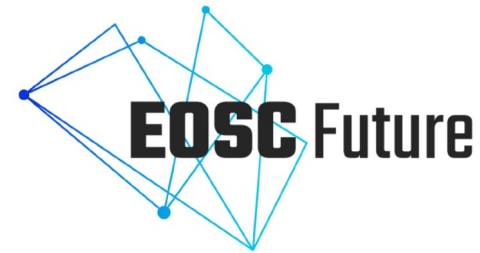


# Integrating the VISA portal into DESY infrastructure



Tim Wetzel, Johannes Reppin, Michael Schuh, Peter van der Reest, Patrick Fuhrmann  
DMA-ST1, Jan 20 2025, Hamburg

The EOSC Future project is co-funded by the European Union Horizon Programme through call INFRAEOSC-03-2020 - Grant Agreement Number 101017536

**HELMHOLTZ** RESEARCH FOR  
GRAND CHALLENGES

**DESY.**



# Overview

- VISA portal originally developed by Institut Laue-Langevin (ILL) during PaNOSC
- Adopted by member institutes of the project itself as well as by ExPaNDS members
- Currently setting up an MoU to
  - Build a community for further development of the portal software and exchange experiences
  - Ensure sustainable reuse of the portal software at multiple institutes
  - Disseminate the use of VISA to a wider audience
- Visa is currently in productive use at ILL and ESRF
- More institutes to join soon
  
- VISA allows scientific users to run data analysis tools in their browser on virtual hardware close to their data
- VISA allows to embrace **FAIR** data principles
  - Data catalogue integration in VISA makes data **Findable**
  - Storage integration at the hosting institutes makes data **Accessible**
  - Standardized data formats per scientific discipline make data **Interoperable**
  - Labelling data for open use after embargo periods makes it **Reusable**

# VISA

Create an instance

Experiments

Select the experiments you wish to associate with your compute resource.

Search for experiments

Search for your experiments using the filters below

Instrument

All instruments

between

2017

and

2021

with open data

included

sort by


date (newest first)

Proposal	Title	Instrument	Start Date	End Date	
p700002	FXE example data	EUXFEL-XMPL	27 Sept 2021	30 Dec 2021	SELECT
p700001	Detector Calibration Test Data	EUXFEL-XMPL	19 Jan 2019	20 Jan 2019	SELECT
CXIDB-ID-98	ExPaNDS Reference Data for Serial Crystallography	EUXFEL-SPB/SFX	30 Aug 2018	03 Sept 2018	SELECT
CXIDB-ID-103	Advances in long-wavelength native phasing at X-ray free-electron lasers	SwissFEL-Aivra	07 Aug 2018	10 Aug 2018	SELECT
p700000	Example Data	EUXFEL-XMPL	08 Nov 2017	31 Dec 2017	SELECT

Results per page51 - 5 of 5 experiments


## Computing Environment

Choose an environment



VISA\_Apptainer

VISA image with Apptainer (former Singularity) preinstalled.



VISA\_CrystFEL

VISA Image with latest CrystFEL installed.

## Choose hardware requirements

4 Cores

8GB memory

Large

8 Cores

16GB memory

XLarge

The screenshot shows a JupyterLab interface with a file browser on the left and a code editor on the right. The file browser shows a directory with various Jupyter notebooks and files. The code editor shows a Bash kernel with several code cells.

**File Browser:**

Name	Last Modified
indexamajig.7018	seconds ago
cowork_demo.ipynb	34 minutes ago
crystfel_tutorial.ipynb	34 minutes ago
demo_bash_nb.ipynb	34 minutes ago
files.lst	seconds ago
gui_demo.ipynb	34 minutes ago
hpc_demo.ipynb	34 minutes ago
jupyterlab_demo.ipynb	8 minutes ago
links.txt	34 minutes ago
slurm_script.sh	34 minutes ago
stream.out	seconds ago

