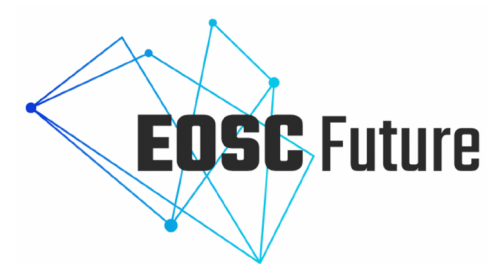


# Integrating the VISA portal into DESY infrastructure



Tim Wetzel, Johannes Reppin, Michael Schuh, Peter van der Reest, Patrick Fuhrmann  
IBERGRID 2023, 25th – 29th September 2023, Benasque, Spain

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**

### 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	<a href="#">FXE example data</a>	EUXFEL-XMPL	27 Sept 2021	30 Dec 2021	<a href="#">SELECT</a>
p700001	<a href="#">Detector Calibration Test Data</a>	EUXFEL-XMPL	19 Jan 2019	20 Jan 2019	<a href="#">SELECT</a>
CXIDB-ID-98	<a href="#">ExPaNDS Reference Data for Serial Crystallography</a>	EUXFEL-SPB/SFX	30 Aug 2018	03 Sept 2018	<a href="#">SELECT</a>
CXIDB-ID-103	<a href="#">Advances in long-wavelength native phasing at X-ray free-electron lasers</a>	SwissFEL-AIvra	07 Aug 2018	10 Aug 2018	<a href="#">SELECT</a>
p700000	<a href="#">Example Data</a>	EUXFEL-XMPL	08 Nov 2017	31 Dec 2017	<a href="#">SELECT</a>

Results per page 5 1 - 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 displays a directory structure with files like 'indexamajig.7018', 'cowork\_demo.ipynb', 'crystfel\_tutorial.ipynb', 'demo\_bash\_nb.ipynb', 'files.lst', 'qui\_demo.ipynb', 'hpc\_demo.ipynb', 'jupyterlab\_demo.ipynb', 'links.txt', 'slurm\_script.sh', and 'stream.out'. The code editor shows a notebook cell with a bash kernel, containing the following code and output:

```
[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 )

```
[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
```

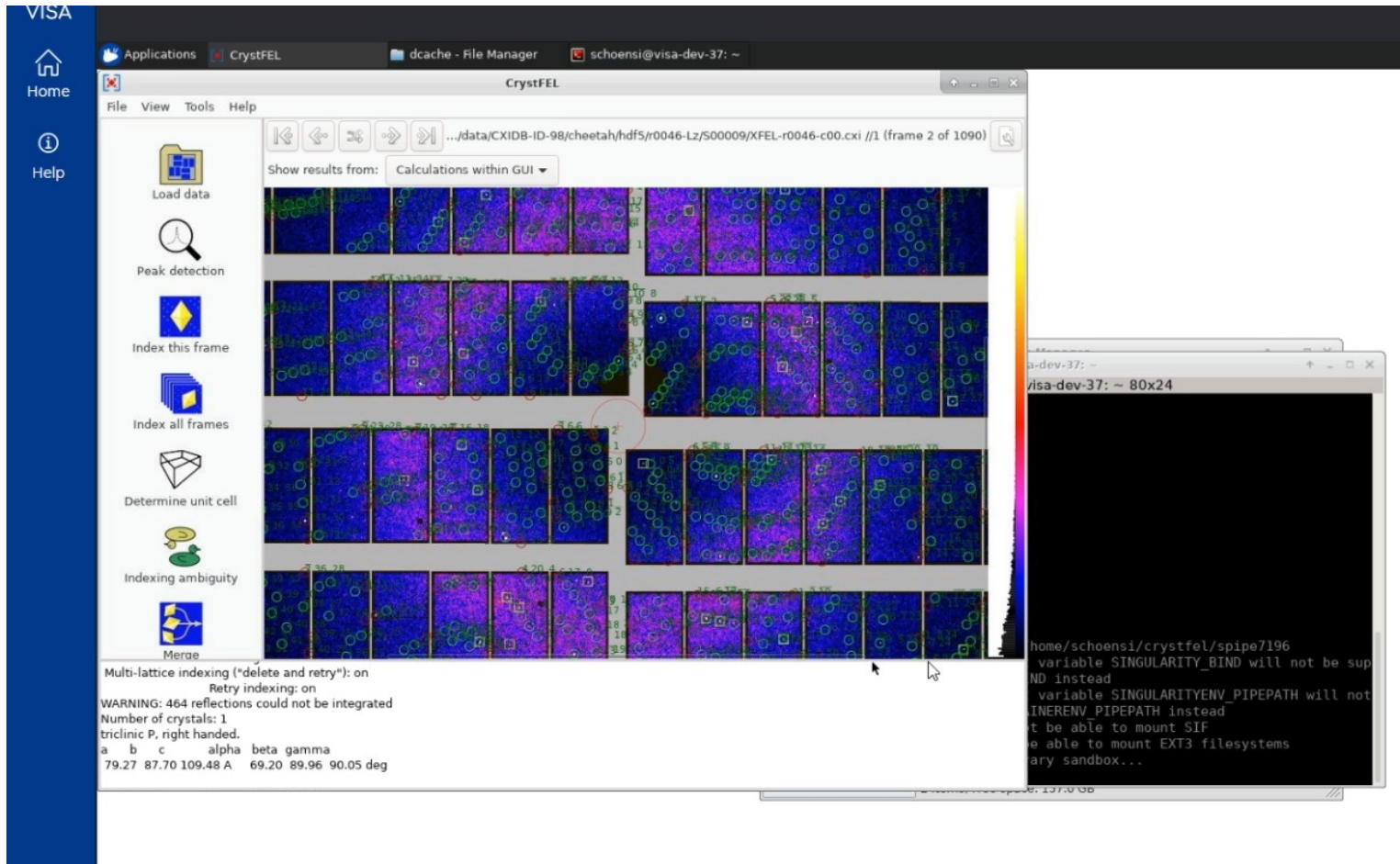
**Run CrystFEL command**

