



# Welcome to the PaNOSC closing event

Vincent Favre-Nicolin

ESRF-Experiments Division  
Algorithms & scientific Data Analysis group



| The European Synchrotron

# EOSC DREAMS...



**EUROPEAN OPEN  
SCIENCE CLOUD**

- Ubiquitous access to data and compute resources
- Standardise remote access to compute resources
- Open data for open science

... We do not (yet?) get *all* of that, because:

- (large) data transfers are not magical
- Lack of economical model for the resources

But we now have all essential tools/bricks available

# OPEN SCIENCE, OPEN DATA



« *non-reproducible single occurrences are of no significance to science* »

Karl Popper

- The current period is less enlightened than most of us hoped from the *open information* era of the internet: "alternate facts", "filter bubbles", etc...
- We can't solve all those problems...
- But we have a duty to increase the trust in scientific outputs
- And the cornerstone for that is open - reliable & FAIR- data

# OPEN SCIENCE, OPEN DATA

Data Portal My Data Open Data Closed Data Shipping My Beamlines Manager Feedback Log out Vincent FAVRE NICOLI

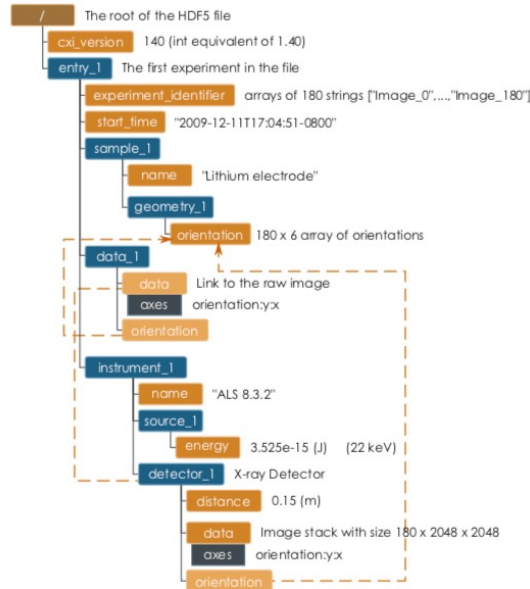
My Data

Filter between dates Filter between dates Open Jupyter Notebook Search

Proposal	Beamline	Start	Title	A-Form	Datasets	Files	Release	DOI
BLC-13866	ID16B	14/06/2022	Test of inline reconstruction	<a href="#">BLC-13866</a>				
MI-1352	ID01	09/12/2020	Magnetic Coherent Diffraction Imaging with hard X-rays	<a href="#">MI-1352</a>	8	1.7 GB	13/12/2023	<a href="#">DOI 10.1515/ESRF-ES-119464351</a>
HC-4050	ID01	04/11/2020	In situ and operando coherent diffraction imaging of nanostructures during chemical reactions; towards atomic resolution and time-resolved e	<a href="#">HC-4050</a>	2	1.3 GB	07/11/2023	<a href="#">DOI 10.1515/ESRF-ES-152611353</a>
BLC-12258	ID01	22/07/2020	ID01 KB alignment + test experiments	<a href="#">BLC-12258</a>				
IH-MA-69	ID01	18/11/2018	Coherent diffraction imaging of LNMO batteries, Si reference structures, Co microcrystals and Ge microdisks				20/11/2021	<a href="#">DOI 10.1515/ESRF-ES-119464351</a>
HC-3800	ID01	10/09/2018	Strain imaging in suspended GeSn micro-Bridges for laser application using multi-angle Bragg projection ptychography		0	0 Bytes	14/09/2021	<a href="#">DOI 10.1515/ESRF-ES-100129017</a>
MI-1328	ID16A	08/05/2018	High resolution, high throughput pink beam far field Ptychography		209	9.1 MB	11/05/2021	<a href="#">DOI 10.1515/ESRF-ES-91421585</a>
MA-3864	ID01	09/03/2018	Strain in operando AlGaInGaN High-Electron-Mobility Transistor		13	12.4 GB	13/03/2021	
MA-3571	ID01	10/11/2017	Strain imaging in suspended GeSn micro-disks for laser application using Bragg ptychography		7	1.0 GB	14/11/2020	
MA-2625	ID01	10/02/2017	Strain fluctuations in strained, ultrathin Silicon and Silicon-Germanium layers and lines using pink beam nano-diffraction		6	92.6 MB		

### 3.3 A CXI file with raw tomograph

This file exemplifies the use of Scans.



