SciCat

advance your data to advance your science

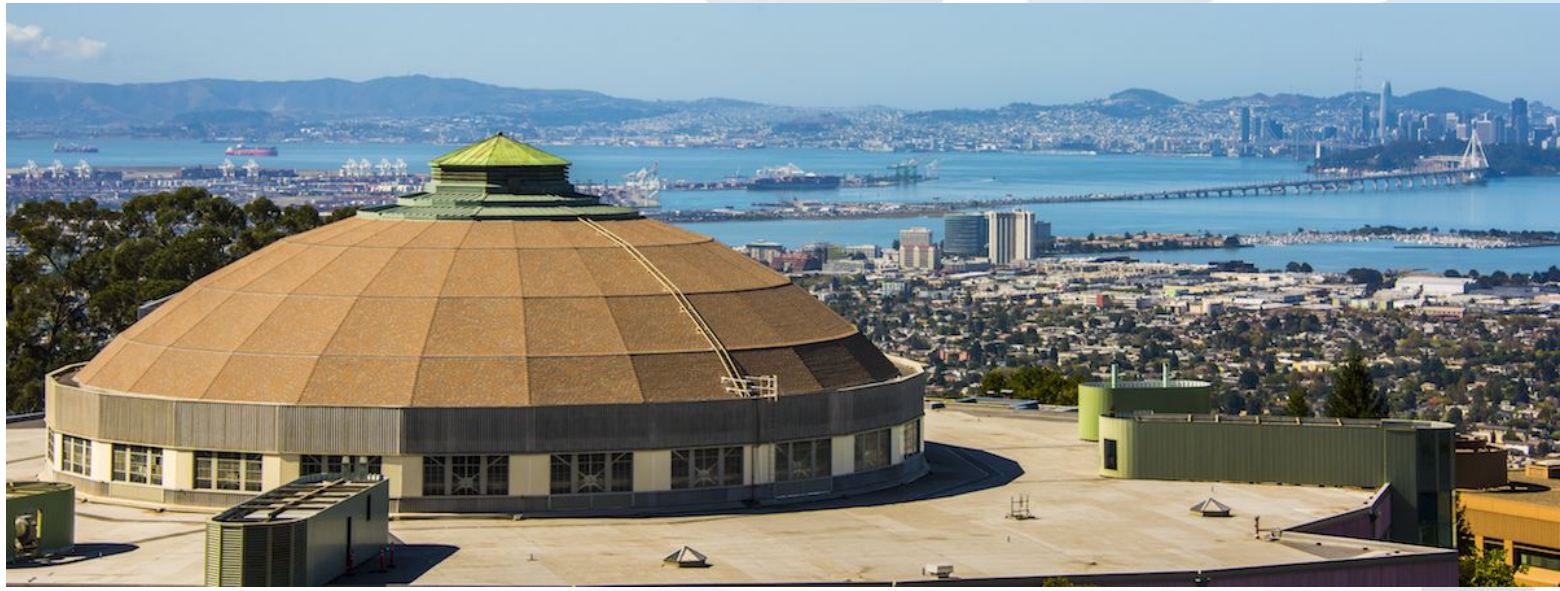
ALS Facility Updates

Dylan McReynolds

ALS Computing Group

ScicatCon June, 2025

ALS



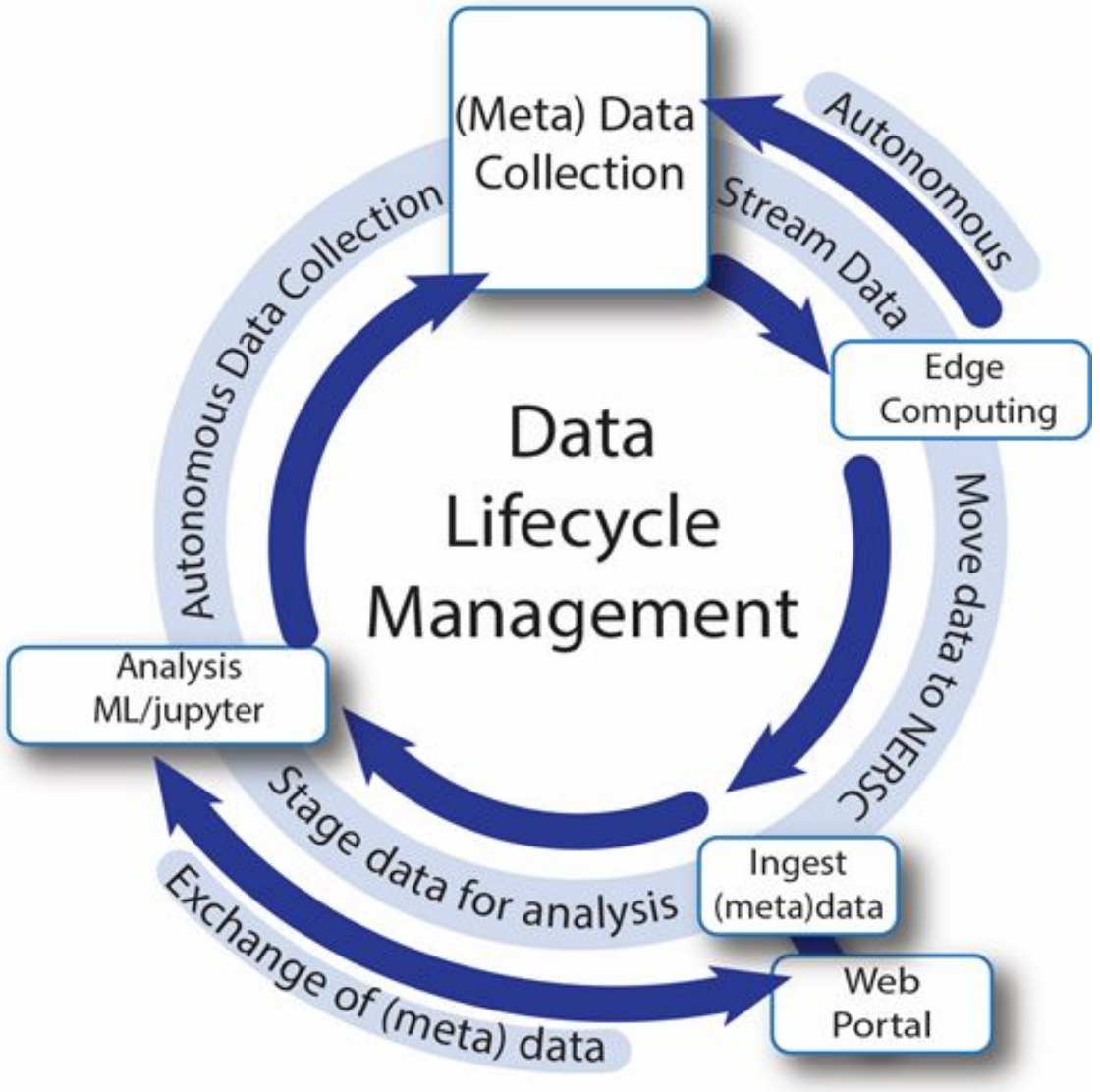
- 40 beamlines
- 2000 users per year
- 3rd generation
- 1 SciCat beamline
- Photon Sciences Computing Group
  - AI/ML
  - Workflows and Visualization
  - Data Infrastructure

ALS-U

- Dark in 2028