```
[*]: 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):

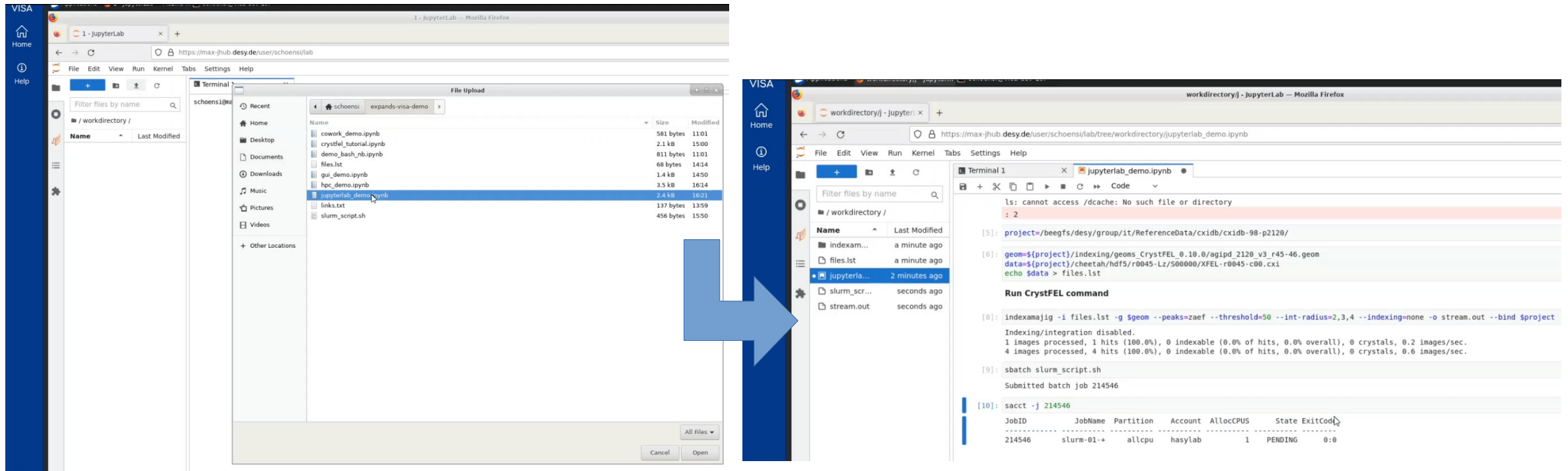
Using CrystFEL Docker Images to run Singularity Container in a Jupyter Notebook (on a Bash Kernel).