**Code Editor:**

```
[4]: crystfel --version

host listening to pipe for commands: /home/schoensi/crystfel/spipe6854
WARNING: DEPRECATED USAGE: Environment variable SINGULARITY_BIND will not be supported in the future, use APPTAINER_BIND instead
WARNING: DEPRECATED USAGE: Environment variable SINGULARITYENV_PIPEPATH will not be supported in the future, use APPTAINERENV_PIPEPATH instead
INFO: squashfuse not found, will not be able to mount SIF
INFO: fuse2fs not found, will not be able to mount EXT3 filesystems
INFO: Converting SIF file to temporary sandbox...
CrystFEL: 0.10.1-206-gf3645139
License GPLv3+: GNU GPL version 3 or later <http://gnu.org/licenses/gpl.html>.
This is free software: you are free to change and redistribute it.
There is NO WARRANTY, to the extent permitted by law.

Written by Thomas White and others.
INFO: Cleaning up image...
/home/schoensi/crystfel/crystfel: line 31: 6856 Terminated (. $INSTALL_DIR/slurm_listener.sh )

Define sample data

[5]: ls /dcache

CXIDB-ID-103 CXIDB-ID-98

[6]: project=/dcache/CXIDB-ID-98

[7]: geom=${project}/indexing/geoms_CrystFEL_0.10.0/agipd_2120_v3_r45-46.geom
data=${project}/cheetah/hdf5/r0045-Lz/500000/XFEL-r0045-c00.cxi
echo $data > files.lst

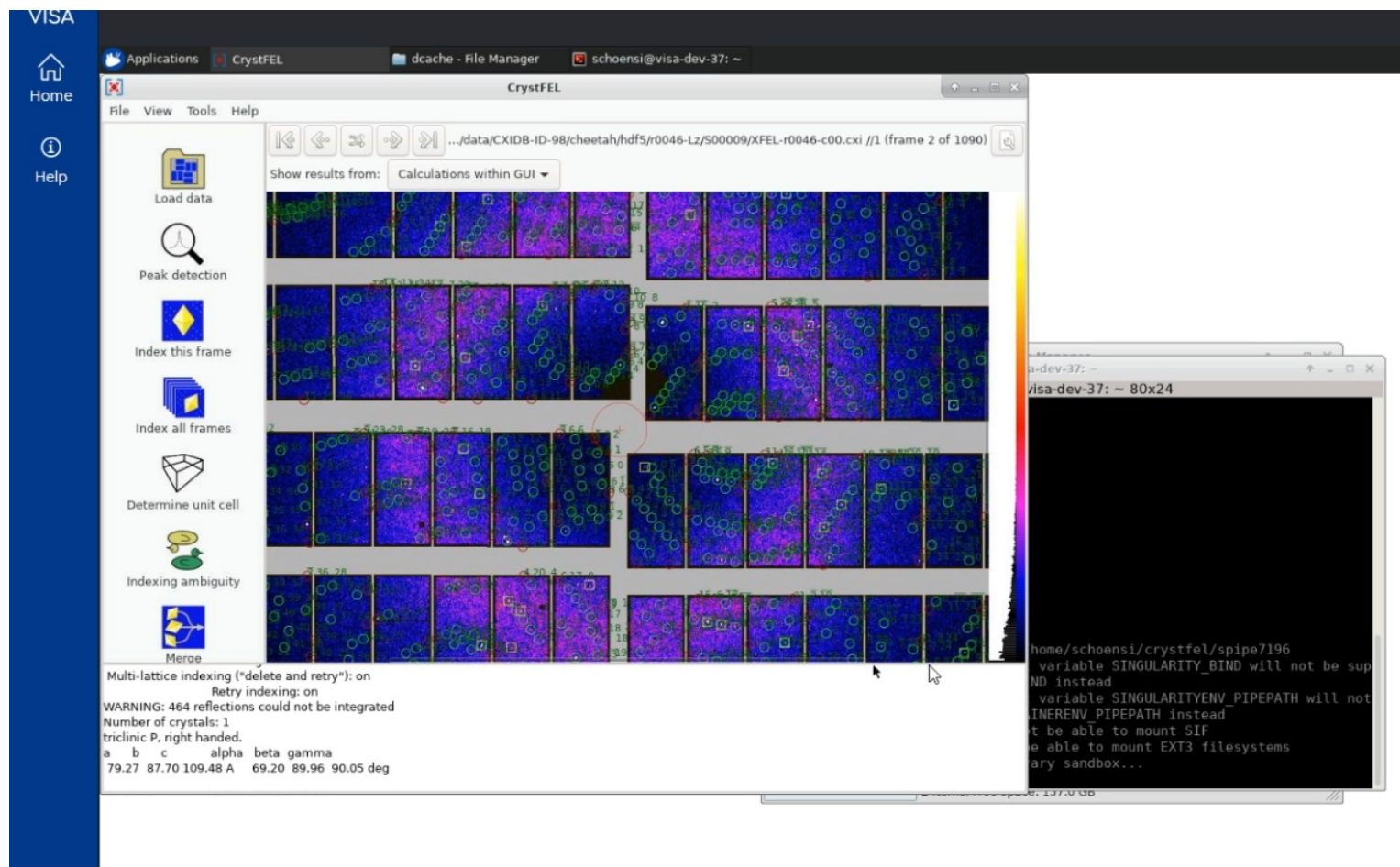
Run CrystFEL command

[*]: indexamajig -i files.lst -g $geom --peaks=zaef --threshold=50 --int-radius=2,3,4 --indexing=none -o stream.out

WARNING: DEPRECATED USAGE: Environment variable SINGULARITY_BIND will not be supported in the future, use APPTAINER_BIND instead
INFO: squashfuse not found, will not be able to mount SIF
INFO: fuse2fs not found, will not be able to mount EXT3 filesystems
INFO: Converting SIF file to temporary sandbox...
Indexing/integration disabled.
0 images processed, 0 hits (-nan%), 0 indexable (-nan% of hits, -nan% overall), 0 crystals, 0.0 images/sec.
1 images processed, 1 hits (100.0%), 0 indexable (0.0% of hits, 0.0% overall), 0 crystals, 0.2 images/sec.
```