- ~ 2yr Darktime
- 1PB a day?

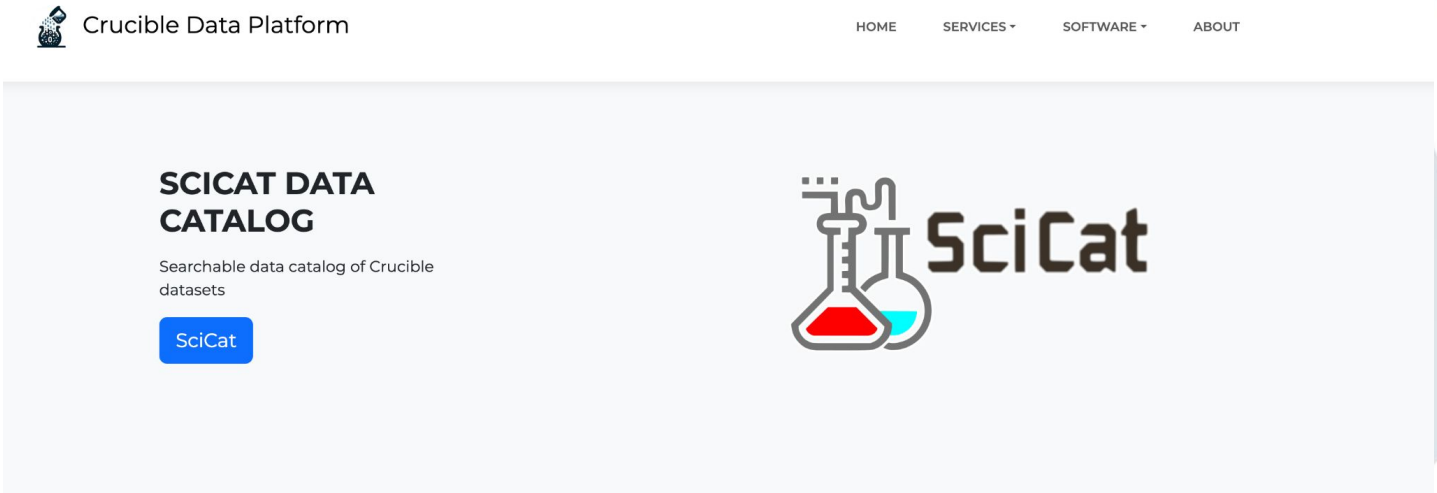
ALS Computing Program

The ALS computing program's mission is to **develop solutions for the mounting data challenge** and **help facilitate compute solutions for the Advanced Light Source**. The computing program **addresses the main challenges of the data lifecycle in close collaboration with the other four DOE light source facilities**.