- ✓ Open science policies since the 2010s: ISIS, ILL, ESRF, EuXFEL...
- ✓ Data portals (institutions & federated)
- Hdf5 as standard data format (still painful for some communities) with rich data & metadata
- Progress needed on FAIR formats & their implementation

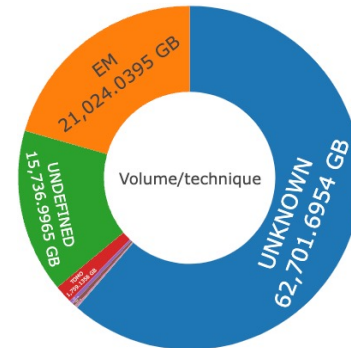
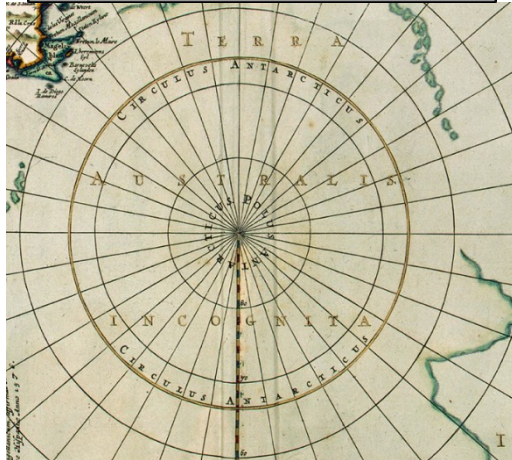
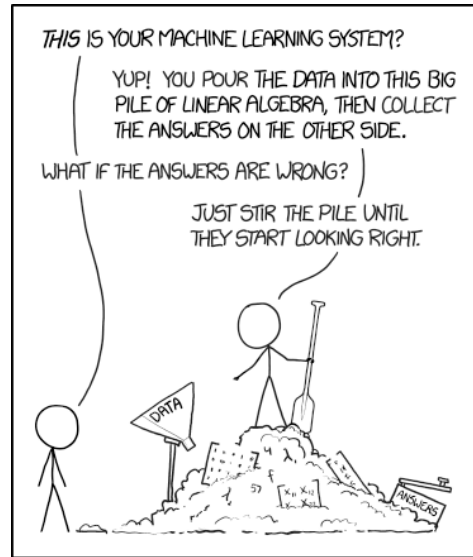
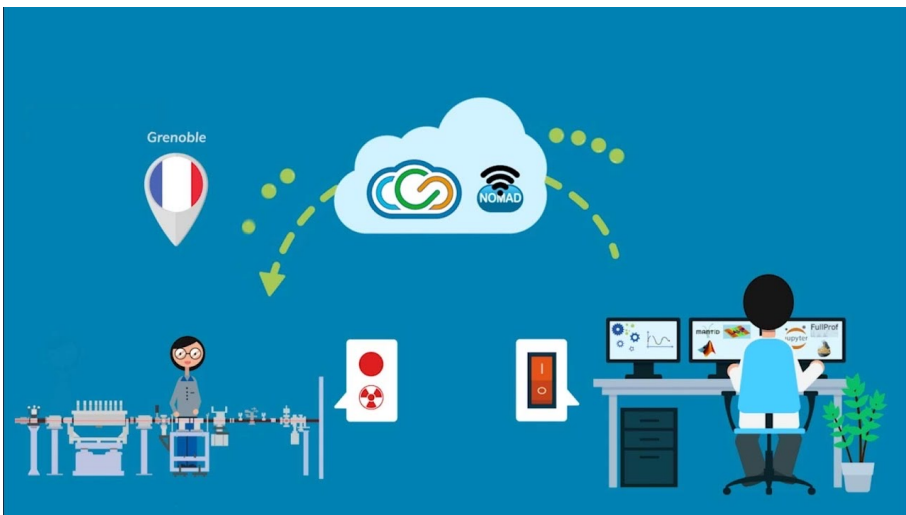


Figure 3: Diagram of a CXI file for storing raw tomographic data.