Example by Silvan Schön (DESY/FS-SC):

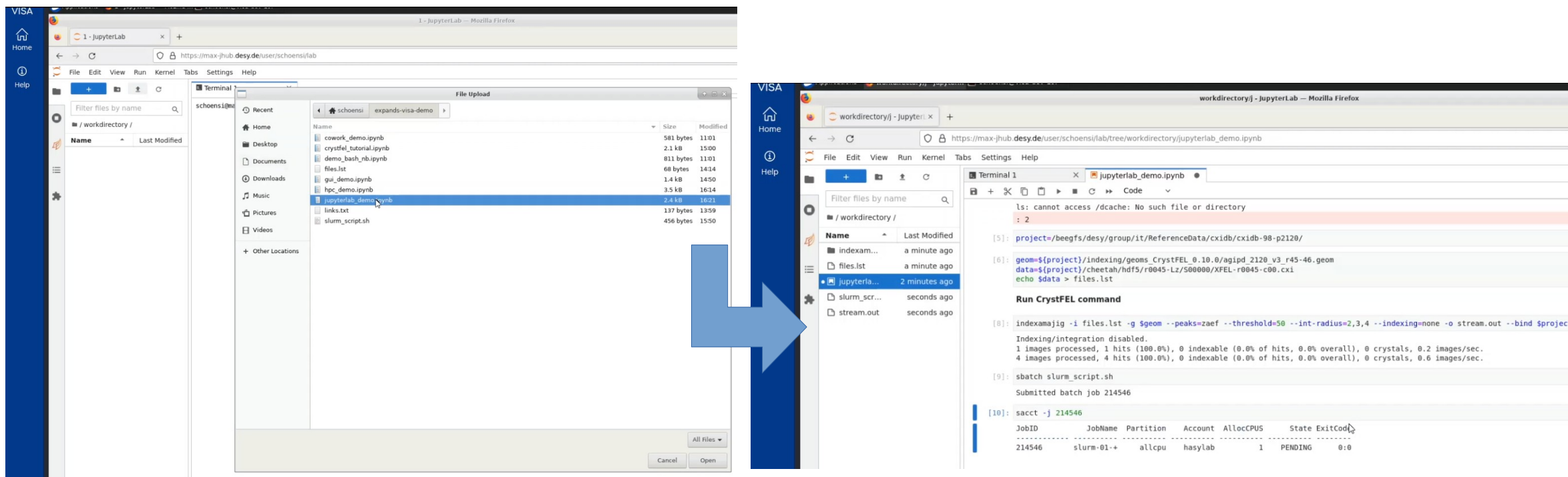
Using CrystFEL Docker Images to run Singularity Container in a Jupyter Notebook (on a Bash Kernel).