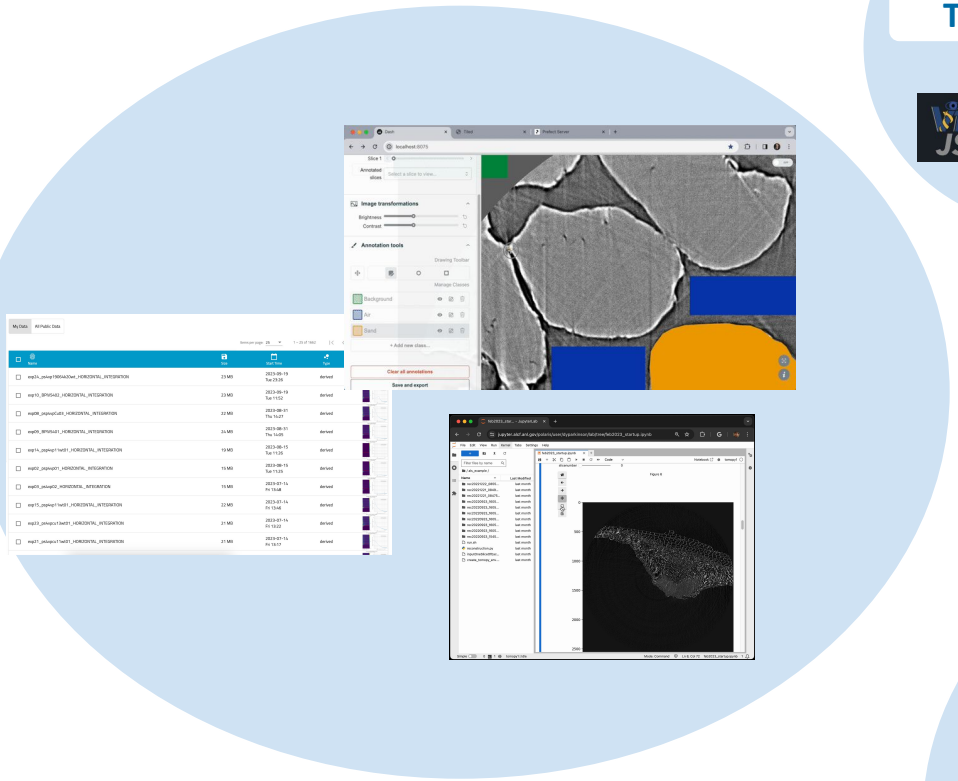


The Molecular Foundry

- Sibling facility
- Synthesis / NCEM
- 25% of ALS users are also TMF users
- Crucible Data Platform



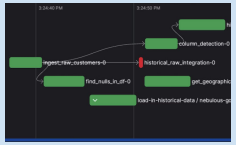
Data Orchestration



Tiled

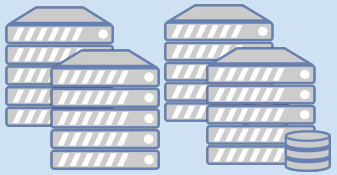
jupyter

plotly



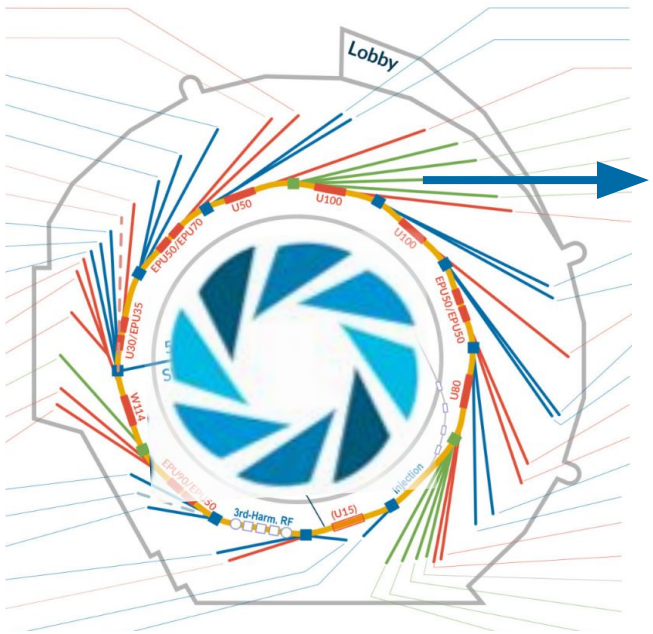
NERSC

beamline



Argonne NATIONAL LABORATORY

ASCR IRI



Measure

Raja S. and Wiebke K. prepare a sand sample and take a measurement.

globus

Computing Workflow

NERSC

Run full-scale reconstruction at NERSC on-demand, in near real-time.

