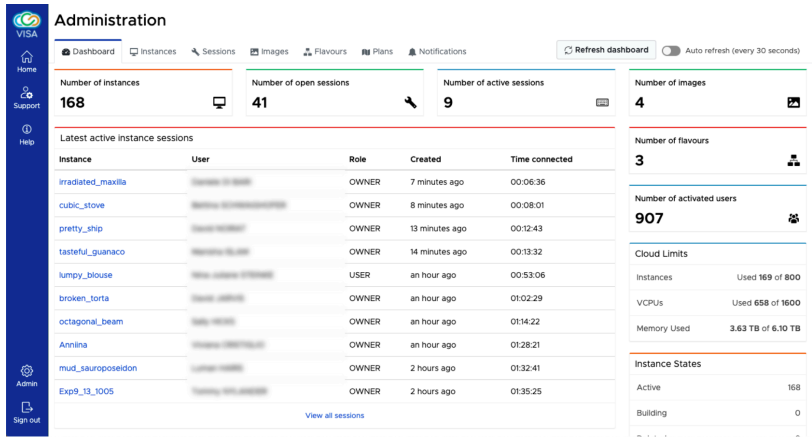
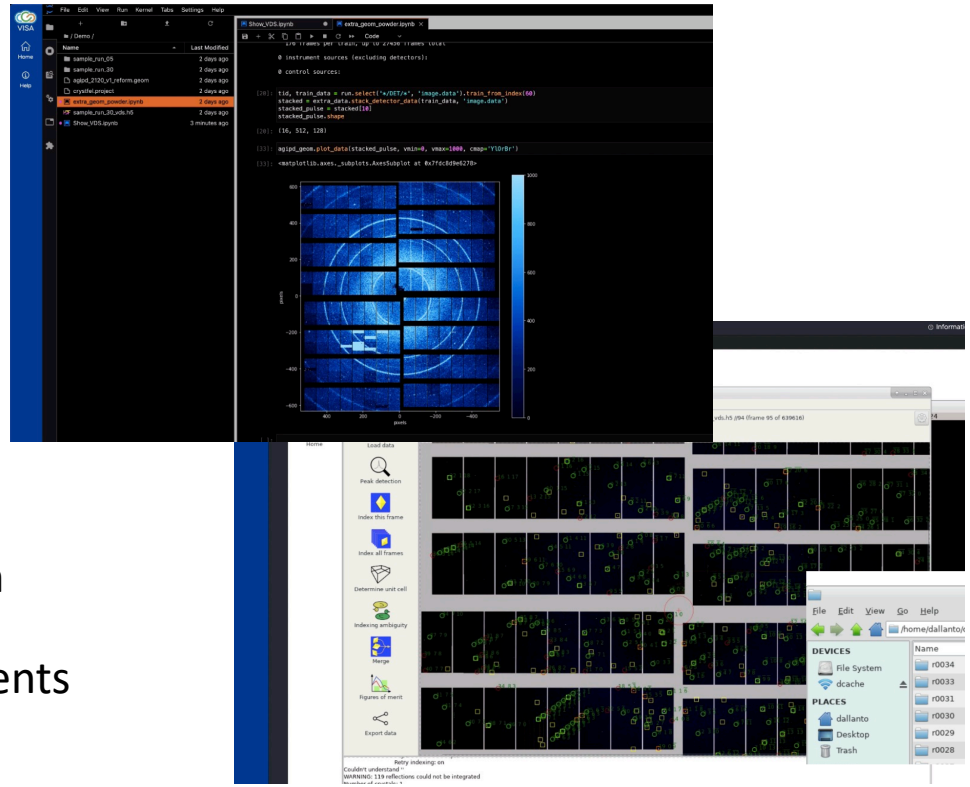
Lessons learned from conducting workshops and trainings using the VISA platform at the European XFEL