Example by Silvan Schön (DESY/FS-SC)

Using CrystFEL Docker Images to run Singularity Container and work with Crystfel 10 Graphical Interface.

## Interoperable workflow that seamlessly extends to Maxwell HPC Infrastructure



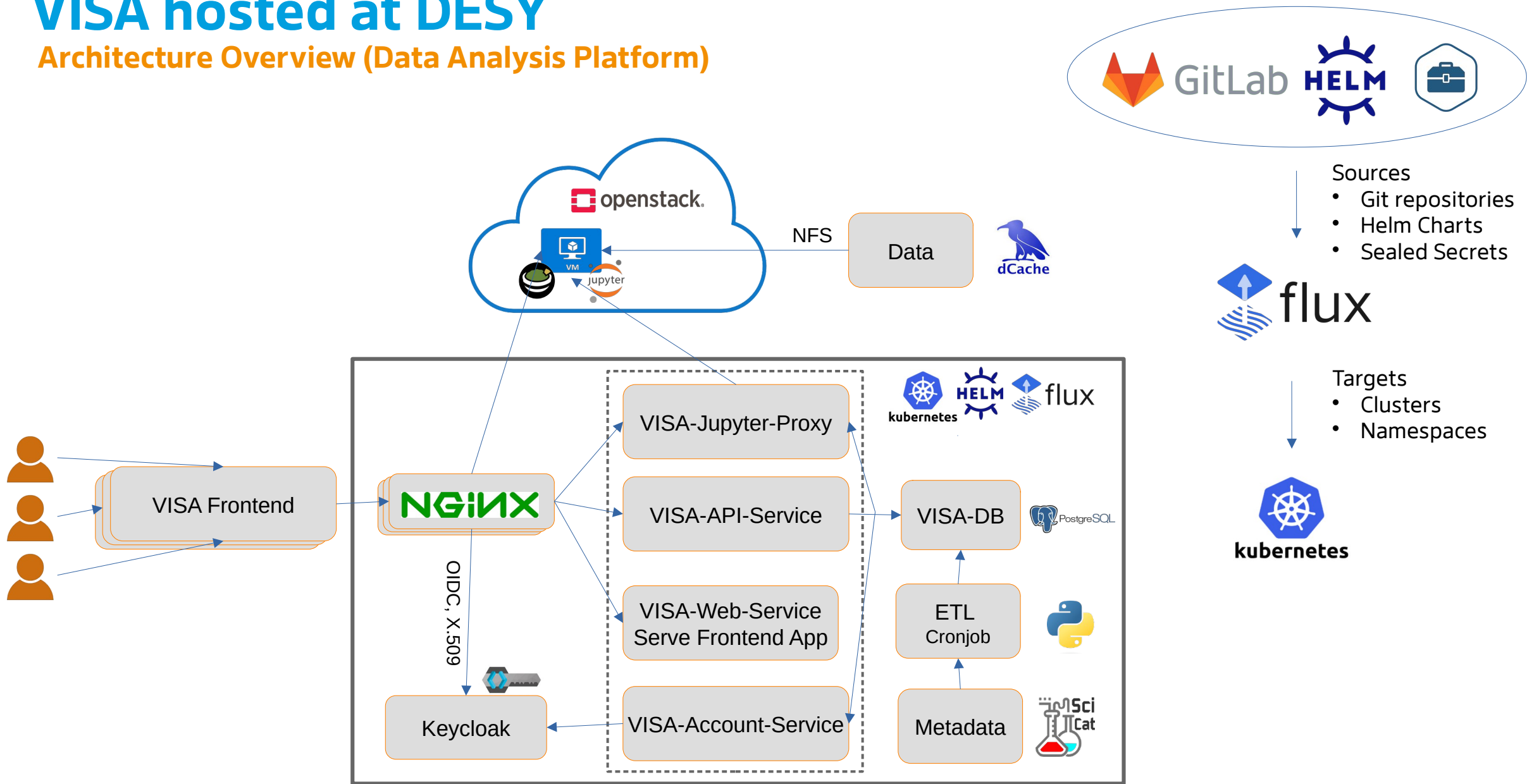
Example by Silvan Schön (DESY/FS-SC)

Use VISA to upload notebook to Maxwell HPC Cluster, Notebook + Slurm Job

Interoperable: Run same analysis, same data (CrystFEL 10 Singularity Container)

# VISA hosted at DESY

## Architecture Overview (Data Analysis Platform)





CLOUD APPLICATIONS AT DESY

English v

Sign in to your account

Username

schuhm

Password

.....

