

# **NOBUGS 2022**

Monday, 19 September 2022 - Thursday, 22 September 2022

Paul Scherer Institute, Villigen, Switzerland



## **Book of Abstracts**



# Contents

A hitchhiker's guide to the easyScience galaxy . . . . .	1
scipp: Reduction of large event data . . . . .	1
Current state of data analysis at the Ncsa Neutron Diffraction Facility . . . . .	2
The Scientific Computing Strategy for the Upgraded Advanced Photon Source . . . . .	2
H5Web: a web viewer for HDF5/NeXus files . . . . .	3
DARTS on-demand computing : heavy-duty data treatment for all . . . . .	4
pyStxm: A python application for STXM data collection using Qt, BlueSky and Epics . . . . .	5
C2 Data Viewer: Visualization tool for EPICS7 Data Streaming . . . . .	6
McXtrace: simulating X-ray beamlines and experiments, with samples . . . . .	7
Recent Developments in the Fit-Benchmarking Python Package . . . . .	9
Lossless data compression: an overview . . . . .	10
ECXAS: A data aggregation tool for battery study in ROCK beamline . . . . .	10
Monte Carlo Multiple Scattering Corrections In Neutron Scattering Experiments . . . . .	11
Legacy of scientific code . . . . .	11
Polarized diffraction and spectroscopy data reduction in Mantid . . . . .	12
Graphical User Interfaces for ILL users in Mantid . . . . .	12
Into the future: ILL Endurance program . . . . .	13
NOMAD OASIS, a Laboratory Data Management Platform . . . . .	13
Experimental Data Infrastructure with BENTEN for Fuel Cell Project at SPring-8 . . . . .	14
Processing Neutron Time-of-Flight Laue Data in DIALS . . . . .	15
The Mamba Data Acquisition Software Project for HEPS . . . . .	15
Minimalist beamline control and experiment software . . . . .	16
Ophyd v2 . . . . .	17

Recent developments in the MSlice software package . . . . .	17
XRD2: MX beamline at Elettra - A mix recipe . . . . .	18
3D Grid Scans: A Bluesky prototype at Diamond Light Source . . . . .	19
Building a PaN Data Commons on the outcomes of the PaNOSC and ExPaNDS EOSC projects . . . . .	19
Daiquiri: a web based user interface framework for beamline control and data acquisition	20
Multi-purpose Regulation Loops in BLISS . . . . .	20
py-ISPyB - A new implementation of a LIMS for experiments in structural biology . . . .	21
Experience with developing FPGA accelerated data reduction in DevCloud . . . . .	21
The Blissdata Project: new perspectives on BLISS Data Management . . . . .	22
NXtomomill: more than just a data conversion tool . . . . .	22
Unified access to photon and neutron data throughout Europe - the SciCat implementation of the PaNOSC Search . . . . .	23
Automated Scientific Metadata Recording and Viewing During Experiments at MAX IV .	23
Current and future developments of European XFEL scan tool Karabacon . . . . .	24
Overview of BLISS, the ESRF-EBS beamline control system . . . . .	25
diffcalc-core: Diffraction condition calculation package for a six-circle diffractometer . .	25
Image Annotation at European XFEL . . . . .	26
Data-Modelling Patterns for Experimental Characterisation . . . . .	26
The Karabo Control System . . . . .	27
Zocalo: a high-throughput data processing framework . . . . .	28
Elettra.chat: from Chat Service to Logbook For Experiments . . . . .	29
A SEDCNN Machine Learning Model for Textured SAXS/WAXD Image Denoising . . . .	29
Taking the pain out of generating complete and compliant NeXus files for dynamic setups	30
An Integrated Data Acquisition and Analysis Software for XRF Mapping Experiments . .	31
Synchronization of commercial camera data at the European XFEL . . . . .	31
Experiences with Datalogging to InfluxDB at the European XFEL . . . . .	32
KaraboGui - The Cockpit of the Supervisory Control and Data Acquisition System Karabo	32
User interfaces at the PolLux STXM beamline . . . . .	33
SciCat Present and Future . . . . .	33

SDU: Software for high throughput automated data collection at the Swiss Light Source	34
Adopting NICOS at SINQ . . . . .	35
A MHz sampling DAQ system for sub-second QEXAFS at the SLS-2.0 "Debye" Beamline	35
Next generation User Office software . . . . .	36
PaN-Training e-Learning: education and training for scientists and students . . . . .	37
Data streaming at SINQ . . . . .	37
Ewoks –a meta-workflow system . . . . .	38
Mantid Imaging . . . . .	38
Environmental sustainability for scientific software . . . . .	39
Fast analysis feedback with automated data processing pipelines at PETRA-III . . . . .	39
Towards real-time data reduction in serial-crystallography . . . . .	40
Experiment data streaming and aggregation at the European Spallation Source . . . . .	41
The new tomography data processing software of the ESRF . . . . .	41
Scientific Data Management at European XFEL. . . . .	42
High-Rate Data Acquisition, Streaming, and Processing at the EMBL PETRA III Beamlines	42
A common control and readout software for different X-ray detector systems . . . . .	43
Bluesky at the Advanced Photon Source . . . . .	44
Muon Galaxy –an open web platform for computational muon science . . . . .	45
Muon Galaxy –an open web platform for computational muon science . . . . .	45
Machine-learning driven beamline alignment at EuXFEL . . . . .	46
Managing Experiment Data with Ease at the Advanced Photon Source . . . . .	47
Unified recording of Channel Access and beam-synchronous data for PSI . . . . .	48
EasyDiffraction: A user-friendly interface for diffraction data analysis . . . . .	48
Scaling Diffuse Scattering Workflows with Hybrid HPC Workflows . . . . .	49
The UK Ada Lovelace Centre . . . . .	50
CamServer: Data Processing Pipelines @ SwissFEL . . . . .	50
smargopolo: A new control system for SmarGon/MCS2 using ROS . . . . .	51
Integrated Real-Time Auto-Processing at Diamond Light Source . . . . .	52
DIALS as a Toolkit . . . . .	52

PyBluIce: Modular data acquisition software for long duration, scripted automation, and high speed protein crystallography experiments . . . . .	53
SOLEIL digital transformation toward the upgrade of the facility . . . . .	53
SciLog –An Electronic Logbook for User Experiments . . . . .	54
SLS 2.0 Controls and Science IT Sub Project Status Update . . . . .	54
#29 Daiquiri: a web based user interface for beamline control and data acquisition . . . .	55
#22, The Mamba Data Acquisition Software Project for HEPS . . . . .	55
Welcome by Thierry Straessle, PSI Deputy Director . . . . .	55
Discussion and NOBUGS 2022 - Closing . . . . .	55
C2 Data Viewer: Visualization tool for EPICS7 Data Streaming . . . . .	55

2

## A hitchhiker's guide to the easyScience galaxy

**Author:** Simon Ward<sup>1</sup>

**Co-authors:** Andrew Sazonov<sup>2</sup>; Andrew McCluskey<sup>2</sup>; Piotr Rozyczko<sup>2</sup>

<sup>1</sup> ESS - DMSC

<sup>2</sup> ESS

**Corresponding Author:** simon.ward@ess.eu

**easyScience**[1] is an initiative from the European Spallation Neutron Source (ESS) to unify simulation software across neutron scattering. **DON'T PANIC!** While this goal seems an unsurmountable challenge, it is achievable as demonstrated by our current releases. The **easyScience** project has the following aims; Provide a unified method to interact the most popular technique specific simulation software/libraries, a professional and welcoming graphical interface for new users, *Jupyter* notebooks for experienced users, unified data structures and workflows across multiple techniques. As an opening to this project, diffraction and reflectometry techniques were chosen to demonstrate the easy philosophy. These techniques have multiple complex calculation engines available, which it is unrealistic to expect users to master. **easyReflectometry** and **easyDiffraction** unifies these calculation engines for their respective techniques and provides a complete, feature rich and easy to use interface. In the future QENS and spectroscopy will also be targeted. As a bonus, the technologies behind the **easyScience** programs allow for advanced modelling and statistical analysis techniques with the ability to scale for large datasets.

Behind these programs is **easyCore**, a unified simulation, optimisation and analysis package. easy-Core is built on the latest techniques and libraries including *scipp* (developed at ESS) for dataset handling, *jax* for machine learning and *PyMC* for Bayesian analysis. Hence all these features are available for all **easyScience** software. We present the main features of **easyScience**, where it came from, where it's going and how it will be used to enhance the analysis workflow with the latest analysis techniques.

[1] <https://github.com/easyScience>

**Email address of presenting author:**

simon.ward@ess.eu

**I agree to recordings of my presentation being made at NOBUGS 2022:**

3

## scipp: Reduction of large event data

**Author:** Jan-Lukas Wynen<sup>1</sup>

**Co-authors:** Neil Vaytet<sup>1</sup>; Simon Heybrock<sup>1</sup>

<sup>1</sup> European Spallation Source ERIC

**Corresponding Author:** jan-lukas.wynen@ess.eu

scipp is a software stack developed at ESS with in-kind contributions from ISIS for the reduction of neutron scattering data. It consists of multiple Python libraries:

- *scipp*: General purpose multi-dimensional arrays with labelled dimensions, physical units, and support for non-destructive binning of event data.
- *scippnexus*: Low-level utilities for reading and writing NeXus files.

- *scippneutron*: Cross-facility routines for neutron data reduction.
- *ess*: Functionality bespoke to ESS.

scipp's lean and flexible data structures allow it to scale to larger data than other software. In addition, its high-level design prevents many common mistakes and makes scipp accessible to people with little programming experience.

**Email address of presenting author:**

jan-lukas.wynen@ess.eu

**I agree to recordings of my presentation being made at NOBUGS 2022:**

4

## Current state of data analysis at the Necsca Neutron Diffraction Facility

**Author:** Deon Marais<sup>1</sup>

<sup>1</sup> *The South African Nuclear Energy Corporation (Necsca) SOC Limited*

**Corresponding Author:** deon.marais@necsca.co.za

This presentation will give an overview of the current state of ScanManipulator, the data reduction, analysis, and visualisation software implemented at the Neutron Diffraction Facility at the SAFARI-1 research reactor in South Africa. ScanManipulator was developed in-house to automate a number of processes associated with the Materials Probe for Internal Strain Investigation (MPI SI) and Powder Instrument for Transition in Structure Investigations (PITSI) angular dispersive neutron diffractometers.

Notable functionalities include: automated flat field, geometric, 2D and 3D neutron attenuation correction; data combining; sample surface determination through analytical entry curve functions; Gauss and Voigt peak fitting procedures; 2D and polar plots through Matplotlib; 3D plots through Mayavi; artificial neural networks through Fann2; data export to Excel, FullProf and GSAS; near real time peak fitting and statistical analysis to optimise counting time.

An outlook on the future Neutron Beam Line Centre at the envisaged new Multi-Purpose Reactor will also be given.

**Email address of presenting author:**

deon.marais@necsca.co.za

5

## The Scientific Computing Strategy for the Upgraded Advanced Photon Source

**Author:** Nicholas Schwarz<sup>1</sup>

<sup>1</sup> *Argonne National Laboratory*

**Corresponding Author:** nschwarz@anl.gov



The Advanced Photon Source (APS) at Argonne National Laboratory (ANL) will replace the entire storage ring with a ring based on a multi-bend achromat lattice design. The new storage ring will increase the APS's brilliance by factors of 100-1,000s, depending on x-ray energy, and make the APS the brightest hard x-ray synchrotron source in the world. Because of the greatly enhanced brightness, coherence, and signal at high x-ray energies along with new state-of-the-art high-bandwidth commercial detectors, beamlines require significant improvements in networking, controls and data acquisition, computing, workflow, data reduction, and analysis tools to operate effectively.

All aspects of APS operation depend on computation, but data analysis software and beamline control and computing infrastructure are of particular importance for facility productivity. Demands for increased computing at the APS are driven by new scientific opportunities, which are enabled by new measurement techniques, technological advances in detectors, multi-modal data utilization, and advances in data analysis algorithms. The priority for the APS is to further improve its world-class programs that benefit most from high-energy, high-brightness, and coherent x-rays. All of these require advanced computing. The revolutionized high-energy synchrotron facility that the APS will deliver will increase brightness and coherence, leading to further increases in data rates and experiment complexity, creating further demands for advanced scientific computation.

Over the next decade, the APS anticipates a multiple-order-of-magnitude increase in data rates and volumes generated by APS instruments. This necessitates 10s of petaflop/s of on-demand computing resources and increased data management and storage resources to process and retain this data and the analyzed results. Advanced data processing and analysis methods will be required to keep up with the anticipated data rates and volumes and to provide real-time experiment steering capabilities.

The key elements of the scientific computing strategy at the upgraded APS, include upgrading networking infrastructure within the APS and between the APS and the Argonne Leadership Computing Facility (ALCF), deploying state-of-the-art experiment control software at beamline instruments, expanding the capabilities and use of common data management and workflow tools and science portals, deploying sufficient local and edge computing resources, and utilizing new supercomputers at the ALCF for large on-demand data processing and analysis tasks, developing high-speed, highly parallel data processing and analysis software, and extensively applying novel mathematical and AI/ML methods to solve challenging data reduction and analysis problems, and collaborating with other light sources, experimental facilities, large-scale computing and networking facilities, and the APS user community.