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



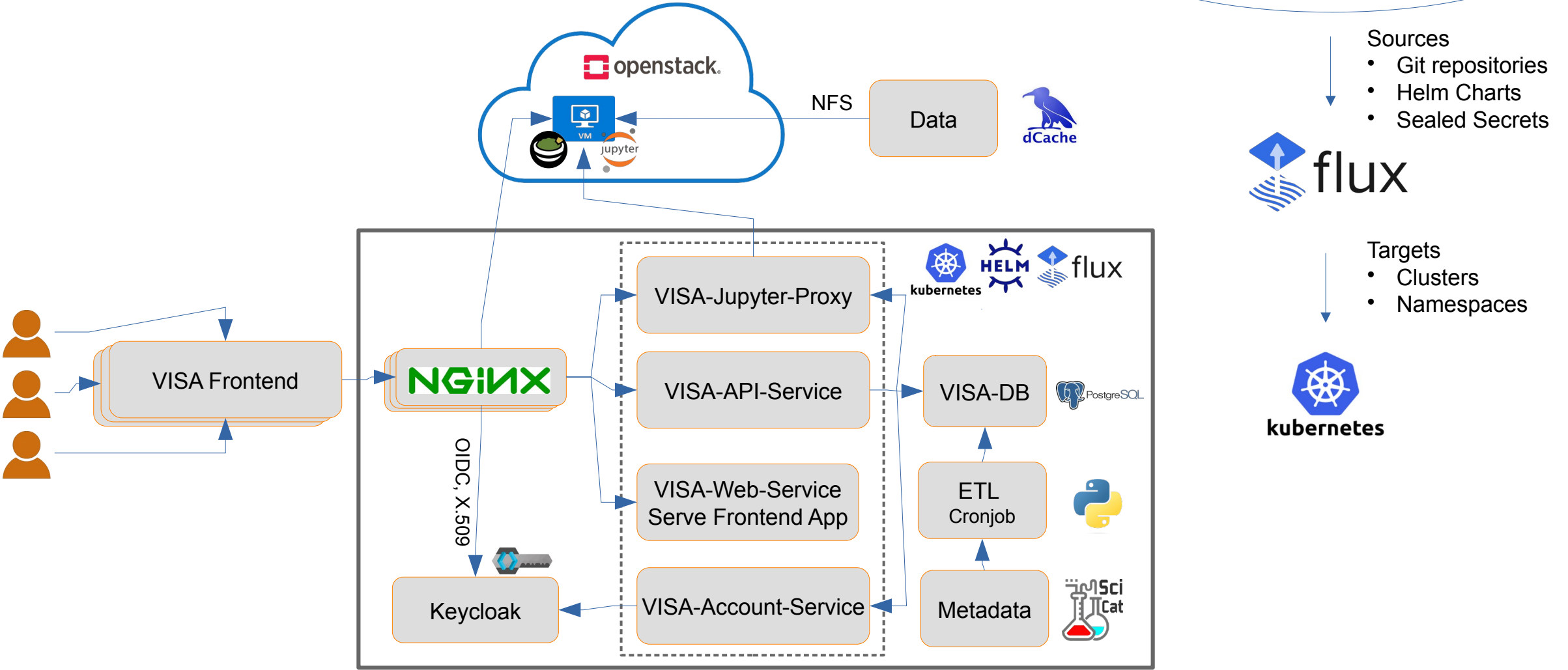
Example by Silvan Schön (DESY/FS-SC)

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)