- The next decade should see a large increase in Big Data analysis (combining many independent experiments)
- Feed open data to machine learning algorithms (or more classical ones)
- **Need to work on FAIR / open data formats with the scientific communities, via:**
  - Conferences
  - Scientific societies (IUCr,..)
- ... & implement it on instruments

# REMOTE DATA VIEWING & ANALYSIS



- VISA – modern virtualisation
- + customised software environment
- + support regular desktop tools
- + ability to share remote environment

h5web

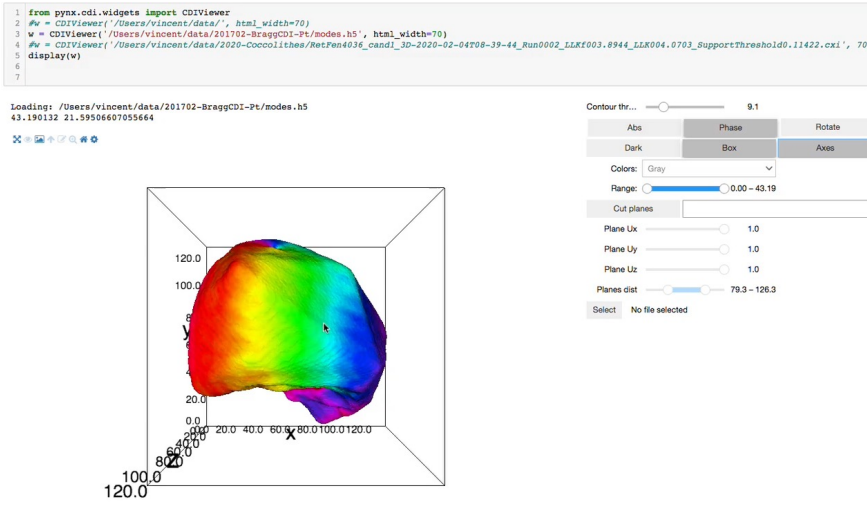
The screenshot shows the h5web interface. At the top, there is a search bar and a table of files. The table has columns for 'Preview', 'Location', and 'Size'. Below the table, there are tabs for 'Initialize', 'View Data', 'Treatment', 'Reduce Spectrum', 'Fit Data', 'Plot', 'Logbook', and 'Readme'. The 'Plot' tab is active, showing a plot of 'Normalized NEXAFS intensity for shifted spectra' for the dataset 'thorondor'. The plot shows several peaks in the energy range of 780 to 805 eV. Below the plot, there is a 'Save Plot' button. To the right, there is a 'Summed Intensity' plot for the dataset 'Aample\_LOB74\_h1\_v4\_0002000\_hplc\_h5'.



Toolkits & widgets for desktop or web data analysis in python

# EOSC DREAMS (CTD)

<http://eosc.eu/?doi=10.15151%2FESRF-ES-119464351>



- From: anywhere (preferably Earth)
- Just give the DOI with your data
- Get an instance transparently paid by an institution (yours, another...)
- Analyse / develop !
- Economical model ?
- Most tools are there thanks to PaNOSC / EOSC developments
- Training material is there, too !

- Cloud resources to be adopted by researchers need to be:
  - *Simple* enough (not *applying* for non-big compute resources)
  - *Ubiquitous* enough so scientific software developers will adopt them



**EUROPEAN OPEN  
SCIENCE CLOUD**



This is the PaNOSC *closing* event...

.. but now the exploitation of all the  
open data & computing tools *begins* !