\*Work supported by U.S. Department of Energy, Office of Science, under Contract No. DE-AC02-06CH11357.

**Email address of presenting author:**

nschwarz@anl.gov

6

## H5Web: a web viewer for HDF5/NeXus files

**Authors:** Loïc Huder<sup>1</sup>; Axel Bocciarelli<sup>1</sup>

<sup>1</sup> ESRF

**Corresponding Author:** loic.huder@esrf.fr

HDF5 (with NeXus) is becoming the standard in many X-ray facilities. HDF5 viewers are needed to allow users to browse and inspect the hierarchical structure of HDF5 files, as well as visualize the datasets inside as basic plots (1D, 2D).

H5Web is a web-based HDF5 viewer developed at the ESRF to fulfill this need. Designed with remote access and modularity in mind, it aims to provide easy access to the data and high interactivity to

users with performant WebGL visualisations.

This presentation will demonstrate some of H5Web's features, including NeXus support, and show examples of where it can be used, including in the browser (demo), in JupyterLab, and in Visual Studio Code.

This project has received funding from the European Union's Horizon 2020 research and innovation program under grant agreement No. 823852.

**Email address of presenting author:**

loic.huder@esrf.fr

7

## **DARTS on-demand computing : heavy-duty data treatment for all**

**Author:** Emmanuel FARHI<sup>1</sup>

<sup>1</sup> *Synchrotron SOLEIL*

**Corresponding Author:** emmanuel.farhi@synchrotron-soleil.fr



Figure 1: DARTS

The **Data Analysis Remote Treatment Service (DARTS)** [1] is a remote desktop service that launches virtual machines in the cloud, and displays them in your browser.

These machines can be used for e.g. scientific data treatment (reduction/analysis). A user indicates how many CPU cores and how much memory he wishes, can optionally request a GPU for heavy-duty computations, and can be customized at start-up. The sessions typically start in about 10 seconds, and can be shared with colleagues within their life-time (typically 7 days). Each user can manage all his active sessions (connect, stop, share).

This service can be installed and configured within a few minutes (`sudo apt install qemu-web-desktop`), and it requires minimal maintenance. The source code is fully open-source, and relies on very simple elements (Apache web-server and QEMU). Give it a try !

We provide at Synchrotron SOLEIL a set of prepared environments with plenty of scientific software for e.g. crystallography, spectroscopy, imaging, tomography, etc. The usual environment is based on Debian/Ubuntu with hundreds of pre-installed scientific software and libraries. A Windows 10 system is also available. A specialized environment has also been set-up for AlphaFold protein structure prediction. These environments are available on the beam-lines, and connected to our data storage.

---

[1] DARTS <https://gitlab.com/soleil-data-treatment/soleil-software-projects/remote-desktop>

**Email address of presenting author:**

emmanuel.farhi@synchrotron-soleil.fr

8

## pyStxm: A python application for STXM data collection using Qt, BlueSky and Epics

**Author:** Russ Berg<sup>1</sup>

**Co-authors:** A.F.G Leontowich<sup>1</sup>; chithra karunakaran<sup>1</sup>; Jian Wang<sup>1</sup>

<sup>1</sup> *Canadian Light Source*

**Corresponding Author:** russ.berg@lightsources.ca

### Abstract

Traditionally, data collection software written by individuals whose primary responsibility is not software have prioritized the User Interface (UI) of their software somewhere near the bottom as this often can become very time consuming as well as the source of many software bugs. For software developers whose primary responsibility it is to deliver software the goal should be to produce reliable software that allows users of all experience levels the ability to collect data in a timely manner, automate monotonous and error prone tasks as well as to provide conveniences where possible that promote user efficiency while at the beamline. A fully featured user interface is key to this delivery. pyStxm is a data collection application that is developed with this motivation in mind, and is the user interface for STXM (Scanning Transmission X-ray Microscopy) data collection developed at the Canadian Light Source 10ID1 beamline. Not wanting to reinvent the wheel the user interface incorporates ideas inspired from successful existing commercial and open source software, primarily Adobe Photoshop [5] and the open source 3D animation software project, Blender[4]. These two software applications were used as inspiration because of their ability to provide a user interface that was able to organize complex data that would scale with time into panels and areas of the screen that facilitated work flow as well as user learning. Not only do these two particular applications do a great job at organizing complex data they also allow for that complex data to scale with future feature enhancements. Along with user efficiency the goal was also standardization of the data file

format that produces NEXUS[6] files that conform to the NXstxm[7] NEXUS application definition. pyStxm is developed in python and uses several freely available frameworks. Qt [3] is a fully featured application development framework and is used as the basis for the user side of the application while the BlueSky [1] framework from Brookhaven National Lab is used to manage scanning and interface to the underlying distributed control system which is EPICS 2. What is presented are some of the key features of pyStxm that were implemented to attempt to facilitate user efficiency during data collection.

#### References

1. Daniel Allan, Thomas Caswell, Stuart Campbell & Maksim Rakitin. Bluesky's Ahead: A Multi-Facility Collaboration for an a la Carte Software Project for Data Acquisition and Management. Synchrotron Radiation News, Volume 32, 2019 –Issue 3.
2. Dalesio, L.R., Kozubal, A.J., Kraimer, M.R. EPICS architecture (Conference: International conference on accelerator and large experimental physics control systems, Tsukuba (Japan), 11-15 Nov 1991)
3. Qt application framework, <https://www.qt.io>
4. Blender open source 3D animation creation suite, <https://www.blender.org/>
5. Adobe Photoshop, <https://helpx.adobe.com/ca/photoshop/user-guide.html>
6. Mark Könnecke, Frederick A. Akeroyd, Herbert J. Bernstein, Aaron S. Brewster, Stuart I. Campbell, Björn Clausen, Stephen Cottrell, Jens Uwe Hoffmann, Pete R. Jemian, David Männicke, Raymond Osborn, Peter F. Peterson, Tobias Richter, Jiro Suzuki, Benjamin Watts, Eugen Wintersberger and Joachim Wuttke. The NeXus data format. Journal of Applied Crystallography. Volume 48, Part 1, February 2015, Pages 301-305
7. Benjamin Watts and Jörg Raabe, A NeXus/HDF5 based file format for STXM, AIP Conference Proceedings 1696, 020042 (2016); <https://doi.org/10.1063/1.4937536> Published Online: 28 January 2016

**Email address of presenting author:**

[russ.berg@lightsources.ca](mailto:russ.berg@lightsources.ca)

9

## C2 Data Viewer: Visualization tool for EPICS7 Data Streaming

**Authors:** Elaine Chandler<sup>1</sup>; Guobao Shen<sup>1</sup>; Sinisa Veseli<sup>1</sup>; Thomas Fors<sup>1</sup>

<sup>1</sup> Argonne National Laboratory

**Corresponding Author:** [echandler@anl.gov](mailto:echandler@anl.gov)

A high-performance data acquisition system (DAQ) has been under active development to meet APS-U needs. It takes data from underneath FPGA (Field Programmable Gate Array) and streams it to its downstream users. The APS-U DAQ system software framework is implemented as a major portion of APS-U new control system software infrastructure, which is called C2. To visualize the DAQ data on the fly, C2 Data Viewer (C2DV) is implemented using Python, which can be used for displaying live PV data streams for monitoring, troubleshooting and diagnostics purposes. It is now capable of handling both EPICS pvAccess (PVA) and Channel Access (CA) data and includes several different applications: a scope viewer for plotting PVA waveforms, an image viewer for displaying Area Detector image data, and a striptool for monitoring PVA as well as CA scalar PVs. In this presentation, we discuss various C2DV features, its usage at the Advanced Photon Source, as well as plans for future development.

**Email address of presenting author:**

echandler@anl.gov

I agree to recordings of my presentation being made at NOBUGS 2022:

10

## McXtrace: simulating X-ray beamlines and experiments, with samples

**Author:** Emmanuel FARHI<sup>1</sup>

**Co-authors:** Antoine PADOVANI<sup>1</sup>; Stéphane BAC<sup>1</sup>

<sup>1</sup> *Synchrotron SOLEIL*

**Corresponding Author:** antoine.padovani@synchrotron-soleil.fr

**McXtrace** <https://www.mcxtrace.org> is a general Monte Carlo ray-tracing software for simulation X-ray beamlines and experiments. It benefits from the acquired experience gained from the McStas <https://www.mcstas.org> neutron code.

Compared to other X-ray modelling software (SRW, Shadow, XRT), McXtrace has been built in a modular way, allowing anybody to **contribute** with minimal involvement. Also, **sample models** are included (SAXS, MX, XRD, XAS, Tomography, and soon to come IXS). McXtrace can handle beam coherence, and runs on clusters and GPU's.

We have modelled a set of source-to-detector beam-lines at Synchrotron SOLEIL

- ANATOMIX (tomography )
- DISCO (imaging, UV)
- MARS (powder diffraction, XRD)
- PX2a (protein crystallography, MX)
- ROCK (absorption spectroscopy, XAS)
- SWING (small angle, SAXS)

We shall present the work done in our group, with these BL, as well as contributed components for the package.