- File-based
- Streaming

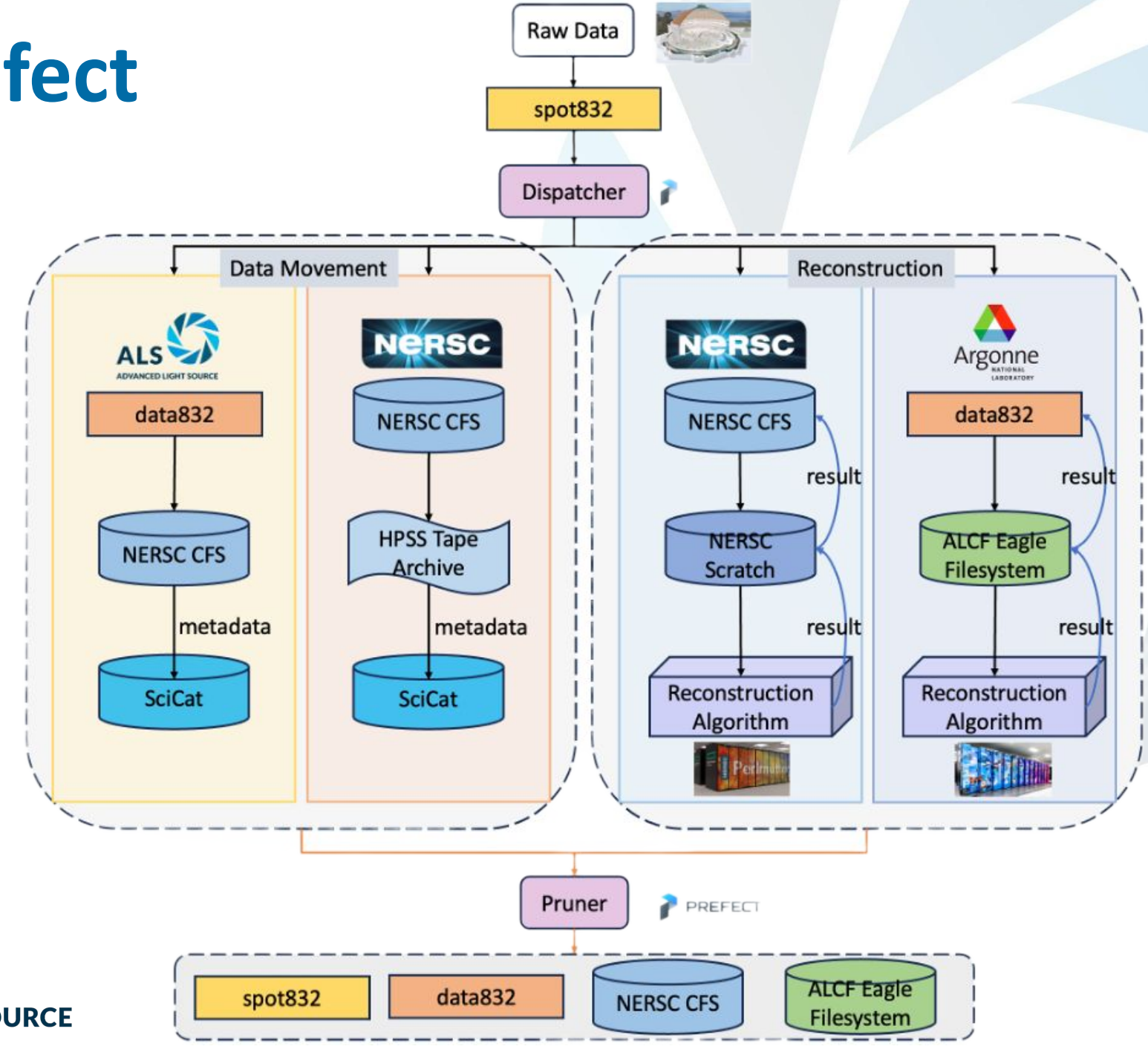
globus

Visualize Results

kitware jupyterhub

View processed scans with intuitive data-driven interactive web applications.