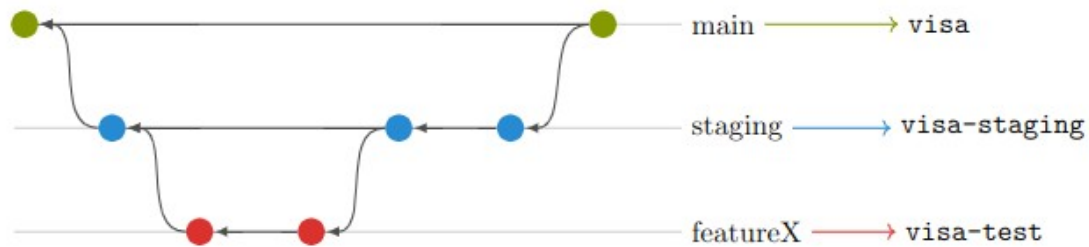
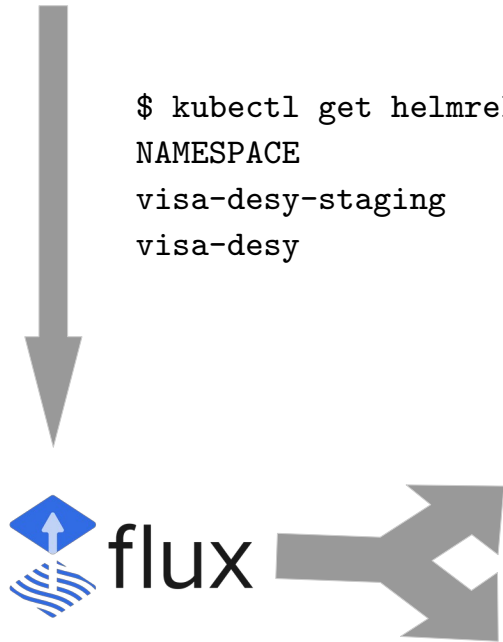
# Deploy to ... Production, Staging

Out of many: VISA Portal, Helmholtz Cloud Portal, ...

```
$ helm repo list | grep visa  
visa-stable    https://gitlab.desy.de/api/v4/projects/2921/packages/helm/stable  
visa-staging   https://gitlab.desy.de/api/v4/projects/2921/packages/helm/staging  
visa-testing   https://gitlab.desy.de/api/v4/projects/2921/packages/helm/testing
```

\$ kubectl get helmrelease -A | grep visa

NAMESPACE	NAME	AGE	READY	STATUS
visa-desy-staging	visa-desy-staging	278d	True	Release reconciliation succeeded
visa-desy	visa-desy	278d	True	Release reconciliation succeeded



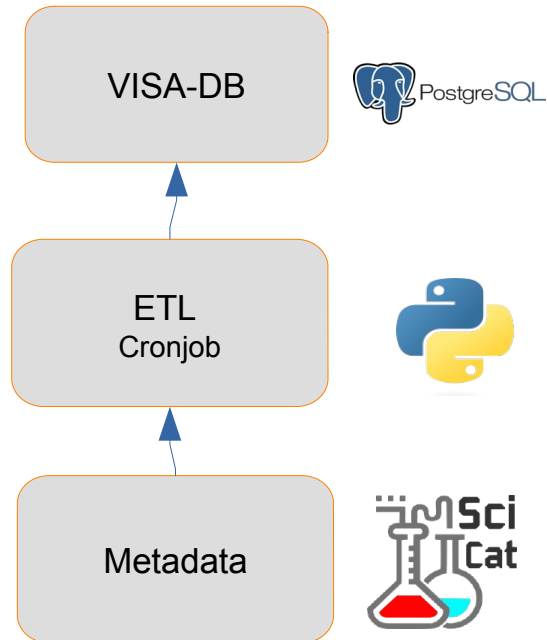
- <https://visa.desy.de>
- <https://visa-staging.desy.de>
- <https://visa-test-83a9eabf.nip.io>