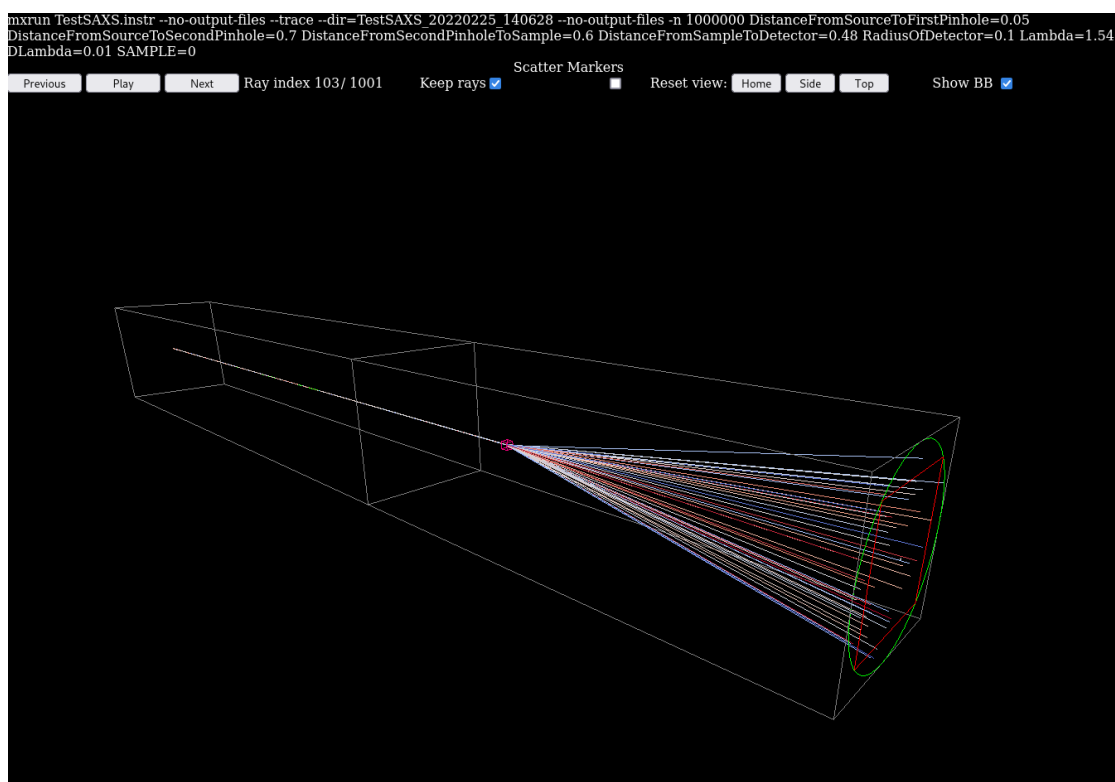


Figure 2: McXtrace: SAXS model

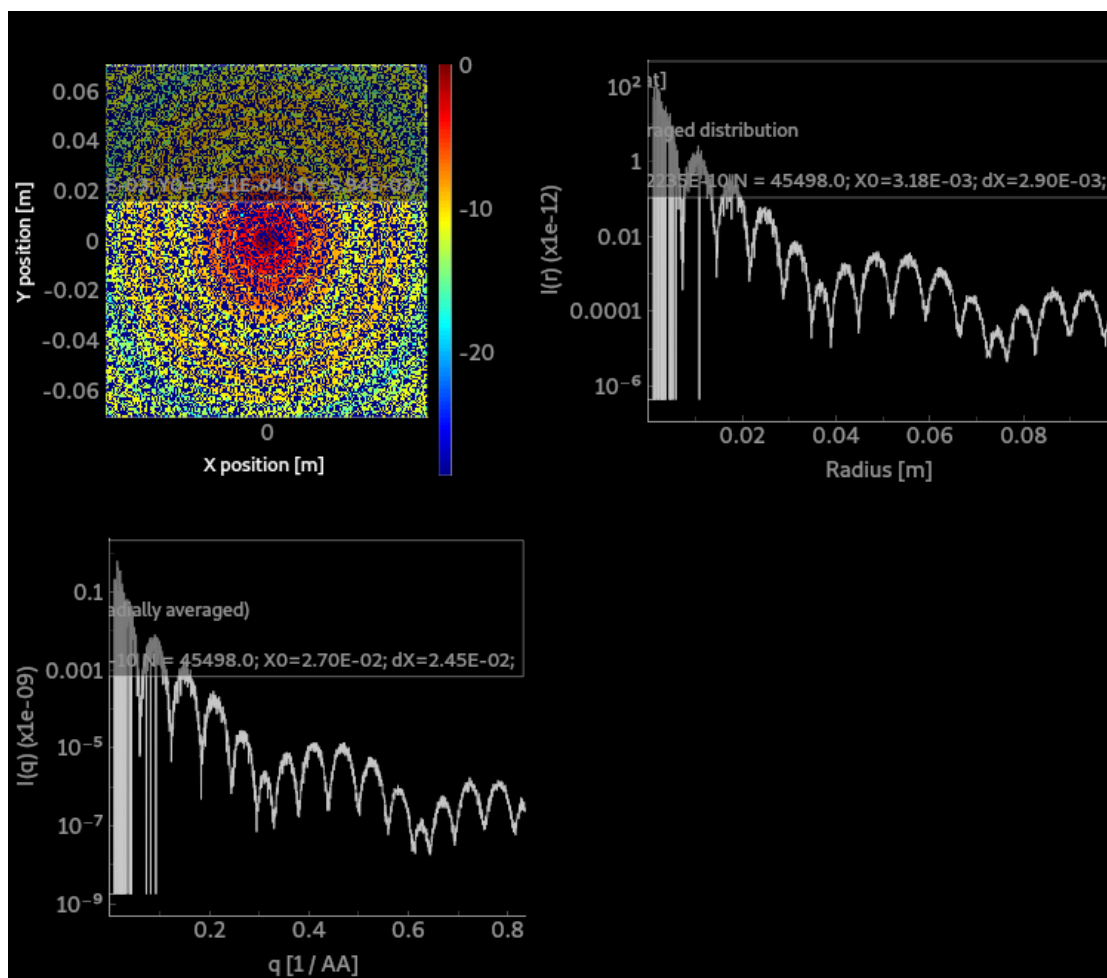


Figure 3: McXtrace: SAXS image

Email address of presenting author:

emmanuel.farhi@synchrotron-soleil.fr

I agree to recordings of my presentation being made at NOBUGS 2022:

11

## Recent Developments in the Fit-Benchmarking Python Package

Author: Robert Applin<sup>1</sup>

<sup>1</sup> STFC

Corresponding Author: robert.applin@stfc.ac.uk

A poster about the recent developments and future plans for the FitBenchmarking python packages developed by the Science and Technology Facilities council, in collaboration with Diamond Light Source.

Email address of presenting author:

robert.applin@stfc.ac.uk

I agree to recordings of my presentation being made at NOBUGS 2022:

12

## Lossless data compression: an overview

**Author:** Gamil CASSAM-CHENAI<sup>1</sup>

<sup>1</sup> *Synchrotron SOLEIL*

**Corresponding Author:** [gamil.cassam-chenai@synchrotron-soleil.fr](mailto:gamil.cassam-chenai@synchrotron-soleil.fr)

Large research infrastructures, such as synchrotron facilities, generate large amounts of data (up to a few tens of terabytes) every day. This data is very valuable, as it is the result of elaborate scientific experiments, which are likely to be performed only once.

Storing this data efficiently in addition to previously accumulated data, being able to transfer it quickly, and accessing it efficiently for visualization and scientific analysis is both a necessity and a challenge that digital data compression can address.

In this review of lossless data compression, I will present the metrics to use when considering compression from a temporal perspective, some strategies for improving compression and a few tools for evaluating compression algorithms with an example based on tomography data obtained at Soleil.

**Email address of presenting author:**

[gamil.cassam-chenai@synchrotron-soleil.fr](mailto:gamil.cassam-chenai@synchrotron-soleil.fr)

13

## ECXAS: A data aggregation tool for battery study in ROCK beam-line

**Author:** Lucía PÉREZ RAMÍREZ<sup>1</sup>

**Co-authors:** Emmanuel FARHI <sup>1</sup>; Stéphanie BELIN <sup>1</sup>; Olga ROUDENKO <sup>1</sup>

<sup>1</sup> *Synchrotron SOLEIL*

**Corresponding Author:** [lucia.perez@synchrotron-soleil.fr](mailto:lucia.perez@synchrotron-soleil.fr)

Present-day material characterization requires insight from multiple techniques if a complete understanding of the material's properties and behaviors is desired. This is often the case with synchrotron-based experiments in which, besides the collected data from a beamline itself, additional data is obtained by simultaneous measurements using other experimental probes and under different in situ environments. This multimodal approach along with the continuously increasing time/space resolution of synchrotron measurements simply translates into higher data dimension and larger data volume, which adds to its complexity during the analysis step. A tool for semi-automated data processing becomes thus decisive during and after each experiment.

Here, we present the case example of ROCK beamline in which time-resolved X-ray absorption spectroscopy (XAS) is coupled to electrochemical cycling (EC) for the study of battery materials. We introduce a tool developed in Python to aggregate data from multiple sources and assemble multi-dimensional maps. The tool allows for a continuous representation of XAS spectra during charge/discharge cycles along with the relevant electrochemical information, thus facilitating the



deduction of present phases over the course of the experiment. The initiative is part of the European BIG-MAP (Battery Interface Genome - Materials Acceleration Platform) project.

**Email address of presenting author:**

lucia.perez@synchrotron-soleil.fr

**I agree to recordings of my presentation being made at NOBUGS 2022:**

14

## Monte Carlo Multiple Scattering Corrections In Neutron Scattering Experiments

**Author:** Danny Hindson<sup>None</sup>

**Corresponding Author:** danny.hindson@stfc.ac.uk

A Monte Carlo integration method has been developed in the Mantid data reduction package that calculates the multiple scattering intensity for a given structure factor, sample shape and instrument definition. The calculation works for both elastic and inelastic instruments (both Direct and Indirect geometries).

The calculation is based on the Fortran DISCUS program that was previously developed by Mike Johnson and Spencer Howells at ISIS, initially in the 1970s.

The calculation doesn't require an assumption that the scattering is isotropic and can explicitly calculate all scattering orders without an assumption that the ratio between the orders is constant. The calculation uses the Mantid sample geometry engine which is able to calculate track intersections for arbitrary shape types described using CSG or mesh geometries.

Some results of the calculation on real/synthetic samples will be shared and work on how this calculation can be formulated into a correction process will be presented. Some of the software engineering challenges around performance and the process for porting a valuable legacy Fortran program into a modern data reduction platform will also be discussed.

**Email address of presenting author:**

danny.hindson@stfc.ac.uk

15

## Legacy of scientific code

**Author:** Anthony Lim<sup>1</sup>

**Co-authors:** Martyn Gigg<sup>1</sup>; Stephen Smith

<sup>1</sup> STFC

**Corresponding Author:** anthony.lim@stfc.ac.uk

There are numerous examples of legacy code within the science community. However, these are usually poorly documented and do not meet modern software standards. As time goes by any knowledge of how these codes are supposed to work is lost. Hence, maintaining them is at best difficult and at worst guess work. In this talk, I will present my experience of creating a modern version of the quasilasticbayes code. I will explain why it is beneficial to fail fast during the early days of development and why incremental modernisation improves understanding of the code. I will discuss some

of the main learning points from the modernisation of the code and some of the tools I have found useful.

**Email address of presenting author:**

anthony.lim@stfc.ac.uk

16

## Polarized diffraction and spectroscopy data reduction in Mantid

**Author:** Dominik Arominski<sup>1</sup>

**Co-author:** Gagik Vardanyan

<sup>1</sup> *Institut Laue-Langevin*

**Corresponding Author:** arominski@ill.fr

Polarized neutrons experiments are the only technique allowing to analyze individual contributions from nuclear-coherent, incoherent, and magnetic components of neutron scattering cross-section necessary to study, among others, properties of paramagnetic materials. This work presents all stages of data reduction implemented in Mantid for Institut Laue-Langevin's (ILL) D7 instrument, starting from wavelength and position calibration to the final cross-sections in absolute units. The new reduction workflow supports monochromatic and single-crystal diffraction, as well as time-of-flight measurements, using Z-only, 6-point, or 10-point component-separation methods. All results are benchmarked against legacy ILL software.

**Email address of presenting author:**

arominski@ill.fr

**I agree to recordings of my presentation being made at NOBUGS 2022:**

17

## Graphical User Interfaces for ILL users in Mantid

**Authors:** Gagik Vardanyan<sup>1</sup>; Guillaume Communie<sup>1</sup>; Mathieu Tillet<sup>1</sup>; Paolo Mutti<sup>1</sup>

**Co-author:** Dominik Arominski<sup>1</sup>

<sup>1</sup> *Institut Laue-Langevin*

**Corresponding Author:** arominski@ill.fr

Graphical user interfaces (GUI) are extremely useful tools for interacting with data and allowing for simplified workflows where users with minimal coding experience can obtain meaningful results, thus bringing the software to a wider audience. Mantid already contains a number of interfaces allowing for data processing and visualization, with the most important point of entry being the workbench. A number of additional interfaces has been identified as crucial and recently implemented for the smooth processing of data at

the Institut Laue-Langevin, which streamline raw data exploration (Raw-data explorer), simultaneous plotting of many curves (Superplot), and interaction with reduction algorithms in the background (DrILL). The use cases and benefits of new Mantid GUIs for ILL and for wider Mantid public are presented, and future developments are discussed.

**Email address of presenting author:**

arominski@ill.fr

**I agree to recordings of my presentation being made at NOBUGS 2022:**

18

## Into the future: ILL Endurance program

**Authors:** Dominik Arominski<sup>1</sup>; Gagik Vardanyan<sup>1</sup>; Guillaume Communie<sup>1</sup>; Mathieu Tillet<sup>1</sup>; Miguel Gonzalez<sup>1</sup>; Paolo Mutti<sup>1</sup>; Stephane Rols<sup>1</sup>

**Co-authors:** Eric Pellegrini<sup>1</sup>; Remi Perenon<sup>1</sup>

<sup>1</sup> *Institut Laue-Langevin*

**Corresponding Author:** arominski@ill.fr

Institut Laue-Langevin (ILL) is looking forward into the future of reactor-based neutron scattering facility with an ongoing upgrade and maintenance programme Endurance phase 2, that span years 2020-2023. This programme is intended to ensure continued output of quality science on 40 number of instruments and remain a competitive offer for users of neutron scattering. A part of this effort is Better Analysis Software Tools for ILL Experiments (BASTILLE) project, currently in its second phase, aiming to bring full Mantid support for 17 instruments at the ILL, with techniques from Small Angle Scattering, time-of-flight spectroscopy to reflectometry and liquid diffraction, with polarised and un-polarised neutrons. This presentation lays out the goals of the project, its place within the ILL programme, and discusses the current status and the future software developments at the ILL.

**Email address of presenting author:**

arominski@ill.fr

19

## NOMAD OASIS, a Laboratory Data Management Platform

**Authors:** Markus Scheidgen<sup>1</sup>; Sandor Brockhauser<sup>1</sup>

<sup>1</sup> *Humboldt University, Berlin*

**Corresponding Authors:** markus.scheidgen@physik.hu-berlin.de, sandor.brockhauser@physik.hu-berlin.de

NOMAD provided an efficient data sharing platform for materials science for a long time. As one of the flagship projects of the German National Research Data Infrastructure initiative, NFDI, the project FAIRmat has now the goal of extending the NOMAD platform and providing an integrated solution for FAIR data management also covering the needs of synthesis and experimental characterisation laboratories. Next to enabling local deployments in the laboratories and the option of their integration to the NOMAD data sharing network, the newly developed NOMAD OASIS also offers customisable electronic lab notebook, and data exploration, analysis and visualisation services. The later services require an allocated compute infrastructure and run containerised tools on it as cloud

services made available in users' browser.

Additionally, the early results of standardising and transforming metadata is also presented together with a few real life use cases.