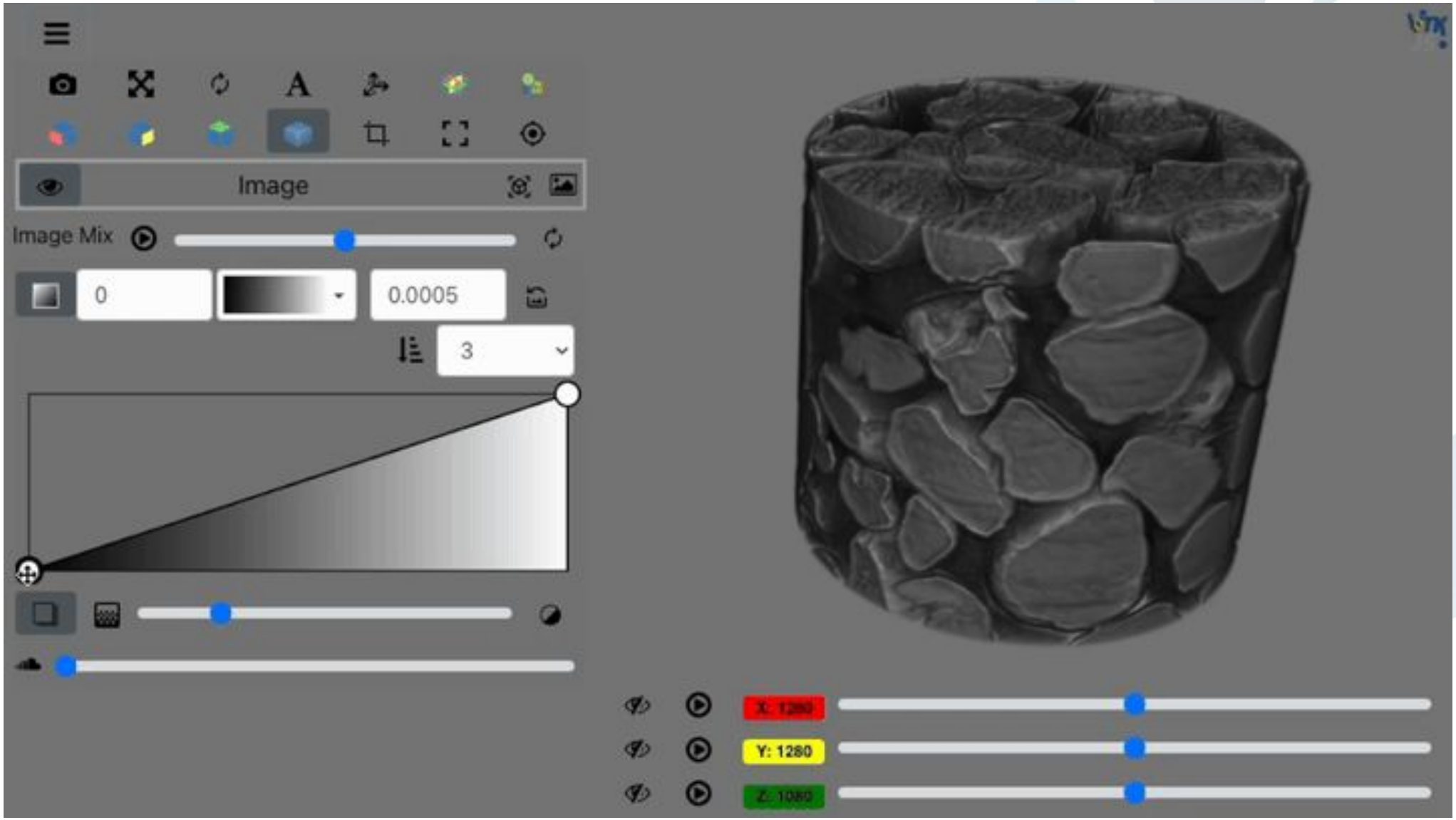
Tomography Prefect Workflow



Workflow Example

- Copy data to HPC
- Ingest metadata into **SciCat**
- Index data in and metadata into **Tiled**
- Run an automated reconstruction
- Create **multi-scale volume pyramid** for visualization
- AI enhanced segmentation
- Copy reconstruction to mid- and long-term storage
- Schedule pruning from intermediary file systems

Visualizations



DOE Data Mandate

- All Data taken at facilities must be open by the end of the 2025
- Will need some catalogs
- A lot of focus on AI

SciCat at ALS by the end of 2025

- 8.3.2 Tomography
- 7.0.1.2 Ptychography
- 7.3.3 Scattering
- 9.3.1 AP-XPS
- 11.0.1.2 AP-XPS/Scattering