Workflow for updating the VISA deployment branches



# Metadata integration

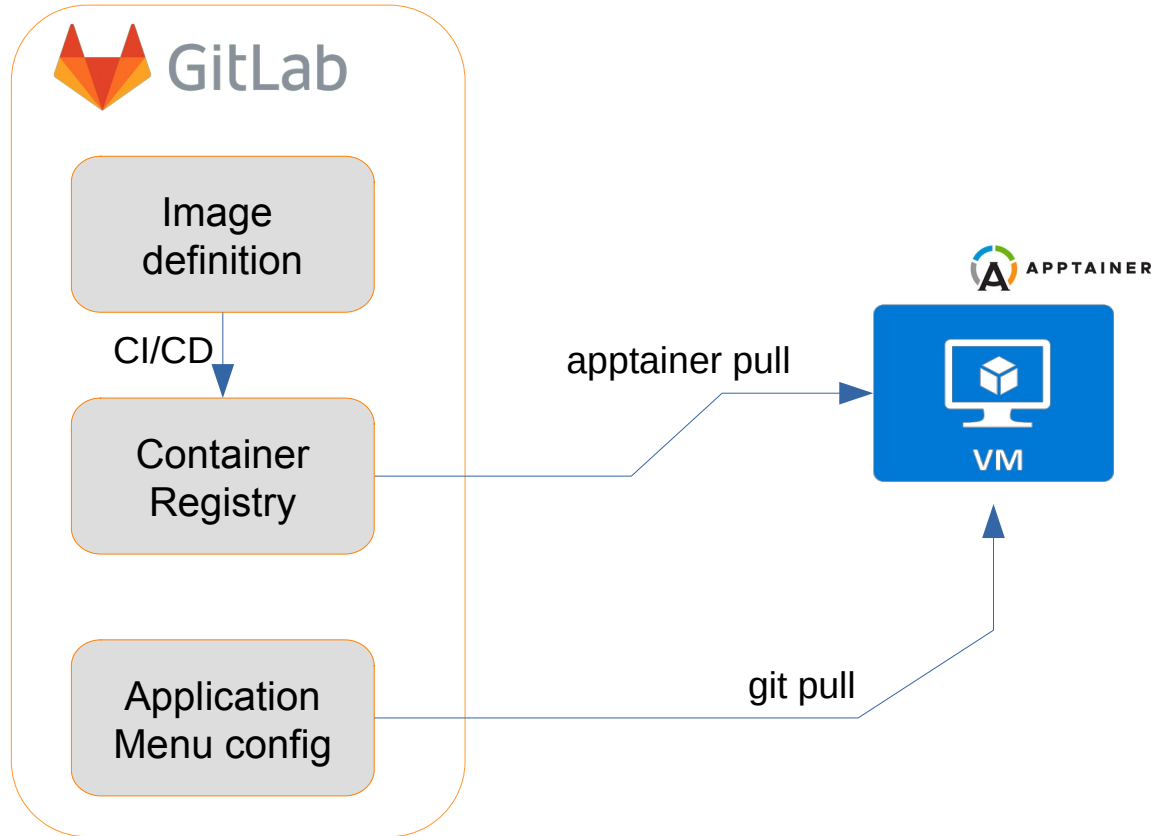
## Custom ETL process



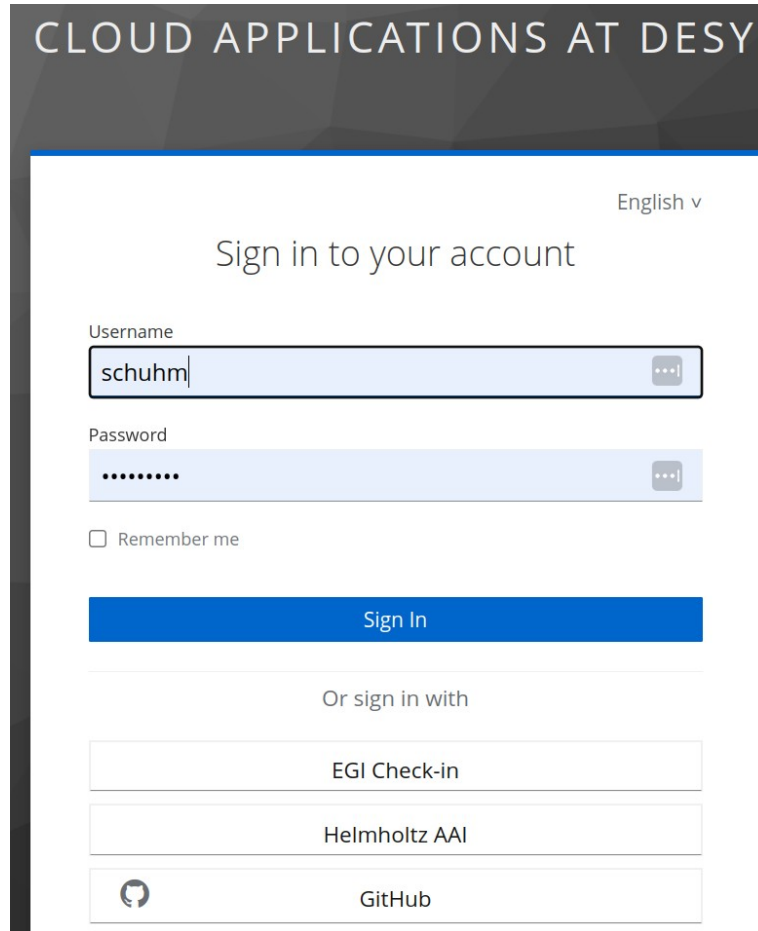
- Python script
- Customizable depending on the metadata source (catalogue API format, authN/Z, ...)
- Can be run once for static data or as a cronjob for dynamic data
- Event-based execution would be nice to have (e.g. webhooks)
  