**Email address of presenting author:**

sandor.brockhauser@physik.hu-berlin.de

**I agree to recordings of my presentation being made at NOBUGS 2022:**

20

## Experimental Data Infrastructure with BENTEN for Fuel Cell Project at SPring-8

**Author:** Takahiro Matsumoto<sup>1</sup>

**Co-authors:** Albert Mufundirwa<sup>1</sup>; Daiki Fujioka<sup>1</sup>; Hideto Imai<sup>2</sup>; Hirokazu Tsuji<sup>1</sup>; Hiroyuki Iwamoto<sup>1</sup>; Shigeru Yokota<sup>1</sup>; Takuma Kaneko<sup>1</sup>; Tetsuya Miyzawa<sup>1</sup>; Tomoya Uruga<sup>1</sup>; Yoshiharu Sakurai<sup>1</sup>; Yoshiharu Uchimoto<sup>3</sup>; Yuki Mizuno<sup>1</sup>

<sup>1</sup> *Japan Synchrotron Radiation Research Institute*

<sup>2</sup> *NISSAN ARC, LTD.*

<sup>3</sup> *Kyoto University*

**Corresponding Author:** matumot@spring8.or.jp

New project to construct experimental database for evaluation of materials and components of Polymer Electrode Fuel Cells (PEFCs), has been founded by New Energy and Industrial Technology Development Organization (NEDO) since 2020 year 1. In this presentation, we report on the experiment data infrastructure for the Fuel Cell project at synchrotron radiation facility, SPring-8.

We newly developed the data infrastructure using experimental data transfer system BENTEN 2. BENTEN provides easy-to-use interface for data registration and data access using REST API. Registered data can be shared with limited members in the project. Experimental data with synchrotron radiation X-ray analysis methods for XRD, XAFS, SAXS, PDF and HAXPES were accumulated using multiple experimental stations. To realize reliable database, we promoted standardization of the procedure for measurement and analysis to have same results even if different person is involved. Standardization of experimental data format is also one of key issues. We attached a metadata file with unified format with YAML. In this file, sufficient metadata such as data contact person, persistent ID, conditions of sample, measurement and analysis etc. are flexibly defined hierarchically. To reduce the cost to prepare metadata, we setup templates to record metadata for each measurement and also promoted auto-generation.

In summary, we developed experiment data infrastructure with BENTEN for PEFCs technology evaluation, and established the procedure to construct experimental database. We plan to develop flexible data retriever for data utilization, and also start to transfer data catalogues into data platform in National Institute for Materials Science (NIMS) to promote material informatics.

**References:**

- 1) [https://www.nedo.go.jp/activities/ZZJP\\_100182.html](https://www.nedo.go.jp/activities/ZZJP_100182.html)
- 2) T. Matsumoto et al., Proceedings of ICALEPCS 2019, p.702-706
- 3) T. Matsumoto et al., AIP Conference Proceedings 2054, 060076 (2019)

**Acknowledgement:**

This work was supported by the Polymer Membrane Fuel Cell Project of the New Energy and Industrial Technology Development Organization (NEDO) of Japan.

**Email address of presenting author:**

matumot@spring8.or.jp

I agree to recordings of my presentation being made at NOBUGS 2022:

21

## Processing Neutron Time-of-Flight Laue Data in DIALS

**Authors:** David McDonagh<sup>1</sup>; David Waterman<sup>1</sup>

<sup>1</sup> STFC

**Corresponding Author:** david.mcdonagh@stfc.ac.uk

The DIALS project<sup>1</sup> provides an open-source, extensible framework to analyse X-ray diffraction data and is now used widely in the X-ray diffraction community. Much of this framework is in principle agnostic to the method used to obtain diffraction patterns. In recent years this has been expanded for continuous-rotation electron diffraction experiments<sup>2</sup>, for example, highlighting how DIALS can be adapted to cope with challenges from electron sources such as low diffraction angles and lens distortion. Continuing with this push towards generalised diffraction integration software, we present preliminary results for how DIALS can be used to process neutron time-of-flight Laue diffraction patterns obtained from the Single Crystal Diffractometer (SXD) at ISIS<sup>3</sup>.

Here we will show how DIALS has been adapted for polychromatic data, allowing not only the processing of time-of-flight Laue data, but opening up the possibility of processing Laue, and quasi-Laue experiments. Changes to refinement, integration, and visualisation will be discussed, including a browser-based GUI for streamlined model reduction workflows.

1 Winter G., Waterman D. G., Parkhurst J. M., Brewster A. S., Gildea R. J., Gerstel M., Fuentes-Montero L., Vollmar M, Michels-Clark T., Young I. D., Sauter N. K., Evans G. (2017). *Acta Crystallogr. D* 74, 85-97

2 Clabbers T. B., Gruene T., Parkhurst J. M., Abrahams J. P., & Waterman D. G. (2018). *Acta Crystallogr. D* 74, 506-518

3 Keen D. A., Gutmann M. J., & Wilson C. C. (2006). *J. Appl. Cryst.* 39, 714-722

**Email address of presenting author:**

david.mcdonagh@stfc.ac.uk

I agree to recordings of my presentation being made at NOBUGS 2022:

22

## The Mamba Data Acquisition Software Project for HEPS

**Author:** Yi Zhang<sup>1</sup>

<sup>1</sup> *Institutes of High Energy Physics, Chinese Academy of Sciences*

**Corresponding Author:** zhangyi88@ihep.ac.cn

Advancements in synchrotron methodology have long been limited by both hardware instrumentation and software. While hardware limitations are gradually being lifted with the emergence of next-generation light sources and beamline instrumentation, high (data) throughput, multimodal, time-resolved, in-situ and dynamic experiments performed at future beamlines are expected to impose tremendous challenges on the software end <sup>1</sup>. Even today, data acquisition software performance has become a key factor hindering advancements in synchrotron methodology.

To address pressing software challenges at High Energy Photon Source (HEPS), the data acquisition team launched the Mamba project in 2020, aiming to develop a systematic Python-based software framework on top of Bluesky (NSLS II) and complete ecosystem for the cutting-edge and data-intensive experiments carried out at the Phase I beamlines of HEPS. After nearly two years of R&D, Mamba's framework design was published online in the Journal of Synchrotron Radiation in May 2022, laying a firm foundation for development of experimental applications at HEPS. During BSRF's first dedicated user operation period in 2022, Mamba-based data collection software applications were made available to experiments at multiple beamlines of a first generation synchrotron source (BSRF). This talk will present the progress of Mamba projects after two years development and our developing plans towards 2025.

References:

- 1 Dong, Y., Li, C., Zhang, Y. et al. Exascale image processing for next-generation beamlines in advanced light sources. *Nat Rev Phys* 4, 427–428 (2022)
- 2 Liu, Y., Geng, Y., Bi, X., Li, X., Tao, Y., Cao, J., Dong, Y. & Zhang, Y. Mamba: a systematic software solution for beamline experiments at HEPS. *J. Synchrotron Rad.* 29, 664-669 (2022)

**Email address of presenting author:**

zhangyi88@ihep.ac.cn

23

## Minimalist beamline control and experiment software

**Author:** Yu Liu<sup>1</sup>

<sup>1</sup> *Institute of High Energy Physics, Chinese Academy of Sciences*

**Corresponding Author:** liuyu91@ihep.ac.cn

This presentation reports the efforts at HEPS to systematically approach the complexity lower bounds both in beamline control and in experiment software:

- The former (Liu 2022a) is done by minimising repetitive work on multiple scales: single devices (with reusable, modular and minimal IOCs), devices on a single beamline (with minimalist package management for EPICS modules), and all beamlines at HEPS (with the comprehensive beamline services).
- The latter (Liu 2022b) is done in the Mamba framework by its command injection / RPC mechanism, the Mamba Data Worker to implement modular data-processing graphs, and experiment parameter generators (EPGs) to abstract irrelevant or repetitive details.

In addition to the contents already in the papers cited above, also presented are some examples which the presenter believes to have satisfactorily approximated the complexity lower-bounds in certain aspects:

- ADXspress3 (with a paper in progress to discuss the techniques used in its refactoring process, which has been successfully applied to numerous projects), in minimising efforts required to adapt the IOC to a different number of Xspress3 boxes or channels.
- ihep-pkg (with recent updates, and also supporting Rocky Linux 8 now), in minimising the efforts required to maintain reproducible RPM packages for EPICS modules (covering the full synApps collection) which also provide reusable modular IOCs.
- (With a paper in progress,) based on an ophyd module implementing full control of PandABox's TCP server, a backend EPG for fly scans providing automated configuration of "Panda Blocks" for constant-speed mapping of various dimensions, as well as generation of scans deliberately fragmented to overcome hardware limits.

**Email address of presenting author:**

liuyu91@ihep.ac.cn

**I agree to recordings of my presentation being made at NOBUGS 2022:**

24

## Ophyd v2

**Authors:** Callum Forrester<sup>1</sup>; Daniel Allan<sup>2</sup>; Thomas Caswell<sup>2</sup>; Thomas Cobb<sup>1</sup>

<sup>1</sup> *Diamond Light Source*

<sup>2</sup> *NSLS-II*

**Corresponding Author:** callum.forrester@diamond.ac.uk

NSLS-II's Bluesky framework has enabled scanning and experiment orchestration at multiple facilities for a number of years. In this talk we present the preliminary work on Ophyd v2, a drop-in replacement for Ophyd that complies with the Bluesky protocols. Ophyd v2 is being developed collaboratively between NSLS-II and Diamond Light Source, it incorporates lessons learned from the original implementation (which is one of the most mature and widely used of the Bluesky libraries) as well as Diamond's expertise and experience with hardware triggered scanning (see pymalcolm). We summarise the proposed changes to the Ophyd API and user experience with associated reasoning, as well as the plans for backwards-compatibility, simultaneous use of both versions and incremental migration strategy, allowed for by the modular nature of Bluesky.

**Email address of presenting author:**

callum.forrester@diamond.ac.uk

25

## Recent developments in the MSlice software package

**Author:** Silke Schomann<sup>1</sup>

<sup>1</sup> *ISIS*

**Corresponding Author:** silke.schomann@stfc.ac.uk

MSlice1 is a Python-based tool for performing and visualizing slices and cuts of inelastic neutron scattering data. It can be used both as a standalone application and as an interface of the data reduction software Mantid2.

We provide an update on recent developments in the MSlice software package as well as on the new packaging system used for deploying the standalone version of MSlice.

This involves improvements for the GUI such as a new plot manager tab, enhanced interaction with Mantid by making MSlice workspaces available in Mantid and upgrades to the script generation.

New functionalities also include various improvements for calculating and displaying cuts. It is now possible to convert intensity information for cut plots. In addition, it is now also possible to overplot powder reflections from materials such as Aluminium, Copper, Niobium and Tantalum.

MSlice differentiates between data from instruments with and without position-sensitive detectors. For both types of data there are now two different cut operations adapted to the respective data type available. In addition to the original method, the rebin cut algorithm, an integration cut algorithm was introduced. Unlike the rebin cut algorithm, this integration cut algorithm will not assume constant signals that can be extrapolated over regions with less data and is therefore more suitable for

integration over energy.

Another area of enhancement is the migration to Conda packaging for the cross-platform deployment of MSlice. Both developers and users profit from the ability to install several versions of standalone MSlice applications independently in separate Conda environments.

1. <https://mantidproject.github.io/mslice/>
2. <https://mantidproject.org/>

**Email address of presenting author:**

[silke.schomann@stfc.ac.uk](mailto:silke.schomann@stfc.ac.uk)

26

## XRD2: MX beamline at Elettra - A mix recipe

**Author:** Annie Héroux<sup>1</sup>

**Co-authors:** Emiliano Coghetto<sup>2</sup>; Nicola Demitri<sup>2</sup>; Raghurama P Hegde<sup>2</sup>; Roberto Borghes<sup>2</sup>

<sup>1</sup> *Elettra Sincrotrone Trieste*

<sup>2</sup> *Elettra Sincrotrone Trieste S.C.p.A.*

**Corresponding Author:** [annie.heroux@elettra.eu](mailto:annie.heroux@elettra.eu)

For the past 5 years, Elettra has offered an MX beamline on its menu. Even though Italian and Indian users are its main patrons, clients from a wide variety of locations are enjoying all it has to offer. We will describe how we blend flavors of software from different facilities with our in-house ingredients to satisfy the utmost efficiency for mainly take-out orders.

More than a decade of experience at other beamlines around the world helped shape the hardware development to address the needs of our scientific community. XRD2 is no exception with a highly automated environment to allow high throughput and streamlined data analysis pipelines. From its inception, XRD2 needed to provide reliable remote access which steered most of our decisions.

Elettra participates in the MXCuBE[1,2] collaboration initiated at the ESRF. Despite being almost 2 decades old, the vibrant community is still going strong and reaches out to facilities all around the globe. The web version is suited for our Indian partners and has proven to be more performant than remote desktop sharing. MXCuBE is compatible with Elettra's Tango Control System for which a large community provides constant new hardware support and new methods are being folded in.

To pair with data collection, we opted for SynchWeb3 from Diamond Light Source with its flavor of ISPyB4 to answer most of the needs involved in sessions at the beamline. Its straightforward interface with our Users' portal, VUO (Virtual Unified Office) enables us to associate scheduling with the relevant people for beamline and data access.

Other aspects such as visualization of the diffraction pattern, data download, live communication with users, and shipping of the dewars rely on in-house solutions. Our facility also provides highly flexible access time, with monthly proposal submission, visits split over several sessions or shifted to match with dewar delivery. This overall blend of technology and flexibility from the facility provides a familiar environment to the users, and manageable support from the staff which has been our recipe for success.