Visa - Virtual Infrastructure for Scientific Analysis



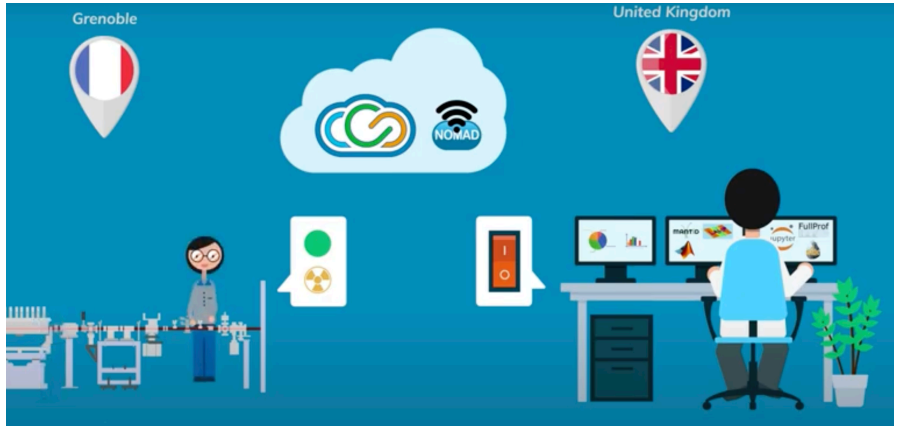
visa.readthedocs.io



- “Abstract away” infrastructure and use case specific configuration
  - Facility integrates and scales generic data analysis environments
    - Jupyter Ecosystem
    - Linux remote desktop
  - Use cases can provide specific data analysis environments, agnostic of the infrastructure