☐ Remember me

Sign In

Or sign in with

EGI Check-in

Helmholtz AAI

GitHub

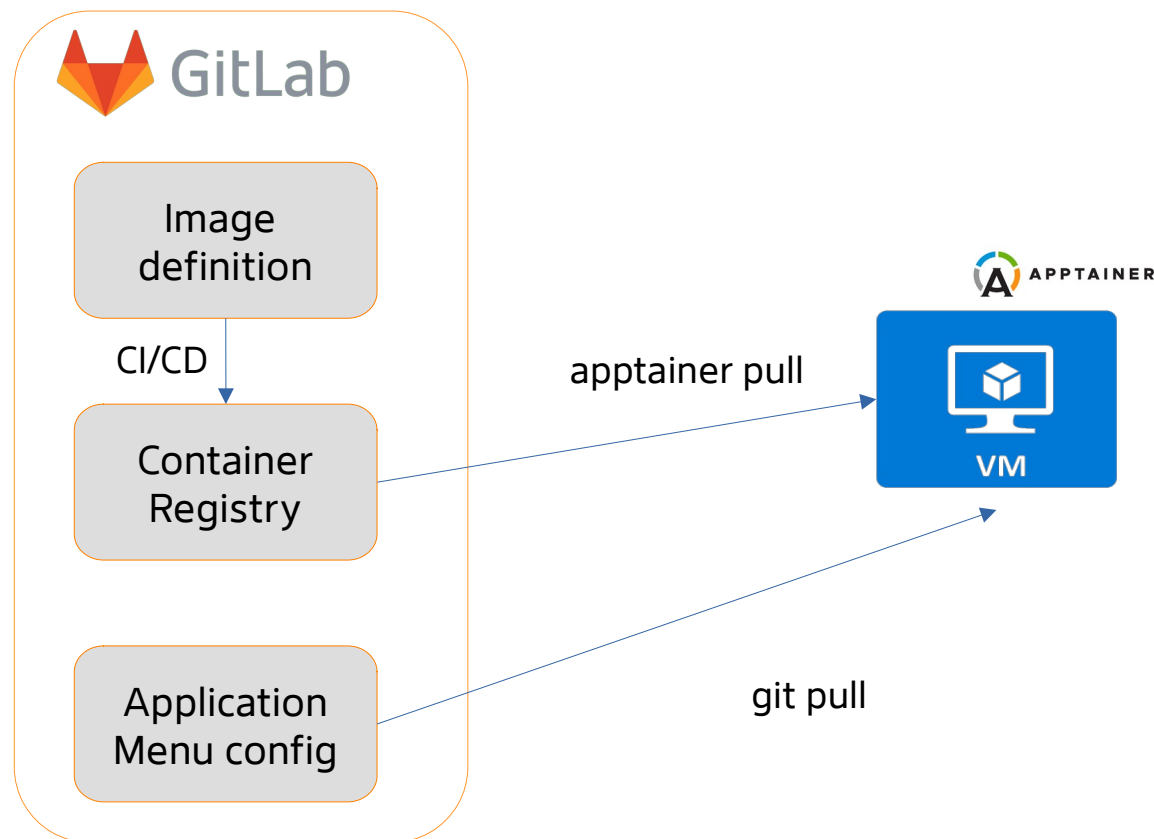
```
dn: uid=*****-dlr-f,ou=people,ou=rgy,o=desy,c=de
[...]
displayName: *****, ***** (*****-dlr-f)
mail: *****@dlr.de
uid: *****-dlr-f          ←--- persistent UUIDs
uidNumber: 1001234         ←--- persistent UID numbers
```

- keycloak.desy.de as SSO solution for DESY
- Local:
  - DESY Registry Accounts (LDAP)
  - DOOR accounts
- Integration with federations:
  - Helmholtz AAI
  - EGI Check-in
  - Github
- Importing accounts and assigning persistent UUIDs
- OIDC clients managed on-premise