1 Oscarsson, M. et al. 2019. "MXCuBE2: The Dawn of MXCuBE Collaboration." *Journal of Synchrotron Radiation* 26 (Pt 2): 393–405.

2 Gabadinho, J. et al. (2010). MxCuBE: a synchrotron beamline control environment customized for macromolecular crystallography experiments. *J. Synchrotron Rad.* 17, 700-707

3 SynchWeb: a modern interface for ISPyB S. Fisher et al., *J. Appl. Cryst.* (2015). 48, 927-932

4 SPyB: an information management system for synchrotron macromolecular crystallography S. Delageniere et al., *Bioinformatics* (2011) 27 (22): 3186-3192



**Email address of presenting author:**

annie.heroux@elettra.eu

**I agree to recordings of my presentation being made at NOBUGS 2022:**

27

## 3D Grid Scans: A Bluesky prototype at Diamond Light Source

**Author:** Dominic Oram<sup>1</sup>

<sup>1</sup> *Diamond Light Source*

**Corresponding Author:** dominic.oram@diamond.ac.uk

To complement the facility's next major upgrade, Diamond Light Source (DLS) is undergoing a programme to modernise its software stack. One area of particular interest is in scanning and experiment orchestration, for which DLS is migrating to the Bluesky framework. As a prototype, a Bluesky implementation of Xray Centring using a 3D grid scan has been implemented on a beamline at DLS. The software has successfully been integrated with existing infrastructure at DLS, including the current DAQ software, the LIMS, NeXus file writing and the analysis pipelines. This talk will present the results from this prototype and the lessons learnt from the project, which will inform the design of the production system and its rollout across the wider facility.

**Email address of presenting author:**

dominic.oram@diamond.ac.uk

**I agree to recordings of my presentation being made at NOBUGS 2022:**

28

## Building a PaN Data Commons on the outcomes of the PaNOSC and ExPaNDS EOSC projects

**Authors:** Andy Gotz<sup>1</sup>; Patrick Fuhrmann<sup>2</sup>

<sup>1</sup> *ESRF*

<sup>2</sup> *Deutsches Elektronen Synchrotron, DESY*

**Corresponding Author:** andy.gotz@esrf.fr

The European Open Science Cloud (EOSC) is a project from the European Commission to enable and facilitate access to a wide range of open data and services in Europe. The EOSC financed two projects dedicated to developing data services for the Photon and Neutron (PaN) community – PaNOSC (<https://panosc.eu>) and ExPaNDS (<https://expands.eu>). This talk will present the outcomes of the two projects and how they can be used to build a PaN Data Commons as part of the EOSC. The Data Commons acts as a single point of entry to search for and find data from the PaN facilities in Europe. The PaN Data Commons will provide a community service available to the EOSC to boost data sharing and re-use and ensure data from the PaN facilities are FAIR.

**Email address of presenting author:**

andy.gotz@esrf.fr

29

## Daiquiri: a web based user interface framework for beamline control and data acquisition

**Author:** Stuart Fisher<sup>1</sup>

**Co-authors:** Jens Meyer<sup>1</sup>; Marcus Oscarsson<sup>1</sup>; Marine Cotte<sup>1</sup>; Valentin Valls<sup>1</sup>; Wout de Nolf<sup>1</sup>

<sup>1</sup> ESRF

**Corresponding Author:** stuart.fisher@esrf.fr

Daiquiri 1 is a web based User Interface (UI) framework for control system monitoring and data acquisition. It provides simple, intuitive, and responsive interfaces to control and monitor hardware, launch acquisition sequences, and manage associated metadata. Daiquiri concerns itself only with the UI layer, it does not provide a scan engine or controls system but can be easily integrated with existing systems. Daiquiri is implemented with a traditional client / server methodology with the intention of producing a generic extensible framework for acquisition. The server is implemented in Python 3 and provides a REST API and SocketIO service for real-time feedback. The client is implemented in javascript es6 making use of the popular front end framework React along with Redux. Daiquiri is currently deployed on a number of beamlines at ESRF including the scanning X-ray microscope beamline ID21, the microfocus X-ray mapping beamline ID13, and the BioSAXS beamline BM29. Deployment is in progress to another four beamlines in 2022. In the future daiquiri will be the standard interface by which users and scientists interact with the controls system on beamlines at ESRF.

Further information can be found at <https://ui.gitlab-pages.esrf.fr/daiquiri-landing>

References:

1. Daiquiri: a web-based user interface framework for beamline control and data acquisition, Fisher et al., J. Synchrotron Rad. (2021). 28, 1996-2002

**Email address of presenting author:**

stuart.fisher@esrf.fr

30

## Multi-purpose Regulation Loops in BLISS

**Author:** Perceval Guillou<sup>1</sup>

**Co-authors:** Anthony Mauro<sup>1</sup>; Jens Meyer<sup>1</sup>; Matias Guijarro<sup>1</sup>; Sebastien Petitdemange<sup>1</sup>; Valentin Valls<sup>1</sup>

<sup>1</sup> ESRF

**Corresponding Author:** perceval.guillou@esrf.fr

BLISS, the ESRF beamline control system, in production since the ESRF-EBS upgrade, provides a framework to ease and standardize the integration of a variety of regulation devices (like temperature or pressure controllers). Thanks to a generic API, users can define custom regulation systems by combining BLISS objects e.g. a probe and a motor can be declared as Input and Output of a software loop with its own PID control. Real-time plotting of PID loop parameters allows easy tuning and monitoring of values directly in Flint, the BLISS data online visualization tool. The framework also brings extra features to control regulation loops directly from the shell or as background tasks. Any BLISS regulation loop can be exported as a Tango device in order to share hardware over multiple clients or to bridge with other systems.

**Email address of presenting author:**

perceval.guillou@esrf.fr

**I agree to recordings of my presentation being made at NOBUGS 2022:**

31

## **py-ISPyB - A new implementation of a LIMS for experiments in structural biology**

**Authors:** Alex de Maria<sup>1</sup>; Maël Gaonach<sup>1</sup>

**Co-author:** Stuart Fisher <sup>1</sup>

<sup>1</sup> ESRF

**Corresponding Author:** mael.gaonach@esrf.fr

ISPyB is a mature Laboratory Information System (LIMS) for synchrotron-based Macromolecular Crystallography (MX), Small Angle Scattering (BioSAXS), and Electron Microscopy (EM) experiments, developed and used by a number of light sources world-wide. The project has been developed over the last 20 years and is based on an ageing JAVA software stack. The ISPyB collaboration has continued to grow and the software is now a critical part of experiments at many synchrotrons. In 2019 the ISPyB collaboration decided to evaluate newer software technologies and redesign the architecture to enable easier maintenance and facilitate extension. A python prototype was developed in 2019 by EMBL Hamburg which has now matured into an extensible framework. The first use case of this new framework, py-ISPyB, is to implement Serial Crystallography (SSX) on the ID29 beamline at the ESRF-EBS. This talk will present the current state and progress of the project including the software design, code, and collaboration.

**Email address of presenting author:**

mael.gaonach@esrf.fr

32

## **Experience with developing FPGA accelerated data reduction in DevCloud**

**Author:** Zdenek Matej<sup>1</sup>

**Co-authors:** Andrii Salnikov <sup>1</sup>; Artur Barczyk <sup>1</sup>

<sup>1</sup> MAX IV Laboratory, Lund University

**Corresponding Author:** zdenek.matej@maxiv.lu.se

Field Programmable Gate Arrays (FPGAs) present interesting tool for accelerating scientific computing and artificial intelligence applications in high performance computing centers and at large scale photon and neutron facilities. Reliable, high throughput and low latency processing of data from X-ray detectors is likely the most exciting application in the latter case. Accessibility of proper hardware infrastructure and testing environment including software is a crucial component of the application development process, including implementation of continuous integration and deployment. Cloud services are well established in providing ecosystems for developing “conventional” software nowadays. Contrary, cloud service instances supporting “hardware” accelerated software are less known and domain specific applications are rather being implemented on “edge” clouds. A

brief overview of public cloud services available for FPGA applications developers is given in this contribution, followed by a report on tools, experience in using Intel DevCloud for a project focused on FPGA accelerated data reduction for synchrotron data 1.

1 *bincount implementation of Azimuthal Integration (AZINT) with FPGAs*, <https://gitlab.com/MAXIV-SCISW/compute-fpgas/bincount> (last visited on August 18th, 2022)

**Email address of presenting author:**

zdenek.matej@maxiv.lu.se

33

## The Blissdata Project: new perspectives on BLISS Data Management

**Authors:** Lucas Felix<sup>1</sup>; Matias Guijarro<sup>1</sup>; Wout de Nolf<sup>1</sup>

<sup>1</sup> ESRF

**Corresponding Author:** lucas.felix@esrf.fr

One of the main features of BLISS, the ESRF beamline control system in production since the ESRF-EBS upgrade, is to decouple data acquisition from online data storage. Redis, an online database, is used as a temporary buffer to store all the data produced by an acquisition or a reference to the data in the case of large data. This allows clients to consume data without perturbing the acquisition, and alleviates real-time constraints when clients want to access data for display (for example in the case of Flint, the BLISS data online visualization tool) or for saving in a HDF5 file (for example in the case of the BLISS Nexus Writer). In addition to display and saving, our goal is to enable online data analysis, with clients accessing the data buffer to produce intermediate results or to process data. The way data is structured within Redis is specific to BLISS and clients need to use the BLISS API. This means data analysis programs must import the whole BLISS library including beamline control parts and dependencies e.g. gevent. In order to provide a lightweight easy to use API library for online data analysis, the BLISS data management code is being refactored into a separate library called “blissdata”. Clients with access to Redis can read data on the fly using an h5py-like API. The h5py compatible API will allow data processing code to access data in the online buffer as a file to process online and offline data in the same way.

**Email address of presenting author:**

lucas.felix@esrf.fr

**I agree to recordings of my presentation being made at NOBUGS 2022:**

34

## NXtomomill: more than just a data conversion tool

**Authors:** Christian Nemoz<sup>1</sup>; Henri Payno<sup>1</sup>; Pierre Paleo<sup>1</sup>

**Co-authors:** Alexander Rack<sup>1</sup>; Marco Di Michiel<sup>1</sup>; Nicola Viganò<sup>1</sup>; Paul Tafforeau<sup>1</sup>; Peter Cloetens<sup>1</sup>; Vicente Armando Solé<sup>1</sup>

<sup>1</sup> ESRF - The European Synchrotron

**Corresponding Author:** nicola.vigano@esrf.fr

The ESRF is rebuilding its acquisition and processing workflows from scratch, with unified solutions whenever possible, to deliver both a homogeneous experience across all its beamlines, and robust high-performance processing software. However, due to the specificity of the different beamlines and techniques: (a) the acquisition data format might still differ; (b) the data itself might require specific pre-processing, before reconstruction.

We are solving points (a) and (b) with a common data format and versatile conversion software for all X-ray tomography techniques and beamlines. NXtomo (from the NeXus international standard) is our choice for a common tomographic data format. NXtomomill is an open source software package, developed at the ESRF, for the transformation of all the required raw tomographic data into NXtomo compliant form.

NXtomomill supports rearranging several input data formats for full-field tomography: This includes traditional ESRF's EDF full-field datasets, and APS'DataExchange. In the future, it will support advanced phase retrieval methods (e.g. for holotomography), through a plug-in that will use specialized software. Similarly, NXtomomill is now also receiving the support for azimuthal integration of XRD-CT scans, and elemental fitting of XRF-CT scans.

NXtomomill guarantees an identical output data format for each ingested raw data format and data type. It decouples data handling from data reconstruction, resulting in uniform user experience, easier development, reduced maintenance costs, and greater robustness of the tomography processing pipeline. This also supports easier data and soft-ware exchanges with other synchrotron radiation facilities.

**Email address of presenting author:**

nicola.vigano@esrf.fr

**I agree to recordings of my presentation being made at NOBUGS 2022:**

35

## Unified access to photon and neutron data throughout Europe - the SciCat implementation of the PaNOSC Search

**Authors:** Carlo Minotti<sup>1</sup>; Fredrik Bolmsten<sup>2</sup>; Massimiliano Novelli<sup>2</sup>; Stephan Egli<sup>1</sup>; Tobias Richter<sup>2</sup>

<sup>1</sup> PSI - Paul Scherrer Institut

<sup>2</sup> European Spallation Source

**Corresponding Author:** max.novelli@ess.eu

As part of the European Open Science project, PaNOSC and the sister project ExPaNDS have developed a federated access API to public data held at any of the participating photon and neutron sources. The API and accompanying web front end support domain specific queries. This presentation will briefly cover the PaNOSC search API but mainly focus on the SciCat implementation, which is the data catalogue in operations at the partners ESS, PSI and MAX IV.

We will explain how the PaNOSC data model is mapped to SciCat data, how filtering by experimental technique and parameters is implemented and importantly how the relevancy of search results are estimated. This "scoring service" assigns a significance score to each dataset result returned to PaNOSC and allows inter-facility sorting of the results. For this the data in the catalogue had to undergo a data curation process, which we will illustrate and explain based on selected queries and their results.