Sample Metadata Tracker

radius

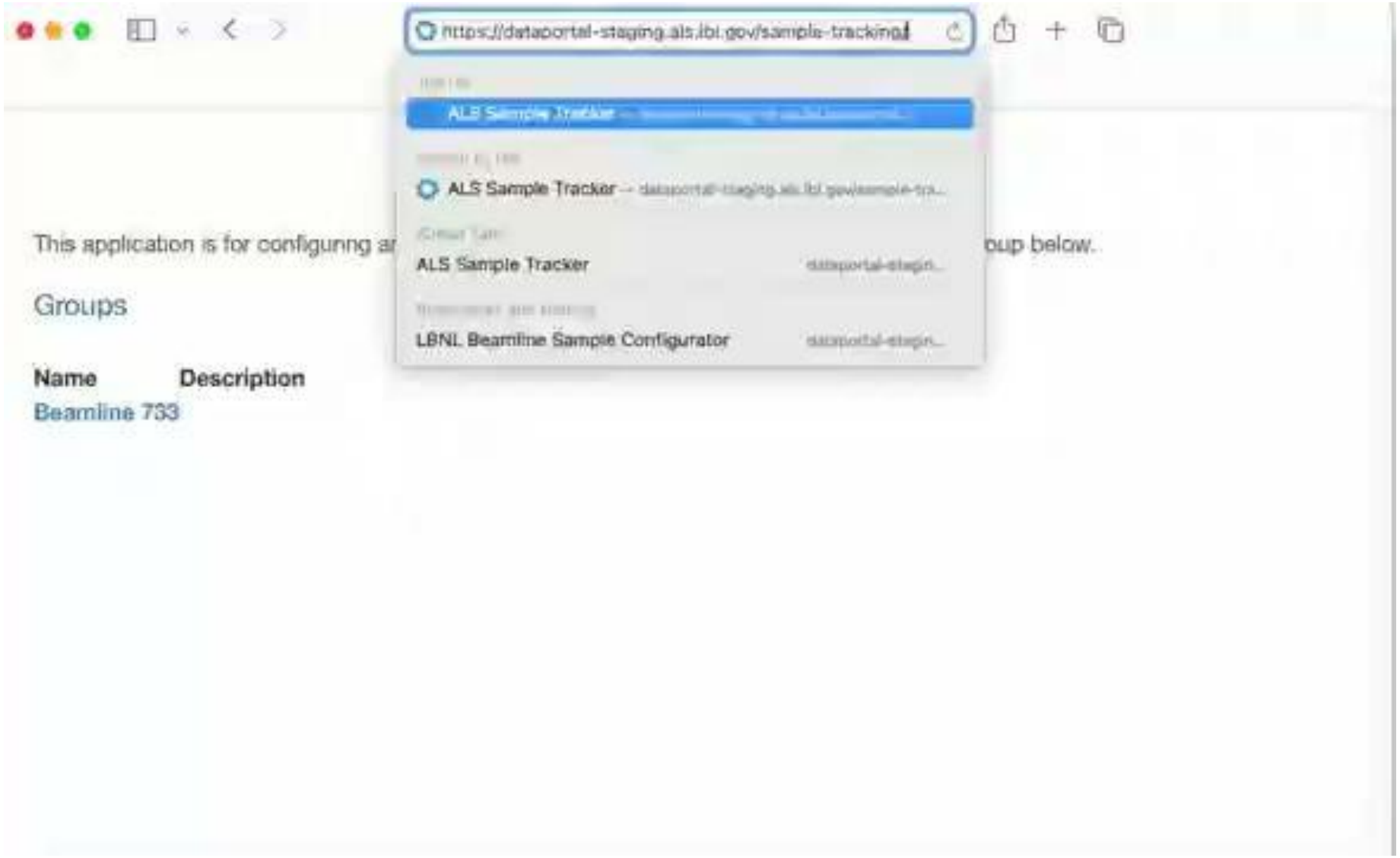
Rapid Deployment of Instrumentation for User Science

Photon Science small projects
Spring 2024

- Website where scientists can enter sample parameters in advance, before arriving at ALS.
- Each sample (and sample group) gets a unique identifier than can be used for all downstream operations.
- Identifier can be printed onto a label and attached to the samples.
- Robots can read the identifier and use it to fetch scan parameters automatically.
- We use SciCat to store data, and provide sample Python code to fetch it.