# Software deployment

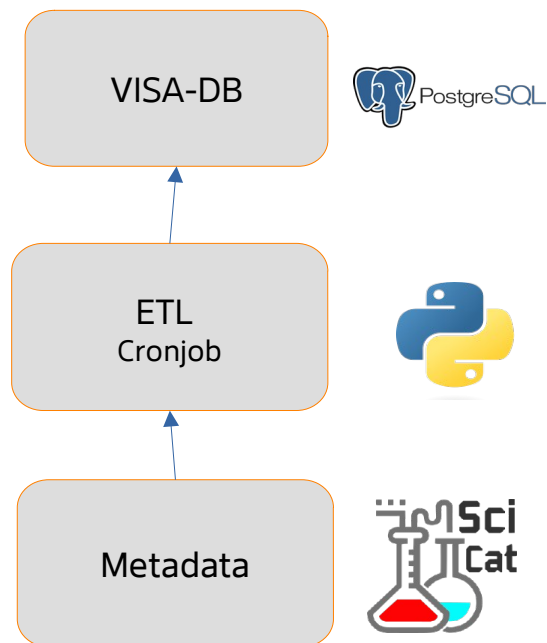
Via Apptainer images built in Gitlab CI/CD



- Software in Apptainer images
  - Many applications already available as Apptainer image from HPC workflows
- Built from .def file in CI/CD pipeline
- Image publicly available in Gitlab registry
- Pulled on application startup
- Application menu entries defined separately in git repository
- Seamless integration into the OS applications
- Menu entries updated from menu config by cronjob pulls the repository regularly
- Updates to the menu by admins

# Metadata integration

## Custom ETL process



- Python script
- Customizable depending on the metadata source (catalogue API format, authN/Z, ...)
- Can be run once for static data or as a cronjob for dynamic data
- Event-based execution would be nice to have (e.g. webhooks)
- Metadata import
  - Experimental specifications
  - Dataset status (embargoed or public)
  - User access rights
  - Storage paths
- Database backup