**Email address of presenting author:**

max.novelli@ess.eu

**I agree to recordings of my presentation being made at NOBUGS 2022:**

36

## Automated Scientific Metadata Recording and Viewing During Experiments at MAX IV

**Authors:** Daphne van Dijken<sup>1</sup>; Magnus Klingberg<sup>1</sup>; Mikel Eguiraun<sup>1</sup>; Vincent Hardion<sup>1</sup>

<sup>1</sup> MAX IV

**Corresponding Author:** daphne.van\_dijken@maxiv.lu.se

Beamline instrumentation has dramatically improved over the years in synchrotron research facilities. Nowadays, detectors can produce thousands of frames in a matter of seconds. Therefore, a well-structured and configurable framework is required to easily access and assess the quality of these enormous amounts of data.

In this communication we present a metadata management solution recently developed and implemented at MAX IV to automatically retrieve and record metadata from Tango devices relevant to the current experiment. User-selected scientific metadata and predefined defaults related to the beamline setup are propagated into the Sardana control system and automatically recorded at each scan using a library, SciFish1. The recorded metadata, stored in the SciCat2 database, can be accessed through a ReactJS-based web-interface, Scanlog 3, to easily sort, filter and extract important information. This tool allows to access the metadata in real-time and is used for monitoring as well as exporting for post-processing.

These new software tools ensure that recorded data is findable, accessible, interoperable and reusable (FAIR4) for many years to come. Collaborations are on-going to develop these tools at other particle accelerator research facilities.

Footnotes:

1 SciFish <https://gitlab.com/MaxIV/lib-maxiv-scifish>

2 SciCat <https://scicatproject.github.io/>

3 Scanlog <https://gitlab.com/MaxIV/svc-maxiv-scanlog>

4 Wilkinson, Mark D., et al. "The FAIR Guiding Principles for scientific data management and stewardship. Sci Data 3: 160018." (2016). <https://www.nature.com/articles/sdata201618>

**Email address of presenting author:**

daphne.van\_dijken@maxiv.lu.se

37

## Current and future developments of European XFEL scan tool Karabacon

**Author:** Ivars Karpics<sup>1</sup>

**Co-authors:** Dennis Göries<sup>1</sup>; Cammille Carinan<sup>1</sup>; Wajid Ehsan<sup>1</sup>; Istvan Mohacsi<sup>2</sup>; Steffen Hauf<sup>3</sup>

<sup>1</sup> European XFEL

<sup>2</sup> PSI - Paul Scherrer Institut

<sup>3</sup> European XFEL GmbH

**Corresponding Author:** ivars.karpics@xfel.eu

User experiments at synchrotron and free electron laser sources typically require longer duration data acquisition while synchronously moving several actuators and motors. Most of the accelerator control systems contain scan engines and tools (for example 1. - 4.) facilitating such experimental

data collections. At European XFEL, the so called scan tool “Karabacon” has been developed, and successfully used [5]. It is an extension of the Karabo [6] control system, and includes a command line and a graphical user interface with real-time and historical plots, basic data analysis tools as well as scan customization extensions. In this contribution, current and future developments of the Karabacon scan tool are presented.

1 SPEC: <https://certif.com/content/spec/>

2 Sardana Spock: <https://sardana-controls.org/users/spock.html>

3 Bliss: [https://bliss.gitlab-pages.esrf.fr/bliss/master/bliss\\_standard\\_scans.html](https://bliss.gitlab-pages.esrf.fr/bliss/master/bliss_standard_scans.html)

4 Bluesky: <https://nsls-ii.github.io/bluesky/plans.html>

[5] Karabacon: <https://rtd.xfel.eu/docs/scantool/en/latest/>

[6] Hauf, Steffen, et al. “The Karabo distributed control system.” *Journal of synchrotron radiation* 26.5 (2019): 1448-1461.

**Email address of presenting author:**

ivars.karpics@xfel.eu

**I agree to recordings of my presentation being made at NOBUGS 2022:**

38

## Overview of BLISS, the ESRF-EBS beamline control system

**Authors:** Cyril Guilloud<sup>1</sup>; Emmanuel Papillon<sup>1</sup>; Gilles Berruyer<sup>1</sup>; Jens Meyer<sup>1</sup>; Laurent Claustre<sup>1</sup>; Lucas Felix<sup>1</sup>; Matias Guijarro<sup>1</sup>; Perceval Guillou<sup>1</sup>; Sebastien Petitdemange<sup>1</sup>; Valentin Valls<sup>1</sup>; Wout De Nolf<sup>2</sup>

**Co-authors:** Alejandro Homs<sup>1</sup>; Anthony Mauro<sup>1</sup>; Antonia Beteva<sup>1</sup>; Clemence Muzelle<sup>1</sup>; Holger Witsch<sup>1</sup>; Marcus Oscarsson<sup>1</sup>; Marie-Claire Lagier<sup>1</sup>; Samuel Debionne<sup>1</sup>; Stuart Fisher<sup>1</sup>

<sup>1</sup> ESRF

**Corresponding Author:** [guijarro@esrf.fr](mailto:guijarro@esrf.fr)

Since the start of ESRF-EBS, the Extremely Brilliant Source upgrade, the new beamline control system, BLISS, is in production on half of the ESRF beamlines since August 2021, with deployment on all beamlines planned for 2024. BLISS is a Python library and a set of tools to empower scientists with the ability to write and to execute complex data acquisition sequences. BLISS is made of mini frameworks, to standardize and ease the integration of various devices like motor controllers, multi-channel analyzers, regulation devices, or 2D detectors. BLISS also includes a powerful command line interface, a visualization tool (Flint) and advanced data management features enabling online data analysis. This talk will present an overview of the BLISS project, focusing on feature highlights and technical information as well as feedback from real experiments conducted with BLISS.

**Email address of presenting author:**

[guijarro@esrf.fr](mailto:guijarro@esrf.fr)

39

## diffcalc-core: Diffraction condition calculation package for a six-circle diffractometer

**Author:** Irakli Sikharulidze<sup>1</sup>

<sup>1</sup> Diamond Light Source

**Corresponding Author:** irakli.sikharulidze@diamond.ac.uk

The Python 3 package `diffract-core` implements diffraction condition calculation for a six-circle x-ray diffractometer according to the methodology described in H. You. Angle calculations for a '4S+2D' six-circle diffractometer. *J. Appl. Cryst.* (1999). 32, 614-623. Diffractometer operation modes are selected based on three constraints that can include combination of detector, sample, reference vector, scattering plane, incident, exit beam or  $\chi$ -plane orientations. All solutions fulfilling the diffractometer equation are returned for a requested reciprocal orientation. Configurations that correspond to a continuous solution space are flagged as insufficiently constrained. UB matrix can be set directly or calculated using reference reflections and/or known crystal orientations. Methods are provided to refine UB matrix based on a single reflection data or using a least-square fitting procedure with data from multiple reflections. The package source is available on GitHub (<https://github.com/DiamondLightSource/diffract-core>) under terms of Apache Software License ver. 2.0.

**Email address of presenting author:**

irakli.sikharulidze@diamond.ac.uk

**I agree to recordings of my presentation being made at NOBUGS 2022:**

40

## Image Annotation at European XFEL

**Author:** Ana García-Tabarés<sup>1</sup>

**Co-authors:** Dennis Göries<sup>1</sup>; Ivars Karpics<sup>1</sup>; Luca Gelisio<sup>1</sup>; Steffen Hauf<sup>2</sup>

<sup>1</sup> *European XFEL*

<sup>2</sup> *European XFEL GmbH*

**Corresponding Author:** ana.garcia-tabares@xfel.eu

Cameras and imaging tools in diagnostic systems are valuable sources of information at photon sources, and instrument scientists rely on their information to perform their experiments. Defining the reference position of the beam during alignment of the instrument setup, monitoring and tuning the beam stability or aligning the position of the target with respect to the beam are examples of tasks performed with imaging cameras. However, at the European XFEL, existing tools do not allow to extract the information in a computer-readable form and thus make tracking the events observed during the different phases of the experiment or, even across experiments, difficult. As part of the AMORE (Automated Metadata annotation Reconstruction Environment) project, the European XFEL Control group has developed a set of tools that allows instrument scientists to extract and integrate metadata from/to existing imaging tools integrated into the control system, as well as to process and store them. This contribution summarizes the tools under development and their applications.

**Email address of presenting author:**

ana.garcia-tabares@xfel.eu

**I agree to recordings of my presentation being made at NOBUGS 2022:**

41

## Data-Modelling Patterns for Experimental Characterisation

**Author:** Sandor Brockhauser<sup>1</sup>



**Co-authors:** Andrea Albino <sup>1</sup>; Carola Emminger <sup>2</sup>; Tommaso Pincelli <sup>3</sup>; Tamás Haraszti <sup>4</sup>; Markus Kühbach <sup>1</sup>; Laurenz Rettig <sup>3</sup>; Christoph Koch <sup>1</sup>; Martin Aeschlimann <sup>5</sup>; Yue Sun <sup>6</sup>; Norbert Esser <sup>7</sup>; Chris Sturm <sup>8</sup>; Ralph Ernstorfer <sup>3</sup>; Sherjeel Shabih <sup>1</sup>; Walid Hetaba <sup>9</sup>; Marius Grundmann <sup>8</sup>; Florian Dobener <sup>1</sup>; Sebastian Brückner <sup>1</sup>

<sup>1</sup> *Department of Physics, Humboldt-Universität zu Berlin, Zum Großen Windkanal 2, 12489 Berlin, Germany*

<sup>2</sup> *Department of Physics, Humboldt-Universität zu Berlin, Zum Großen Windkanal 2, 12489 Berlin, Germany and Felix-Bloch Institut für Festkörperphysik, Universität Leipzig, Linnéstr. 5, 04103 Leipzig, Germany*

<sup>3</sup> *Fritz Haber Institute of the Max Planck Society, Faradayweg 4-6, 14195 Berlin, Germany*

<sup>4</sup> *DWI Leibniz Institute for Interactive Materials, Aachen, 52074 Germany*

<sup>5</sup> *Technische Universität Kaiserslautern, Erwin-Schrödinger Str. 46, 67663 Kaiserslautern*

<sup>6</sup> *Software Engineering Department, Institute of Informatics, University of Szeged, Dugonics tér 13, 6720 Szeged, Hungary and European XFEL GmbH, Holzkoppel 4, 22869 Schenefeld, Germany*

<sup>7</sup> *Leibniz Institut für Analytische Wissenschaften, Schwarzschildstr. 8, 12489 Berlin*

<sup>8</sup> *Felix-Bloch Institut für Festkörperphysik, Universität Leipzig, Linnéstr. 5, 04103 Leipzig, Germany*

<sup>9</sup> *Max Planck institute for chemical energy conversion, Stiftstraße 34-36 45470, Mülheim an der Ruhr*

**Corresponding Authors:** florian.dobener@physik.hu-berlin.de, sandor.brockhauser@physik.hu-berlin.de

In computer science, programming patterns are an invaluable source of creating state-of-the-art software. They propel the invention cycle of software projects, providing the vast availability of software solutions. Further, there exists a lot of tools and standards to manage and store data in the big data and AI economies. However, in scientific research structured data modelling patterns are not widely adopted across fields. This often leads to lab-contained data and programming solutions making the reproducibility and interoperability of such data a barely manageable task. We aim to provide standardisation for lab-sized data management according to the FAIR principles (findable, accessible, interoperable, reproducible) and show how to take advantage of the NeXus standard. NeXus is a common exchange format for neutron, X-ray, and muon experiments, which we used mainly at large beamline facilities. We adapted generic data modelling patterns to a set of example measurements (for various optical and electrical characterisation methods) and also applied it to NeXus, where we encountered several limitations in representing data collections in a standardised and machine actionable way. Here, we present our solutions to these limitations and offer data patterns to tackle common design patterns in organising data collections and their metadata. Although some of the examples are using the NeXus standard as a specific representation, our approach is not tied to a specific format, and we aim to present generic data modelling strategies. Accordingly, these patterns are for general use, and support making FAIR data management a systematically solvable task.