- Metadata import
  - Experimental specifications
  - Dataset status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

# Software deployment

Via Apptainer images built in Gitlab CI/CD



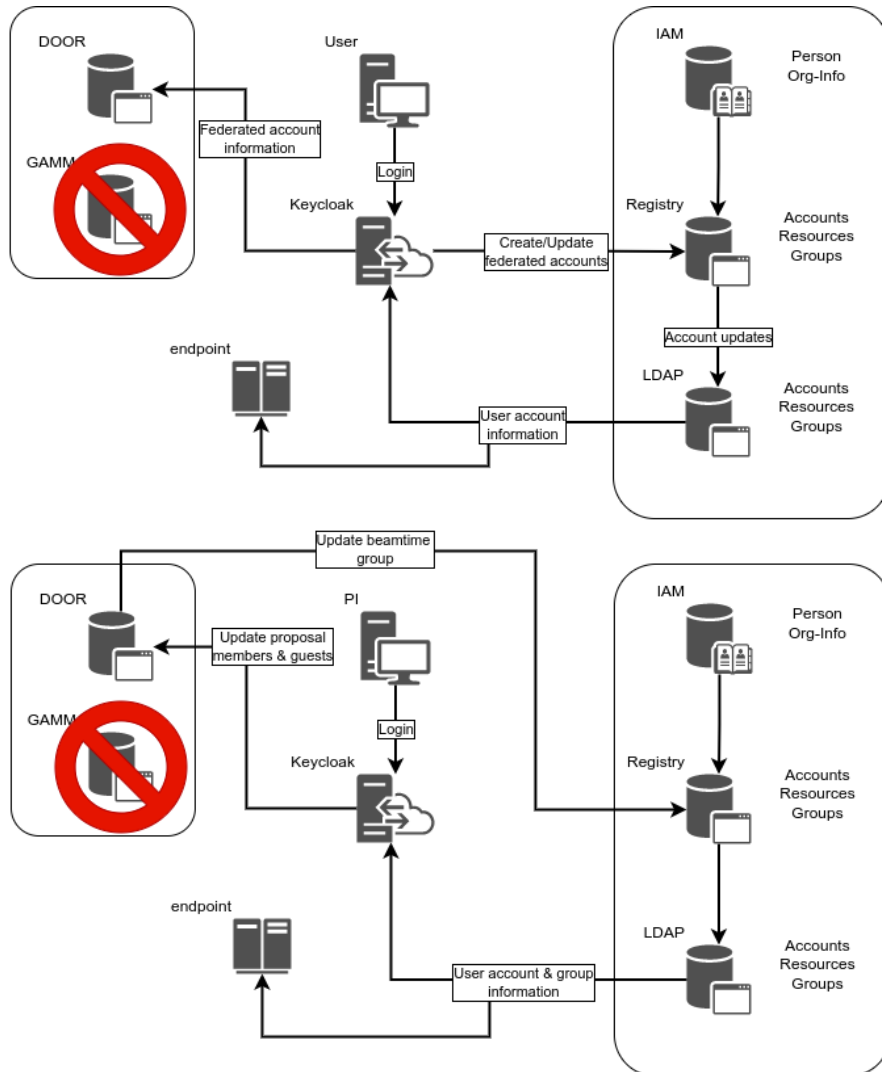
- Software in Apptainer images
  - Many applications already available as Apptainer image from HPC workflows
- Built from .def file in CI/CD pipeline
- Image publicly available in Gitlab registry
- Pulled on application startup
  
- Application menu entries defined separately in git repository
- Seamless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