Visa - Virtual Infrastructure for Scientific Analysis

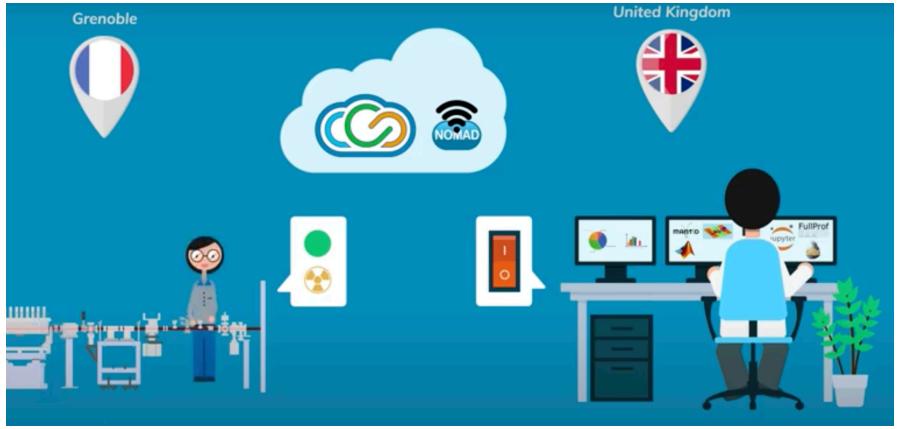
VISA



visa.ill.fr -> youtube.com/watch?v=EAVgbQXRldo

Visa - Virtual Infrastructure for Scientific Analysis

VISA

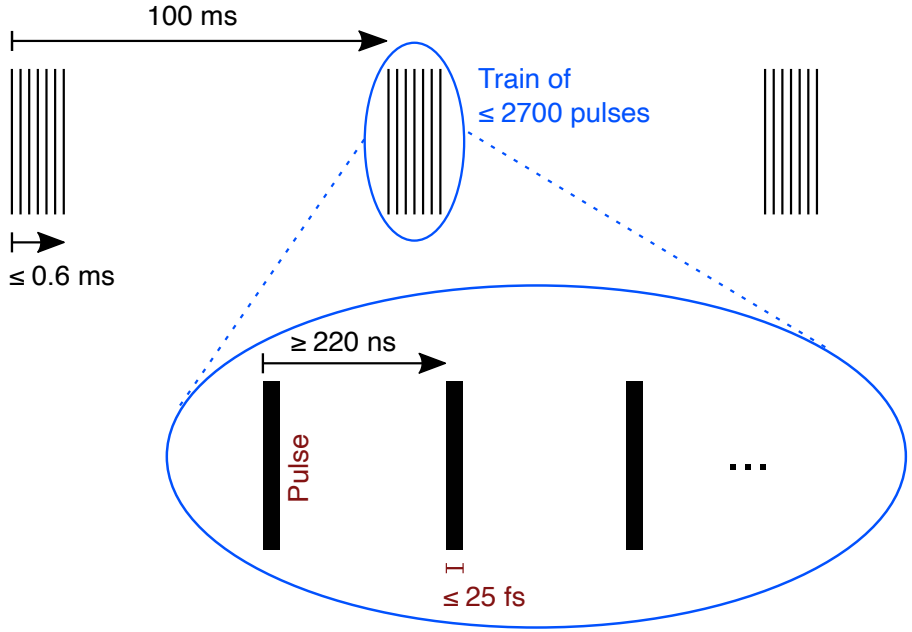


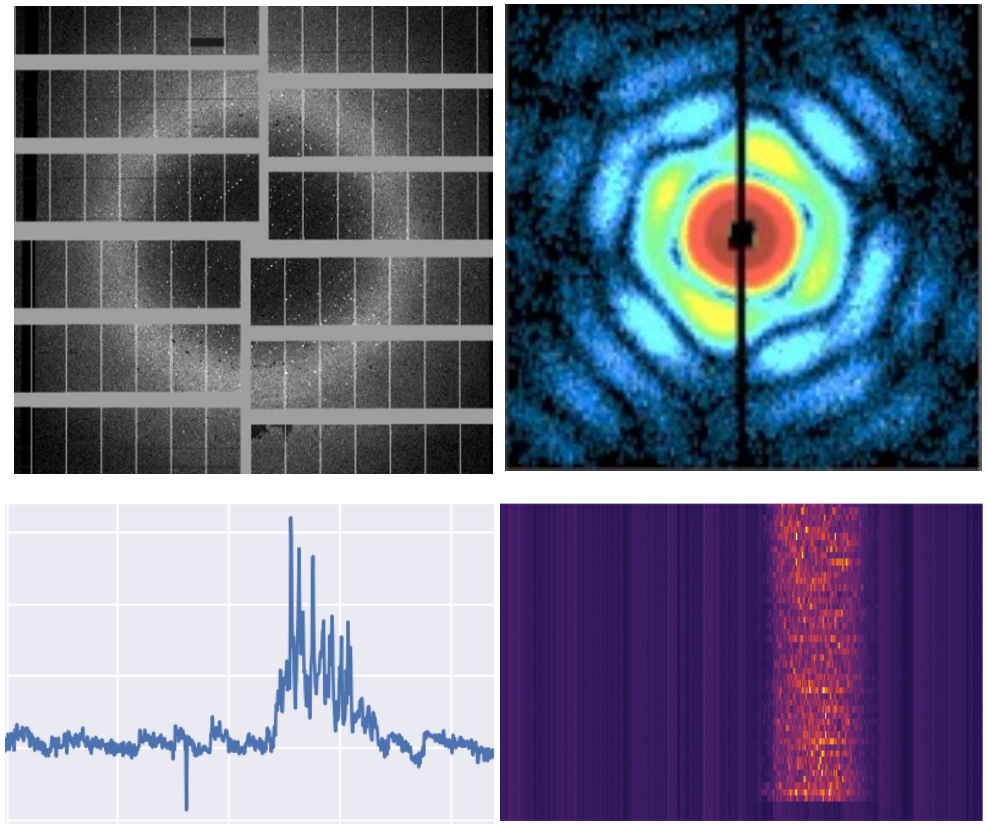
visa.ill.fr -> youtube.com/watch?v=EAVgbQXRldo