Sample Metadata Tracker

- Users log in with ORCiD
- They choose a proposal they have access to.
- Each beamline defines its own parameters.
- Interface validates data, and supports cut/paste and undo/redo.
- Labels print to a standard sheet.
- Persisted into SciCat Samples



Links Enhancement

- Backend Issue #689
- Link templates
- Metadata matching
- Add full links to the detail view
- Currently configuration file
  - selector-like matching based on dataset metadata
- Coordinate with Actions (selection, handlebars)
- Configuration management?

pyscicat

- Fixes for new backend login
- Still thinking about using python SDK

Short Term

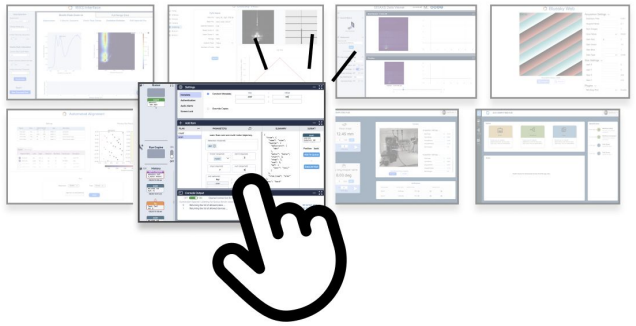
- More beamlines!
- Ingest Derived Data into SciCat
- Ontology / Schema definition and validation
- Integration with more data services
  - Tiled
  - Globus
  - GDrive (TMF Crucible)
- Visualizations
- Migrate deployment to cloud k8s (GKE)
- Implement Elastic

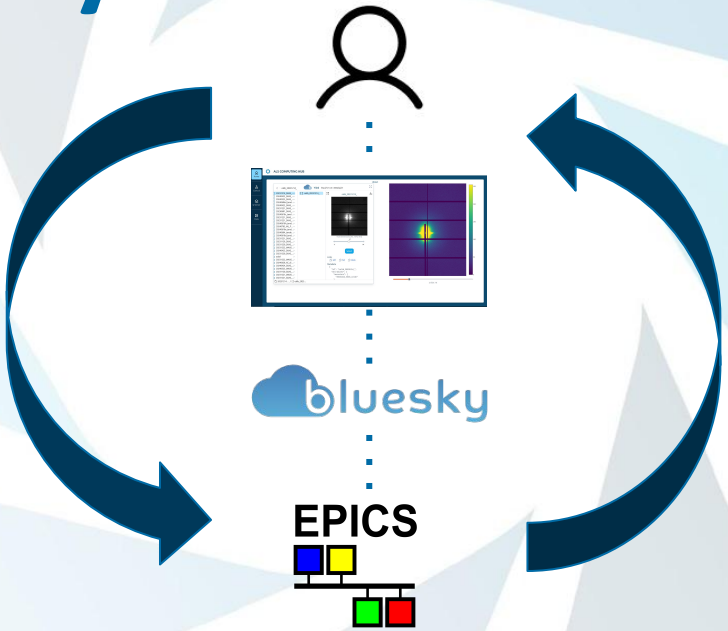
Further Out

- Data Access (globus, better application, tiled)
- Schemas
- Tagging/Labeling
- Multiple Search Engines
  - Elastic
  - LLMs
    - SciCat Agent
  - Federated Search