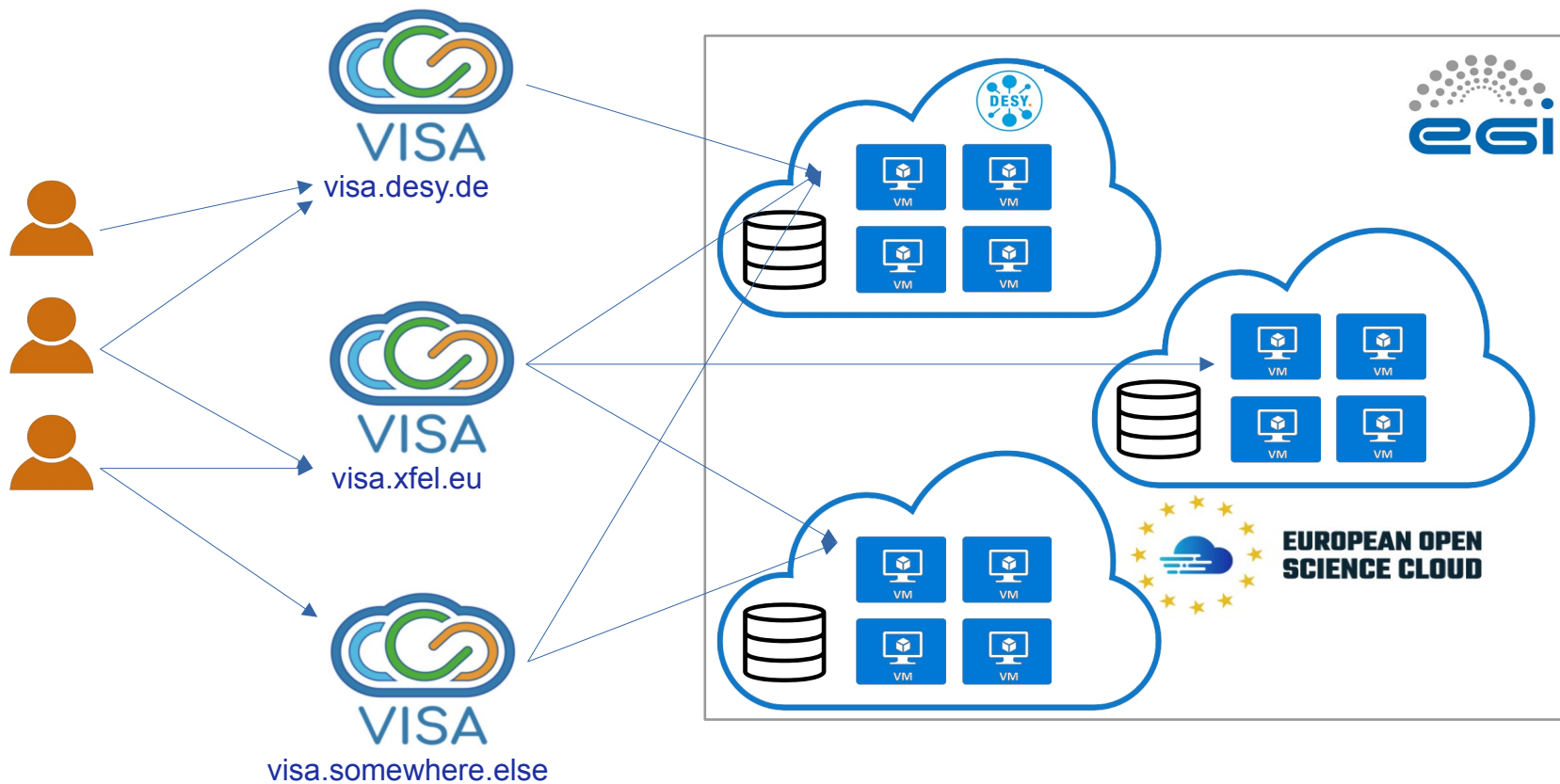
```
dn: uid=*****-dlr-f,ou=people,ou=rgy,o=desy,c=de
[...]
displayName: *****, ***** (*****-dlr-f)
mail: *****@dlr.de
uid: *****-dlr-f          ←--- persistent UIDs
uidNumber: 1001234         ←--- persistent UID numbers
```

- 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 UIDs
- OIDC clients managed on-premise