- Developed by ILL, project open sourced as part of the Photon and Neutron Open Science Cloud (PaNOSC)
  - Remote data analysis, remote experiment control
- EuXFEL joined MoU on VISA Collaboration
  - Provide users with access to open data, software, compute infrastructure and support (IT / scientific)

Use case 1 - open access to reference data and best practices

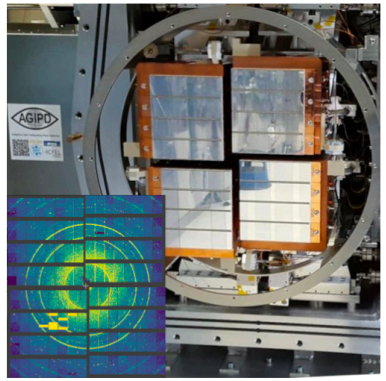
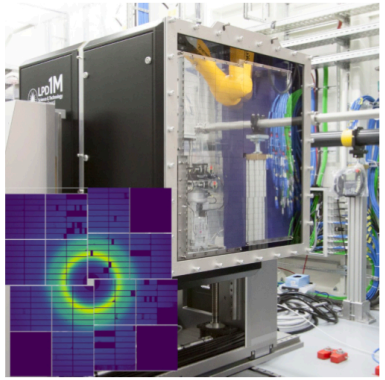
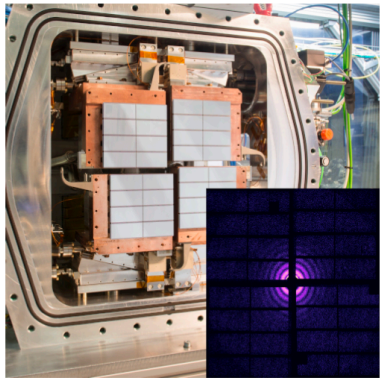
- Diffraction / scattering data from large area detectors
- Spectroscopic data from 1D detectors
- Digitisers, diagnostics devices, control data, …
- High repetition rates at 4.5MHz



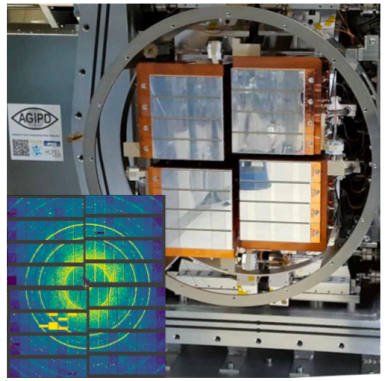
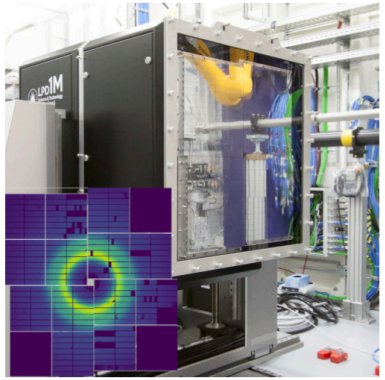
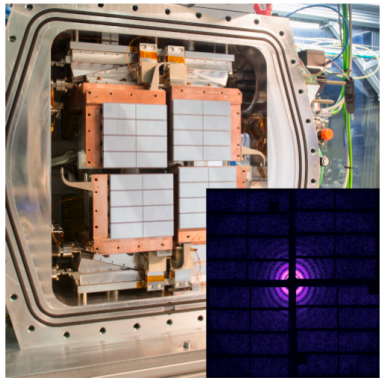


