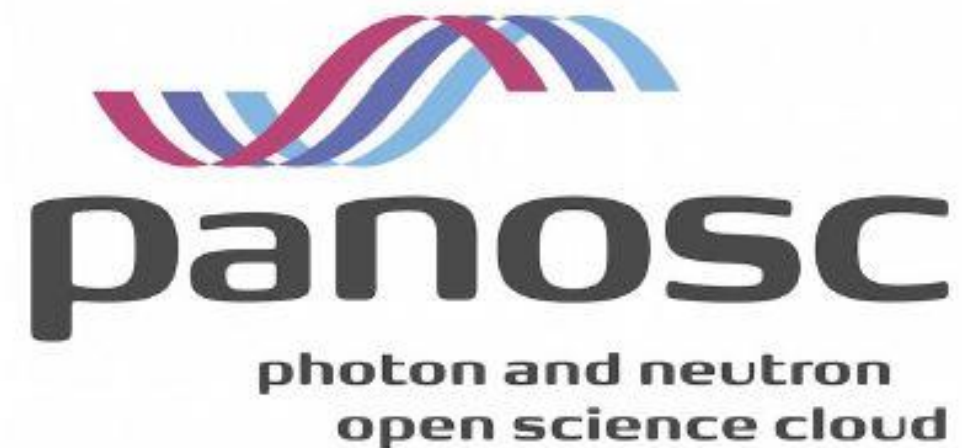


Fig.1 : SR-μ-XRF map (ID21) revealing the presence of a lead phosphate compound in the stratigraphy of a sample from a Leonardo da Vinci painting. SR-μXRD (ID13) analysis revealed the compound to be phosphohedyphane,



The ExPaNDS project has received funding from the *European Union's Horizon 2020 research and innovation programme* under grant agreement No 857641.
The PaNOSC project has received funding from the *European Union's Horizon 2020 research and innovation programme* under grant agreement No 823852.