## DOOR account system



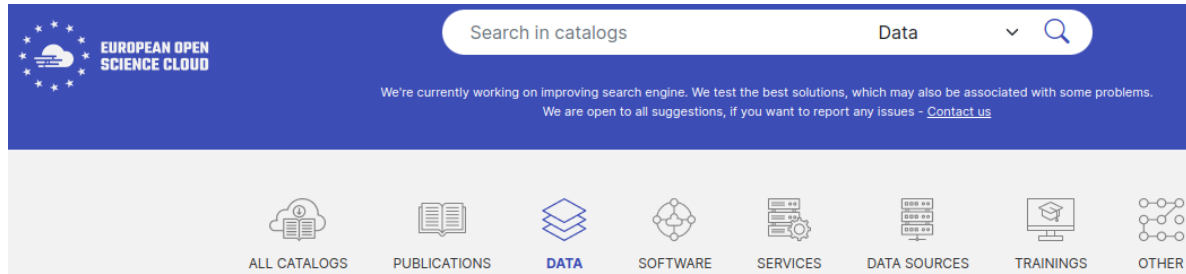
- DOOR as long-standing account system for guests at DESY
- To be reworked to make use of federated accounts via Keycloak as SSO with OIDC
- Integration with
  - HelmholtzAAI
  - EGI Check-In
  - UmbrellaID
- AuthN/Z integration with all other DESY systems
- Several local components to be adapted
  - DOOR portal
  - resource/rights allocation management
  - Data download portal (until now GAMMA)
  - Security relevant services at the beamlines
  - Remote access to beamline control systems
  - ...



- Horizontal infrastructure
- Federated IaaS, cloud interface
- Common AAI
- Shared images, software repos
- Collaborate on monitoring, accounting, development, security, documentation, ...
- Run analysis where the data is

# Storage

## Data Lake and File Transfer Service

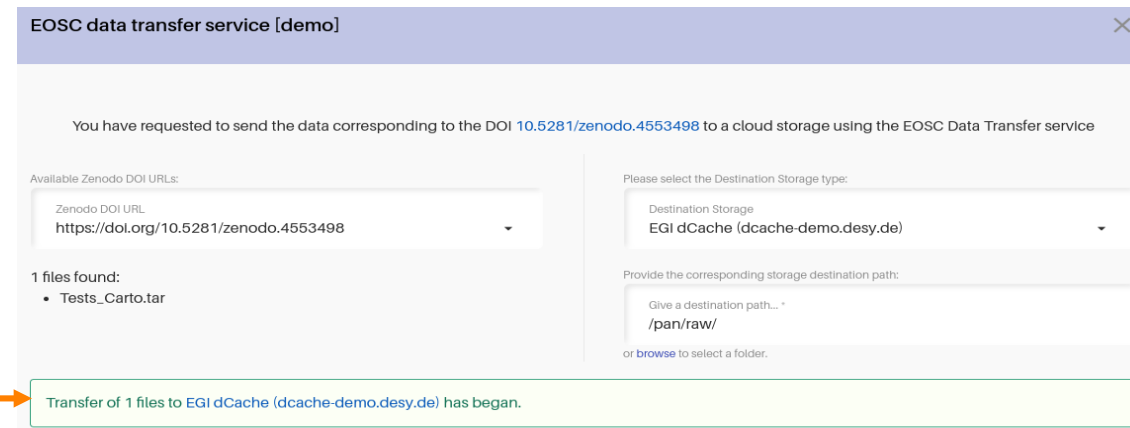
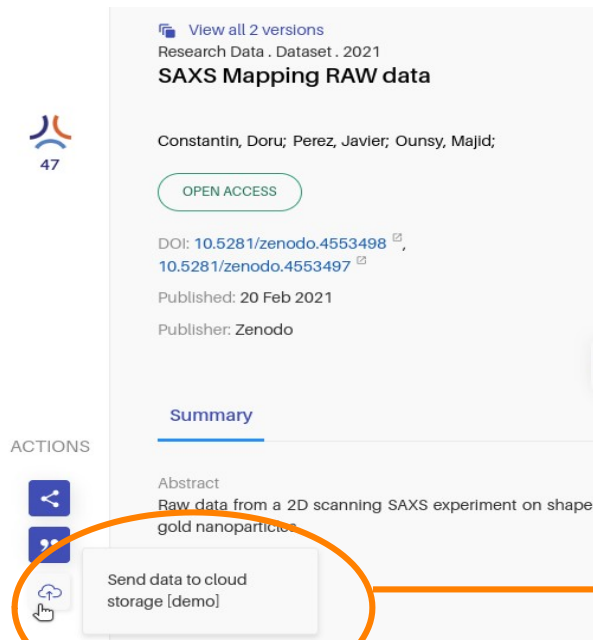


EOSC Marketplace

1207844 (2/8/23) entries in data category



Prototype in place for direct file transfer



# Demo time

Live demo, may not work according to expectations...

[visa-staging.desy.de](https://visa-staging.desy.de)

[visa.desy.de](https://visa.desy.de)



# 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)