Source: Kirkwood et al. (2022) Nat. Sci. Data 9

Use case 1 - open access to reference data and best practices

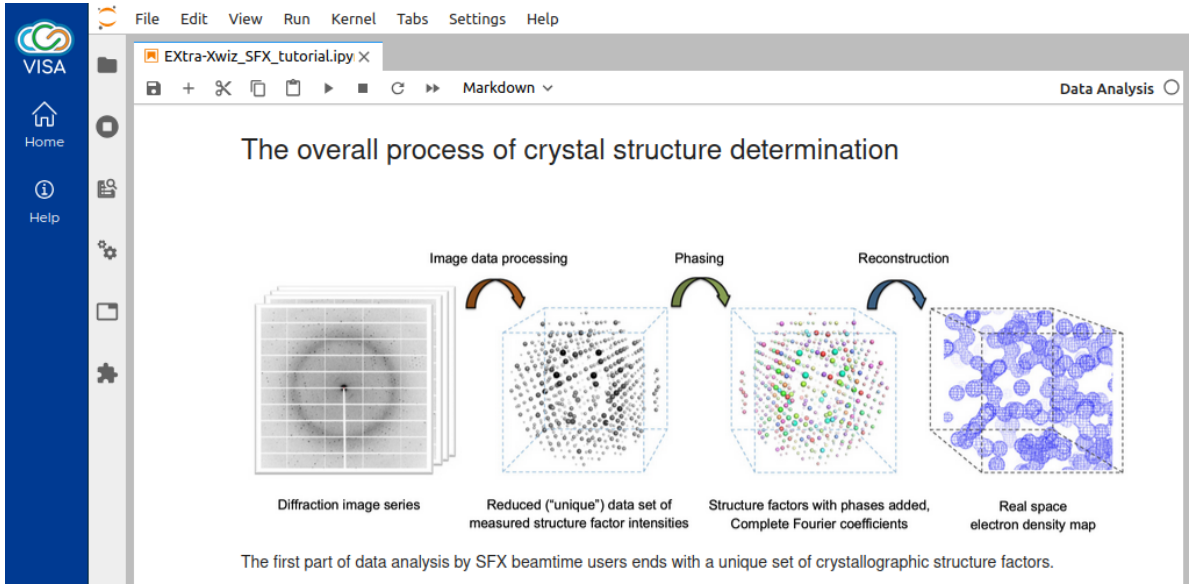
AGIPD-1M	LPD-1M	DSSC-1M	AGIPD-4M
			Coming
21 hours/PiB	29 hours/PiB	19 hours/PiB	6 hours/PiB

Use case 1 - open access to reference data and best practices

AGIPD-1M	LPD-1M	DSSC-1M	AGIPD-4M
			Coming
21 hours/PiB	29 hours/PiB	19 hours/PiB	6 hours/PiB

- Data must be analysed and reduced at the facility
  - Time-limited access to compute infrastructure (allocated for a proposal / beam time)
  - Site specific authentication, authorisation, networking, software repositories, …

Use case 1 - open access to reference data and best practices



- Interactive data analysis environment as additional service for readers of a scientific publication on automating photon science analysis pipeline
- **Data Availability Statement:** Data sets used for the tutorial are freely accessible through [visa.xfel.eu] and, for European XFEL users, at DOI: 10.22003/XFEL.EU-DATA-700000-00
- A reduced data sample was prepared for this use case: `frames_list="indexed_p700000_r0030.lst"`