# DESY Photon Science setup

A high-level view of the world

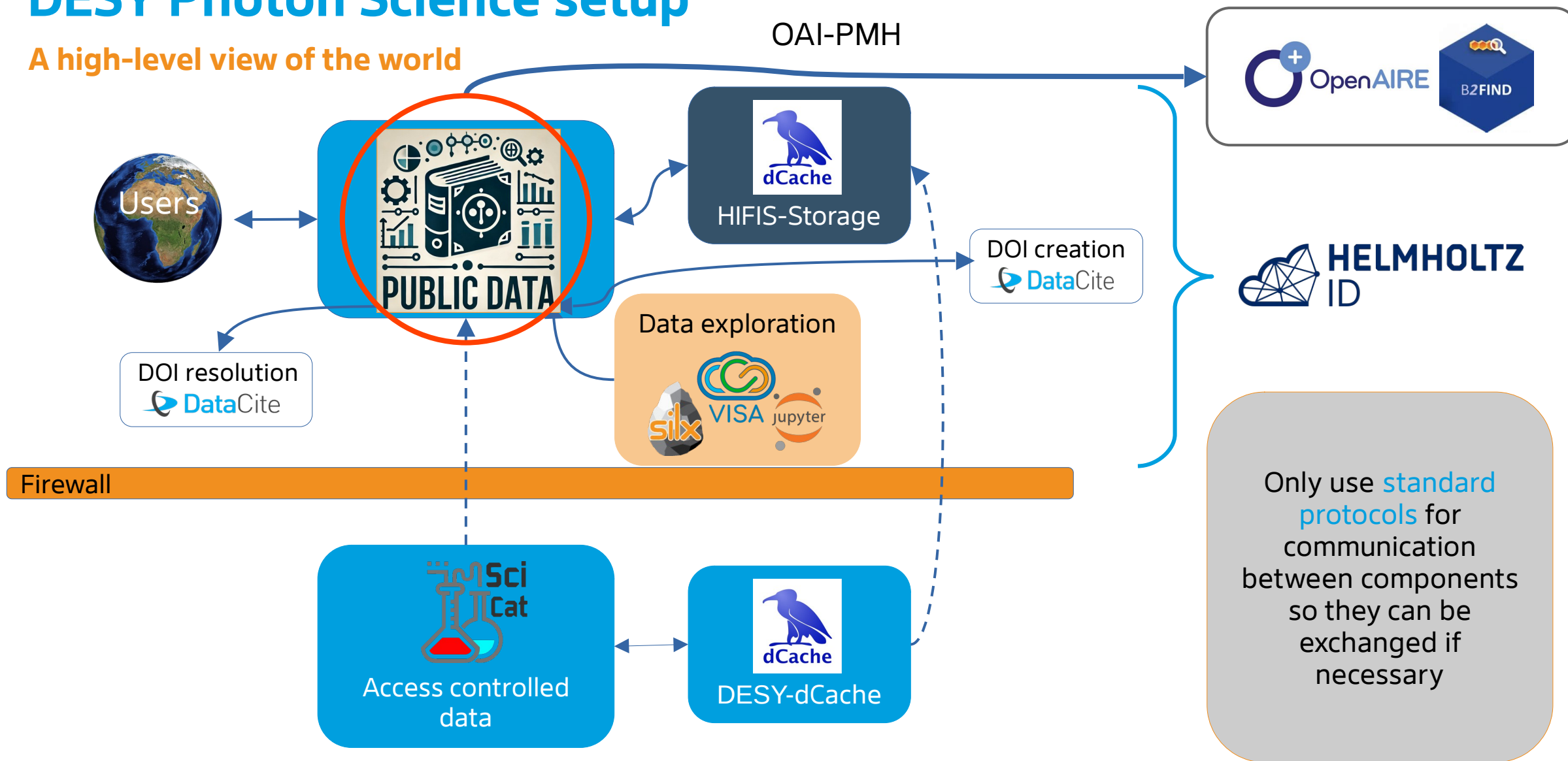


Image adapted from Anton Barty's slide



Search

PID

Text Search

Location

Group

Type

Keywords

Start Date – End Date

Clear

+ Add Condition

Name

Source Folder

Reflectometry curves (XRR and NR) and corresponding fits for machine learning

...do.6497438

spain

.../nfs

General Information

Name

Reflectometry curves (XRR and NR) and corresponding fits for machine learning

Description

This is a compiled dataset of raw X-ray reflectivity (XRR, reflectometry) measurements together with corresponding fit parameters, intentionally published to use as training or test data for machine learning models. (The authors aim to include NR data in further versions of this dataset and plan to include other substrates and materials for XRR. Contributions welcome!)

PID

undefined/10242df2-3868-42cb-bcb2-81c2c44533ec

Type

raw

Creation Time

2024-01-25 18:34

Keywords

Creator Information

Owner

Linus Pithan

Principal Investigator

linus.pithan@desy.de

Contact Email

linus.pithan@desy.de

Owner Group

fsec

Access Groups

File Information

Source Folder

/desy/public-data/upload/daphne4nfdi/10.5281 zenodo.6497438

Scientific Metadata

Search

x

DIP\_1

Experimentalists

Kowark, Stefan

Layer\_CAS

188-94-3

Layer\_formula

C32H16

Layer\_material

Diindenoperylene

Substrate\_temperature

303 (K)

instrument

ESRF, ID10b

q\_max\_fit

0.15 (1/Ang)

year\_experiment

2005

DIP\_2

Path

Size

calc\_xrr.py

2 KB

conda\_env.yml

7 KB

prepare\_plot.py

4 KB

README.html

6 MB

README.ipynb

9 MB

requirements.txt

76 B

xrr\_dataset.h5

254 KB

DESY. | VISA integration at DESY | T Wetzel, M Schuh, J Reppin, P Fuhrmann | DMA-ST1 20th Jan 2025

Page 13

# Thank you

Grateful acknowledgements to

Jamie Hall, Stuart Caunt and Erwan Le Gall (ILL)

Majid Ounsy, Eric Moge, Alain Buteau and Ryan Perseé (SOLEIL)

**Questions?**

## Contact

**DESY.**

Tim Wetzel

IT-RIC

[tim.wetzel@desy.de](mailto:tim.wetzel@desy.de)

Patrick Fuhrmann

IT-RIC

[patrick.fuhrmann@desy.de](mailto:patrick.fuhrmann@desy.de)

Deutsches Elektronen Synchrotron

[www.desy.de](http://www.desy.de)