**Email address of presenting author:**

Sandor.brockhauser@physik.hu-berlin.de

**I agree to recordings of my presentation being made at NOBUGS 2022:**

Yes

42

## The Karabo Control System

**Author:** Steffen Hauf<sup>1</sup>

**Co-authors:** Alessandro Silenzi <sup>2</sup>; Ana García-Tabarés <sup>2</sup>; Andrea Parenti <sup>1</sup>; Anna Klimovskaia <sup>2</sup>; Ayaz Samadli <sup>2</sup>; Christopher Youngman <sup>2</sup>; David Hickin <sup>2</sup>; Dennis Göries <sup>2</sup>; Florian Sohn <sup>2</sup>; Gabriele Giovanetti <sup>2</sup>; Gero Flucke <sup>1</sup>; Hugo Santos <sup>2</sup>; Ivars Karpics <sup>2</sup>; Jawad Tahmeed Bin Taufik <sup>2</sup>; Martin Teichmann <sup>2</sup>; Michael Smith <sup>2</sup>; Milan Staffehl <sup>2</sup>; Noushadali Anakkappalla <sup>2</sup>; Raul Costa <sup>2</sup>; Sergey Esenov <sup>2</sup>; Valerii Bondar <sup>2</sup>; Wajid Ehsan <sup>2</sup>

<sup>1</sup> *European XFEL GmbH*

<sup>2</sup> *European XFEL*

**Corresponding Author:** steffen.hauf@xfel.eu

The Karabo distributed control system has been developed to address the challenging requirements of the European X-ray Free Electron Laser facility 1, including complex and custom-made hardware, high data rates and volumes, and hand-off of data to online data analysis applications for distributed processing and rapid feedback. Karabo is an extensible, distributed application, which implements a broker-based, supervisory control and data acquisition (SCADA) environment as part of a distributed control system 2. Extensions to the core framework, so-called Devices implement control of hardware, monitoring, data acquisition and online processing on distributed hardware. Services for data logging, configuration management and situational awareness through alarm indicators exist. The flexible framework exposes Python and C++ programming APIs, which enable developers to quickly respond to new requirements in instrument control, within an efficient development. Its graphical user interface (GUI), which features an intuitive, coding-free control panel builder, allows non-software engineers to create synoptic control views.

In this contribution the Karabo Control System is introduced, both out of the perspective of application users and software developers. Emphasis is given to Karabo's asynchronous Python 3 programming environment. We share experience of running a large facility like the European XFEL using a clean-sheet developed control system, and discuss the foreseeable release of the system as free and open source software.

1 Tschentscher, Thomas, et al. "Photon beam transport and scientific instruments at the European XFEL." *Applied Sciences* 7.6 (2017): 592.

2 Hauf, Steffen, et al. "The Karabo distributed control system." *Journal of synchrotron radiation* 26.5 (2019): 1448-1461.

**Email address of presenting author:**

steffen.hauf@xfel.eu

43

## Zocalo: a high-throughput data processing framework

**Author:** Richard Gildea<sup>1</sup>

**Co-authors:** Markus Gerstel<sup>1</sup>; Nicholas Devenish<sup>1</sup>; Irakli Sikharulidze<sup>1</sup>; Dan Hatton<sup>1</sup>; Anna Horstmann<sup>1</sup>; Graeme Winter<sup>1</sup>

<sup>1</sup> *Diamond Light Source Ltd*

**Corresponding Author:** richard.gildea@diamond.ac.uk

Zocalo is a data processing system developed to support live analysis at Diamond Light Source across multiple science areas, including crystallography and cryo-EM. Zocalo has been designed to be low-latency and fault-tolerant, and is used to orchestrate a wide variety of data processing tasks, including fast X-ray centring data analysis, automated data reduction with xia2, combining multiple data sets with xia2.multiplex, user-instigated reprocessing via SynchWeb and providing live feedback for serial crystallography experiments (SSX).

Zocalo is made up of several main components. *Services* start in the background and wait for work comprising discrete short-lived tasks, for example finding diffraction spots on a single image, or inserting results into a database. In contrast, *wrappers* are used for longer-running tasks, such as data reduction of a diffraction experiment. *Recipes* describe the connections between services and wrappers, enabling highly flexible and dynamic data processing workflows. Messages are passed between components via an underlying message broker.

Recent improvements to Zocalo include support for the RabbitMQ message broker, and support for running services on Kubernetes, both of which help further improve the high-availability and fault-tolerance of the system. System metrics can be exported via Prometheus enabling monitoring

of system performance and automatic alerting of potential problems with Grafana and Alertmanager.

**Email address of presenting author:**

richard.gildea@diamond.ac.uk

44

## Elettra.chat: from Chat Service to Logbook For Experiments

**Author:** Roberto Borghes<sup>1</sup>

**Co-authors:** Alessandro Olivo<sup>2</sup>; Daniele Favretto<sup>1</sup>; Emiliano Coghetto<sup>2</sup>; Fulvio Billè<sup>1</sup>; Georgios KOUROUSIAS<sup>1</sup>; Ivan Andrian<sup>1</sup>; Roberto Pugliese<sup>1</sup>

<sup>1</sup> Elettra Sincrotrone Trieste

<sup>2</sup> CERIC-ERIC

**Corresponding Author:** roberto.borghes@elettra.eu

Elettra.chat is a live chat service introduced in 2020 in order to improve and facilitate communication among internal teams of Elettra-Sincrotrone Trieste. It is based on the open source Rocket.chat platform and it is fully integrated with the Elettra information system called Virtual Unified Office (VUO). Elettra.chat is intuitive, easy to use and offers multi-platform support, multi-user channels, roles, permissions, file uploading and a REST API. The initial goal of Elettra.chat was to support the remote working but it has proved to be a perfect collaborative tool also for user experiments. In this perspective, the integration between Elettra.chat and VUO has grown considerably over the past two years: for every scientific proposal all its participants are automatically added to a new private chat channel. The contents of the experiment chat channel can be exported in pdf format and saved in the central data storage. In addition, thanks to the REST API, user information can be integrated with metadata coming directly from the acquisition system and screenshots provided by a client application running on the beamline workstations. This work presents and describes all the steps that transformed Elettra.chat from a simple chat service to a logbook for experiments.

**Email address of presenting author:**

roberto.borghes@elettra.eu

**I agree to recordings of my presentation being made at NOBUGS 2022:**

45

## A SEDCNN Machine Learning Model for Textured SAXS/WAXD Image Denoising

**Author:** Zhongzheng Zhou<sup>1</sup>

**Co-author:** Chun Li<sup>1</sup>

<sup>1</sup> Institute of High Energy Physics, Chinese Academy of Sciences

**Corresponding Author:** lichun@ihep.ac.cn

With the advancements on instrumentations of next-generation synchrotron light sources, methodologies for small angle x-ray scattering (SAXS)/wide angle x-ray diffraction (WAXD) experiments have dramatically changed. Such experiments have evolved into dynamic and multi-scale in-situ

characterizations, leaving prolonged exposure time as well as radiation-induced damage a serious concern. However, reduction on exposure time and dose may result in noisy images with much lower signal-to-noise ratio, thus requiring powerful denoising mechanisms for information retrieval. Here, we tackle the problem from an algorithmic perspective by proposing a small, yet effective encoder-decoder-structured machine learning model for experimental SAXS/WAXD image denoising, allowing more room for exposure time and dose adjustment. From preprocessing to architecture design and final performance evaluation, our network provides a bespoke denoising solution for SAXS/WAXD experimental images. Compared with classic image processing models like U-Net, REDCNN, and PMRID for natural images, our proposed model demonstrates superior performance on highly textured SAXS/WAXD images.

**Email address of presenting author:**

lichun@ihep.ac.cn

**I agree to recordings of my presentation being made at NOBUGS 2022:**

46

## Taking the pain out of generating complete and compliant NeXus files for dynamic setups

**Authors:** Kenan Muric<sup>1</sup>; Tobias Richter<sup>1</sup>

<sup>1</sup> *European Spallation Source*

**Corresponding Author:** kenan.muric@ess.eu

NeXus has a powerful and fairly complete structure to hold the geometry of the experiment.

The data reduction package Scipp (links: <https://scipp.github.io/index.html> and <https://scipp.github.io/scippneutron/index.html>) can process data automatically based on the embedded NeXus geometry, that allows scientists to view the histogrammed data in scientific coordinates. Generating an accurate NeXus hierarchy with all the required information is cumbersome, especially for experiments where the setup changes frequently. There are no software tools that provide meaningful support to verify that geometry nodes are correctly connected or that the naming convention is followed according to the NeXus standard.

These points are addressed by the NeXus constructor (link: <https://github.com/ess-dmnc/nexus-creator>) developed and used at the European Spallation Source (ESS) for creating, editing and visualizing NeXus template files. These JSON template files are used by the data file writer to populate NeXus files with information from the data acquisition pipeline. The NeXus constructor also provides a visual confirmation in 3D that the geometry is correct. The interpretation of the instrument geometry in the NeXus constructor has been confirmed to agree with Scipp. That makes successful processing of every written NeXus file very likely at the first attempt and removes the need for a trial-and-error approach.

NeXus constructor is written in Python/Qt. It displays a NeXus tree structure and an instrument 3D view for online visualization of NeXus groups that represent physical components in an instrument. This makes it possible for users to verify the location of different parts of the instrument when adding translations and rotations to individual instrument components. Base geometry shapes are provided by the application for rendering. If the shape of an object has a complex structure, as is the case for many neutron detectors, it is possible to load arbitrary geometries in OFF format as supported by NeXus.

The application supports adding and naming data fields, which can be static or to be filled in from the data acquisition. The NeXus standard naming and data types are always suggested as default options, so making the file compliant is the easy route.

**Email address of presenting author:**

kenan.muric@ess.eu

47

## An Integrated Data Acquisition and Analysis Software for XRF Mapping Experiments

**Author:** xiaoxue bi<sup>1</sup>

<sup>1</sup> *The Institute of High Energy Physics of the Chinese Academy of Sciences*

**Corresponding Author:** bixx@ihep.ac.cn

Mapping experimental mode is very common due to the increasing macro and nanoprobe beamlines in synchrotron facilities. Here we present a python-based software system with integrated data acquisition, analysis and visualization functions developed for a XRF microscopy beamline in Beijing Synchrotron Radiation Facility (BSRF). The control and acquisition part is based on the Mamba framework developed for the future High Energy Photon Source (HEPS). Multiple scanning modes including flying scans are incorporated in the software with a user-friendly Graphical User Interfaces (GUI). PyMCA toolkit is integrated with the acquisition module for real-time data elements distribution analysis. The software has recently been deployed at the XRF beamline of BSRF. With minor modification it has also been implemented at another STXM beamline. The successful deployment of the software lay a firm foundation for the software requirements in future mapping experiments at HEPS.

**Email address of presenting author:**

bixx@ihep.ac.cn

48

## Synchronization of commercial camera data at the European XFEL

**Author:** Andrea Parenti<sup>1</sup>

**Co-authors:** Gabriele Giovanetti <sup>2</sup>; Riccardo Fabbri <sup>1</sup>; Steffen Hauf <sup>1</sup>; Valerii Bondar <sup>2</sup>

<sup>1</sup> *European XFEL GmbH*

<sup>2</sup> *European XFEL*

**Corresponding Author:** andrea.parenti@xfel.eu

Commercial cameras are extensively used at the European XFEL, by the scientific instruments and for beam diagnostics. In order to correlate the data coming from different sources, the train ID is used as a primary key; the train ID counts the number of X-Ray trains supplied since the beginning of the facility operation, is monotonically increasing with a period of 100 ms, and is uniquely defined for the entire facility. Unfortunately, the train ID information coming from the XFEL timing system can generally not be injected into the timing information provided by the commercial cameras in use at the facility.

In this contribution, a reliable and reproducible solution to tag images received from Ethernet cameras with the correct train ID is described.

**Email address of presenting author:**

andrea.parenti@xfel.eu

I agree to recordings of my presentation being made at NOBUGS 2022:

49

## Experiences with Datalogging to InfluxDB at the European XFEL

**Author:** Gero Flucke<sup>1</sup>

**Co-authors:** Alessandro Silenzi<sup>2</sup>; Ana García-Tabarés<sup>2</sup>; Andrea Parenti<sup>1</sup>; Anna Klimovskaia<sup>2</sup>; Ayaz Samadli<sup>2</sup>; Christopher Youngman<sup>2</sup>; David Hickin<sup>2</sup>; Dennis Göries<sup>2</sup>; Florian Sohn<sup>2</sup>; Gabriele Giovanetti<sup>2</sup>; Gianpietro Previtali<sup>1</sup>; Ivars Karpics<sup>2</sup>; Janusz Szuba<sup>1</sup>; Jawad Tahmeed Bin Taufik<sup>2</sup>; Michael Smith<sup>2</sup>; Milan Staffehl<sup>2</sup>; Noushadali Anakkappalla<sup>2</sup>; Raul Costa<sup>2</sup>; Sergey Esenov<sup>2</sup>; Steffen Hauf<sup>1</sup>; Valerii Bondar<sup>2</sup>; Wajid Ehsan<sup>2</sup>

<sup>1</sup> European XFEL GmbH

<sup>2</sup> European XFEL

**Corresponding Author:** gero.flucke@xfel.eu

The event driven Karabo 1 control system has been developed at the European XFEL to operate instruments and photon beam lines. More than ten thousand software devices interact as interfaces to hardware devices, provide system services or high level automation procedures. A core system service is data logging that continuously stores any changes of configurations and slow control data. Since 2020 this data is stored in an InfluxDB 2 database which makes it easily available via the powerful Grafana web interface 3.

This contribution describes the InfluxDB setup with its transparent integration into Karabo, the experiences gained since it is in operation and the resulting adjustments to the setup.

1 Hauf, Steffen, et al. "The Karabo distributed control system." *Journal of synchrotron radiation* 26.5 (2019): 1448-1461.