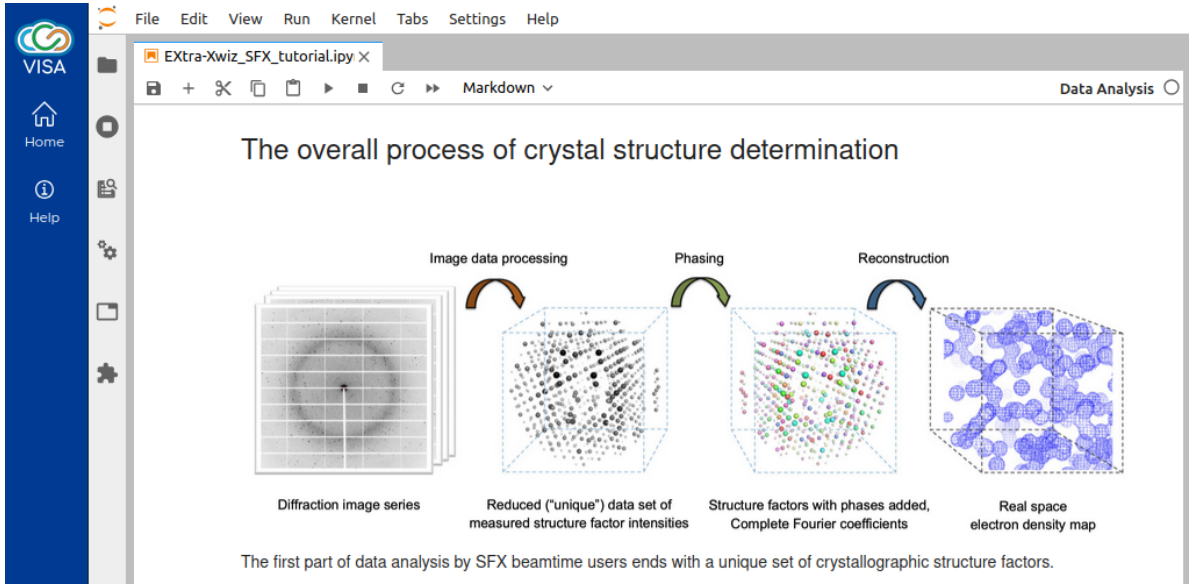
EXtra-Xwiz: A Tool to Streamline Serial Femtosecond Crystallography Workflows at European XFEL

Article *in* Crystals · October 2023

DOI: 10.3390/cryst13111533

mdpi.com/2073-4352/13/11/1533

Use case 1 - open access to reference data and best practices



- Interactive data analysis environment as additional service for readers of a scientific publication on automating photon science analysis pipeline
- **Data Availability Statement:** Data sets used for the tutorial are freely accessible through [visa.xfel.eu] and, for European XFEL users, at DOI: 10.22003/XFEL.EU-DATA-700000-00
- A reduced data sample was prepared for this use case: `frames_list="indexed_p700000_r0030.lst"`

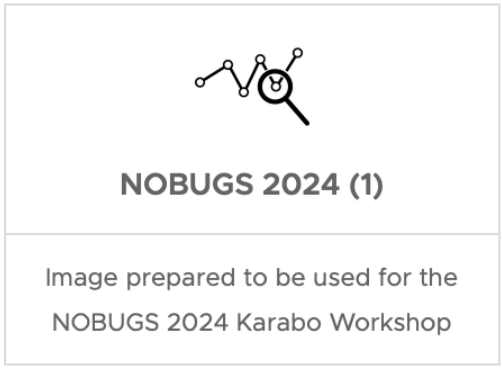
EXtra-Xwiz: A Tool to Streamline Serial Femtosecond Crystallography Workflows at European XFEL

Article *in* Crystals · October 2023

DOI: 10.3390/cryst13111533

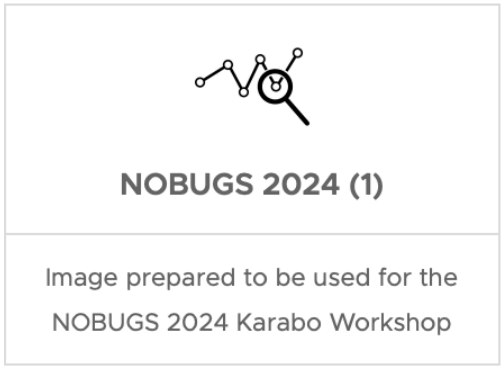
mdpi.com/2073-4352/13/11/1533

Use case 2 - Environments for trainings and tutorials on the Karabo control system