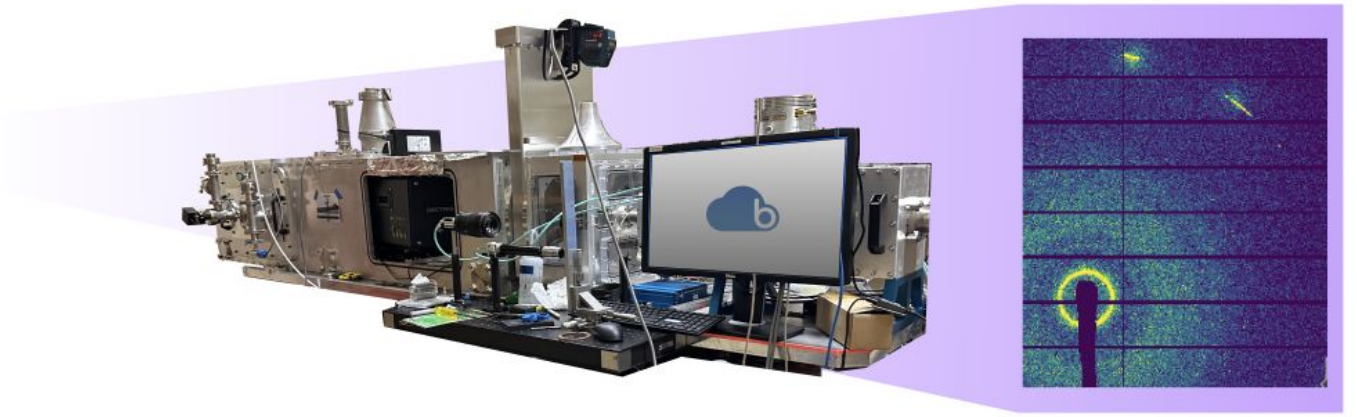
Web Based Beamline Controls/Analysis

Modularized React components from Finch UI library

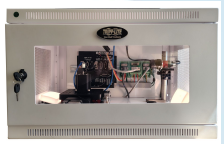


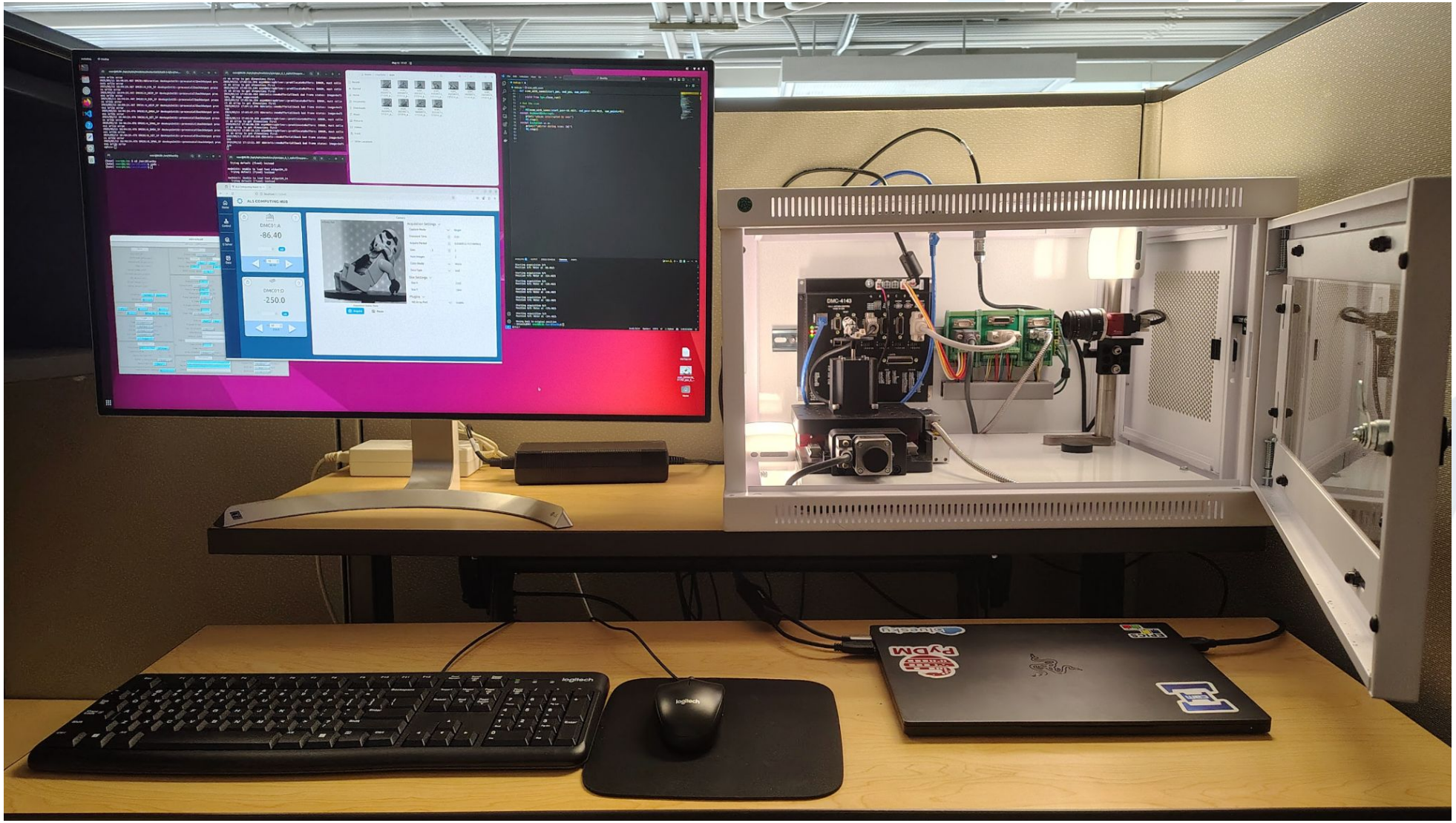


Experimental Orchestration and control with Bluesky



BOLT - Beamline Optics Learning Testbed





Thank You!

ALS

Alex Hexemer
Dula Parkinson
Garrett Birkel
David Abramov
Seij De Leon
Tanny Chavez
Wiebke Köpp
Juliane Reinhardt
Jason Jedd
Tracy Mattox
Mel Sibony
Andrea Jones
Lee Yang
Damon English
Kuldeep Chawla
Karen Fernsler

Scicat Team

All of you!

CRD

Harinarayan Krishnan

NERSC

Bjoern Enders
Spin Team