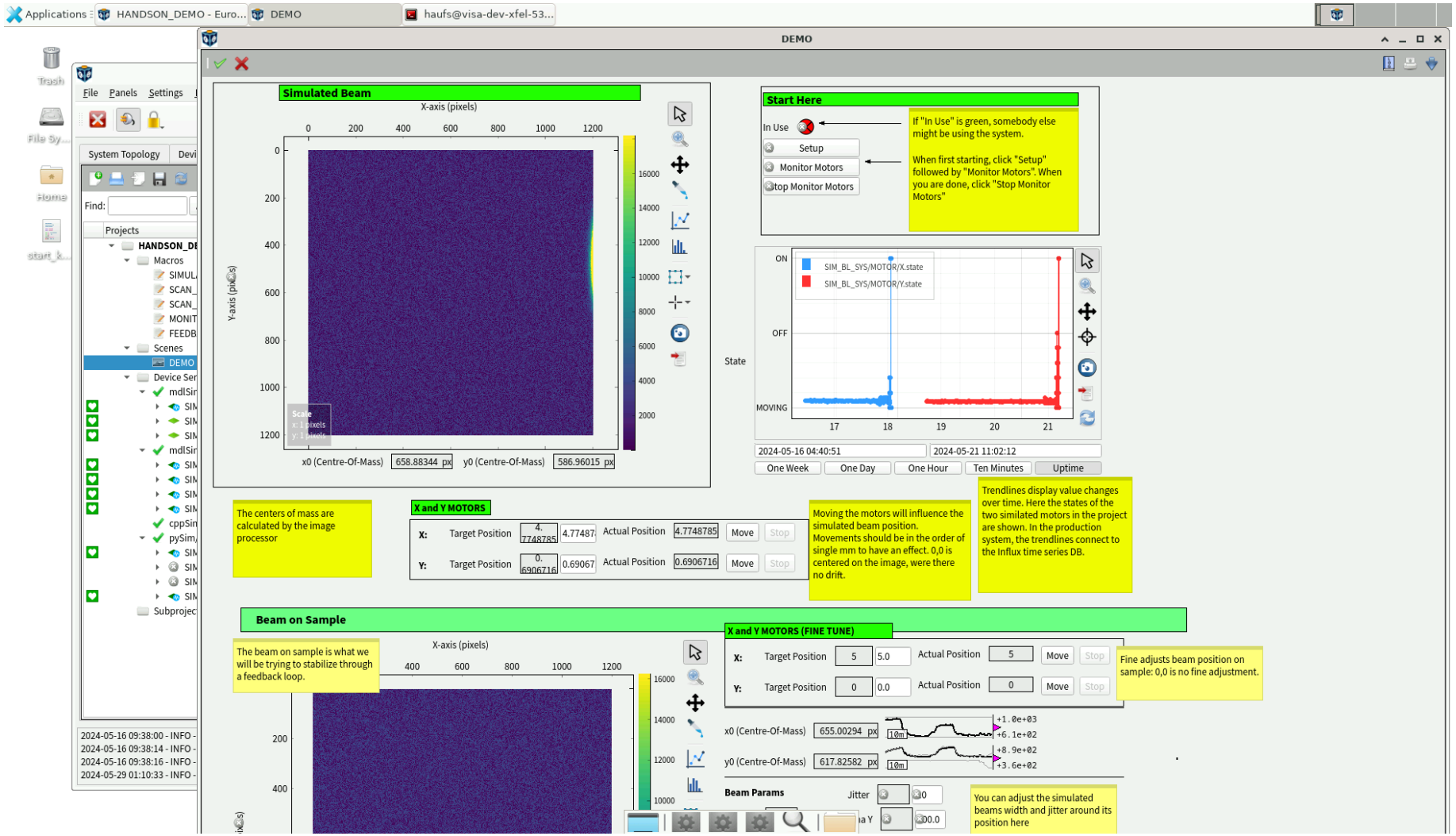
- In-house developed open source system for Supervisory Control and Data Acquisition (SCADA)
- Karabo GUI is the main operator interface at EuXFEL
- Workshop participants interacting with Karabo through their web browser, using the remote desktop
- Image for a standalone Karabo installation, including auxiliary services: project database, message broker
- Workshop content updated outside of the image
  - Tasks implemented as git repo, users checkout tags
  - Resources imported from Nextcloud
  - Mar Internal at EuXFEL
    May Summer school at EIROforum
    Sep Satellite event at NOBUGS 2024

Use case 2 - Environments for trainings and tutorials on the Karabo control system



- In-house developed open source system for Supervisory Control and Data Acquisition (SCADA)
- Karabo GUI is the main operator interface at EuXFEL
- Workshop participants interacting with Karabo through their web browser, using the remote desktop
- Image for a standalone Karabo installation, including auxiliary services: project database, message broker
- Workshop content updated outside of the image
  - Tasks implemented as git repo, users checkout tags
  - Resources imported from Nextcloud
  - Mar Internal at EuXFEL
    May Summer school at EIROforum
    Sep Satellite event at NOBUGS 2024

Use case 2 - Environments for trainings and tutorials on the Karabo control system



- Karabo GUI to build complex hierarchical scenes with Qt Widgets
- Use of labels to explain elements in great detail
- Runs with a simulated experiment, use case needs no data access
- Efficient real-time support by tutors leveraging the collaborative nature of the remote desktop service

Use case 3 - Empirical methods in Research Software Engineering (RSE)

```
xfel.eu data-reduction-recipe v1.0

- AgipdGain <exdf.data_reduction.builtins.AgipdGain>
rechunk-keys     SPB*AGIPD1M*:xtdf    image.data   (-1, 1, None, None)
subslice-keys    SPB*AGIPD1M*:xtdf    image.data   [0, :, :]

- ManualPattern <exdf.data_reduction.builtins.ManualPattern>
select-entries   SPB*AGIPD1M*:xtdf    image              by_id [10:60]

- SomeUserMethod <user.package.module>
...
```

Philipp Schmidt (IFDEPS 2024): Data reduction activities at European XFEL: early results

- Domain Specific Language (DSL) for flexible data reduction pipelines
  - Extendable series of semantic expressions, generated by use case specific plug-ins
  - Operationalise rewriting of files (EXDF based on HDF5)

Use case 3 - Empirical methods in Research Software Engineering (RSE)

```
xfel.eu data-reduction-recipe v1.0

- AgipdGain <exdf.data_reduction.builtins.AgipdGain>
rechunk-keys     SPB*AGIPD1M*:xtdf    image.data   (-1, 1, None, None)
subslice-keys    SPB*AGIPD1M*:xtdf    image.data   [0, :, :]

- ManualPattern <exdf.data_reduction.builtins.ManualPattern>
select-entries   SPB*AGIPD1M*:xtdf    image              by_id [10:60]

- SomeUserMethod <user.package.module>
...
```

- How do users interact with our data reduction tools
- How are manual patterns achieved
- Extend DSL with user supplied operations

Philipp Schmidt (IFDEPS 2024): Data reduction activities at European XFEL: early results

- Domain Specific Language (DSL) for flexible data reduction pipelines
  - Extendable series of semantic expressions, generated by use case specific plug-ins
  - Operationalise rewriting of files (EXDF based on HDF5)

Wrap up

- Use case 1 - open access to reference data and best practices
  - Run data analysis software on sampled, reduced reference data
- Use case 2 - Environments for trainings and tutorials on the Karabo control system
  - Collaborative remote desktop sessions for real-time support from domain experts
- Use case 3 - Empirical research in Research Software Engineering (RSE)
  - Collaborative remote desktop sessions for interviewing users about data reduction recipes