2 InfluxDB, InfluxDATA 2021, <https://www.influxdata.com>

3 Grafana, Grafana Labs 2021, <https://grafana.com>

**Email address of presenting author:**

gero.flucke@xfel.eu

50

## KaraboGui - The Cockpit of the Supervisory Control and Data Acquisition System Karabo

**Author:** Dennis Göries<sup>1</sup>

**Co-authors:** Alessandro Silenzi<sup>1</sup>; Ana García-Tabarés<sup>1</sup>; Andrea Parenti<sup>2</sup>; Anna Klimovskaia<sup>1</sup>; Ayaz Samadli<sup>1</sup>; Cammille Carinan<sup>1</sup>; Christopher Youngman<sup>1</sup>; David Hickin<sup>1</sup>; Florian Sohn<sup>1</sup>; Gabriele Giovanetti<sup>1</sup>; Gero Flucke<sup>2</sup>; Hugo Santos<sup>1</sup>; Ivars Karpics<sup>1</sup>; Jawad Tahmeed Bin Taufik<sup>1</sup>; Michael Smith<sup>1</sup>; Milan Staffehl<sup>1</sup>; Noushadali Anakkappalla<sup>1</sup>; Raul Costa<sup>1</sup>; Sergey Esenov<sup>1</sup>; Steffen Hauf<sup>2</sup>; Valerii Bondar<sup>1</sup>; Wajid Ehsan<sup>1</sup>

<sup>1</sup> European XFEL

<sup>2</sup> European XFEL GmbH

**Corresponding Author:** dennis.goeries@xfel.eu

At the European XFEL, the inhouse developed Supervisory Control and Data Acquisition System (SCADA) Karabo has been steering and facilitating scientific experiments at the photon beamlines

since the free-electron laser started its operation in 2017.

A single Graphical User Interface (GUI), the so-called KaraboGui, has been designed as a multi-purpose application and is the preferred entry point to the control system. Implementing a nowadays standard event-driven asynchronous server-client approach, this application based on Python 2 and Qt 3 is the prime choice for hardware and experiment control and as well configuring detector calibration and data acquisition. Equipped with a generic panel builder and a package updater for external component extensions, the KaraboGui matches the requirements of a present and future graphical user interface for the experiments.

All user-interface panels build in the KaraboGui can be translated into a Scalable Vector Graphics (SVG) representation and loaded or edited with an SVG graphics program (e.g. Inkscape 4).

This contribution describes in detail the client application KaraboGui and its usage at the European XFEL with an outlook of a web representation.

#### References

1 Hauf, Steffen, et al. "The Karabo distributed control system." *Journal of synchrotron radiation* 26.5 (2019): 1448-1461.

2 Python. <https://docs.python.org/3/library/>

3 Qt Framework. <https://www.qt.io/>

4 Inkscape. <http://www.inkscape.org/>

**Email address of presenting author:**

dennis.goeries@xfel.eu

**I agree to recordings of my presentation being made at NOBUGS 2022:**

51

## User interfaces at the PolLux STXM beamline

**Author:** Benjamin Watts<sup>1</sup>

**Co-authors:** Jörg Raabe <sup>1</sup>; Simone Finizio <sup>1</sup>

<sup>1</sup> *PSI - Paul Scherrer Institut*

**Corresponding Author:** benjamin.watts@psi.ch

Scanning transmission X-ray spectro-microscopy (STXM) involves both nanoscale imaging and spectroscopy that can cover multiple absorption edges over a wide energy range. The technique presents users with a dizzying array of parameters that control the dimensions of scans and details of the scanning process. Important quality measures such as the spatial and spectral resolution also depend on these parameters in complicated ways. Users who do not fully understand how to operate the instrument will tend to make suboptimal choices and rely heavily on the beamline staff. The PolLux STXM beamline is addressing these issues with a number of custom software tools:

- \* the Pixelator STXM control software (C++)
- \* scripting interface to Pixelator (Python)
- \* the PolLux Calculator (Python)
- \* the Focus Finder GUI (Python)
- \* thumbnail images embedded in data files (XMP in HDF5)

Here, we will present these software tools and their usage at the PolLux STXM beamline.

**Email address of presenting author:**

benjamin.watts@psi.ch

**I agree to recordings of my presentation being made at NOBUGS 2022:**

52

## SciCat Present and Future

**Author:** Carlo Minotti<sup>1</sup>

**Co-authors:** Massimiliano Novelli<sup>2</sup>; Stephan Egli<sup>1</sup>

<sup>1</sup> *PSI - Paul Scherrer Institut*

<sup>2</sup> *European Spallation Source*

**Corresponding Author:** carlo.minotti@psi.ch

This presentation will introduce SciCat to the general audience, a metadata catalogue developed as a collaboration between PSI, ESS and MAXIV. We will present the guiding principles and the needs that this project strives to address.

We will list the features currently implemented, from ingestion to publication, with an emphasis on the user experience. We will also showcase some external tools that leverage the SciCat API

We will showcase how SciCat is used at PSI and highlight the technical challenges with the underlying software architecture, and the integration work that has been done when a given metadata catalogue, specifically SciCat, is deployed to a given facility.

We will cover the topic of CI/CD driven deployments, in particular for parallel installations into different test and production environments based on Kubernetes.

We will conclude this presentation with the future roadmap for the project.

**Email address of presenting author:**

carlo.minotti@psi.ch

53

## SDU: Software for high throughput automated data collection at the Swiss Light Source

**Authors:** Kate Mary Louise Smith<sup>1</sup>; Ezequiel Panepucci<sup>1</sup>; Jakub Kaminski<sup>2</sup>; Justyna Aleksandra Wojdyla<sup>1</sup>

<sup>1</sup> *PSI - Paul Scherrer Institut*

<sup>2</sup> *Open Eye*

**Corresponding Author:** kate.smith@psi.ch

Presented by Kate Smith on behalf of the SLS MX Group.

The Swiss Light Source Macromolecular Crystallography Group operates three beamlines (X06SA, X06DA and X10SA) served by an in-house developed distributed DA+ software stack, which supports standard and sophisticated data acquisition and analysis[1,2]. Recent hardware upgrades include the TELL robot with increased dewar capacity and sample exchange speed<sup>3</sup>, new TELL gripper design with pin detection, implementation of the fast fragment- and compound-screening pipeline (FFCS)[4,5], new sample environment top camera and backlight, and the installation of SmarGon+ MCS2 with in-house SmargoPolo controls software and calibration routine.

Recent software developments include the extension of DA+ software microservice architecture, implementation of sample spreadsheet validation in TELL GUI and a user web application, deployment of automated loop centering routines, addition of native-sad merging to automatic data processing, and the migration of our samples database to the cloud. In this talk I will present how the



advancements in hardware and software were leveraged in implementing the sophisticated communication and decision making software, Smart Digital User (SDU), for fully automated data collection[6].

1. Wojdyla JA, Kaminski JW, Panepucci E, Ebner S, Wang X, Gabadinho J, et al. DA+ data acquisition and analysis software at the Swiss Light Source macromolecular crystallography beamlines. *J Synchrotron Rad.* 2018 Jan 1;25(1):293–303.
2. Basu S, Kaminski JW, Panepucci E, Huang CY, Warshamanage R, Wang M, et al. Automated data collection and real-time data analysis suite for serial synchrotron crystallography. *J Synchrotron Rad.* 2019 Jan 1;26(1):244–52.
3. Martiel I, Buntschu D, Meier N, Gobbo A, Panepucci E, Schneider R, et al. The TELL automatic sample changer for macromolecular crystallography. *J Synchrotron Rad.* 2020 May 1;27(3):860–3.
4. Kaminski JW, Vera L, Stegmann DP, Vering J, Eris D, Smith KML, et al. Fast fragment- and compound-screening pipeline at the Swiss Light Source. *Acta Cryst D.* 2022 Mar 1;78(3):328–36.
5. Sharpe ME, Wojdyla JA. Fragment-Screening and Automation at the Swiss Light Source Macromolecular Crystallography Beamlines. *Nihon Kessho Gakkaishi.* 2021 Aug 31;63(3):232–5.
6. Smith KM, Panepucci E, Kaminski K, ..., Wojdyla JA. SDU: Software for high throughput automated data collection at the Swiss Light Source. Manuscript in preparation.

**Email address of presenting author:**

kate.smith@psi.ch

54

## Adopting NICOS at SINQ

**Author:** Mark Koennecke<sup>1</sup>

**Co-author:** Michele Brambilla<sup>1</sup>

<sup>1</sup> *PSI - Paul Scherrer Institut*

**Corresponding Author:** mark.koennecke@psi.ch

At SINQ and ANSTO SICS is used as instrument control software. SICS dates back to 1996 and is showing its age. It was decided to switch the SINQ instruments to a combination of EPICS and NICOS. The transition to NICOS is now nearly complete. The authors will present the lessons learned from the collaboration on SICS, from the transition to NICOS and will attempt to quantify the benefits of collaboration.

**Email address of presenting author:**

mark.koennecke@psi.ch

**I agree to recordings of my presentation being made at NOBUGS 2022:**

Yes

55

## A MHz sampling DAQ system for sub-second QEXAFS at the SLS-2.0 "Debye" Beamline

**Author:** Alvin Samuel Acerbo<sup>1</sup>

**Co-authors:** Adam Hugh Clark<sup>1</sup>; Maarten Nachtegaal<sup>1</sup>; Stephan Hitz<sup>1</sup>

<sup>1</sup> PSI - Paul Scherrer Institut

**Corresponding Author:** alvin.acerbo@psi.ch

The upcoming "Debye" beamline at the SLS-2.0 will provide continuous sub-second X-ray absorption spectroscopy and co-located X-ray diffraction under operating conditions with a photon energy range of 4.5 to 60 KeV. Based on a highly successful design at the SuperXAS beamline, the Debye QEXAFS monochromator is designed to produce spectra of monochromatic X-rays at up to 10 Hz by continuous oscillation of the Bragg axis. Continuous sampling of the monochromator Bragg angle and detector channels using National Instruments (NI) hardware is done by oversampling the Bragg angle motor encoder channel and analog input channels (i.e. ion chambers) at a high bit resolution. Low level software controls are provided by the NI supported Python API for the NI-DAQmx library. High level controls of the DAQ pipeline will be implemented in a GUI and will allow for fine control over DAQ parameters, along with selection of input channels. Lossless data reduction of the data stream and optional descriptive statistics are generated in real-time, after which the resulting I0 and I1 signals can be ratioed to produce absorption spectra that are readily analyzed by standard spectroscopy methods, or displayed on consoles.

**Email address of presenting author:**

alvin.acerbo@psi.ch

**I agree to recordings of my presentation being made at NOBUGS 2022:**

56

## Next generation User Office software

**Authors:** Fredrik Bolmsten<sup>1</sup>; Thomas Attwood<sup>2</sup>

<sup>1</sup> ESS

<sup>2</sup> STFC

**Corresponding Authors:** fredrik.bolmsten@ess.eu, thomas.attwood@stfc.ac.uk

The European Spallation Source is developing the next generation of user office software in collaboration with STFC. Handling proposal submission and review as well as experiment scheduling and sample safety declaration the software will address many of the tasks required of the user office for research infrastructures.

The software builds on existing user office software packages, addressing the need for dynamic content, configurable proposal workflows, API first and mobile compatibility. From its inception, the guiding design principle has been to be transparent to the facility: with flexibility and adaptability to allow the software to be used in a number of settings without significant time or cost consuming development work.

In this presentation we will showcase some of the design decisions and technology choices made to achieve the above mentioned goals and the challenges that come with them. We will also outline the roadmap for future development and present our vision of a seamless user journey through a facility.

**Email address of presenting author:**

fredrik.bolmsten@ess.eu

**I agree to recordings of my presentation being made at NOBUGS 2022:**

57

## **PaN-Training e-Learning: education and training for scientists and students**

**Authors:** Alexandre Stefanov<sup>1</sup>; Andrew McCluskey<sup>1</sup>; Kareem Galal<sup>2</sup>; Lottie Greenwood<sup>3</sup>; Mads Bertelsen<sup>1</sup>; Peter Willendrup<sup>4</sup>; Stella d'Ambrumenil<sup>1</sup>; Teodor Ivanoaica<sup>5</sup>; Thomas Rod<sup>1</sup>

<sup>1</sup> *European Spallation Source ERIC*

<sup>2</sup> *European Spallation*

<sup>3</sup> *European Spallation Source*

<sup>4</sup> *Technical University of Denmark, Physics Department*

<sup>5</sup> *Extreme Light Infrastructure ERIC*

**Corresponding Author:** andrew.mccluskey@ess.eu

Training of new neutron and photon scientists is an important aspect in the longevity of large-scale facilities. Typically, these training take place at centrally organised training events. These events are (by necessity) selective, with only space for a limited number of students per iteration.

The PaN-training e-Learning platform (<https://e-learning.pan-training.eu>) looks to increase access to interactive training in the neutron and photon sciences. This platform gives access to a Moodle training platform where rich content can be prepared and shared with students. The students are then able to work through this material in their own time.

In addition to offering this MOOC-style experience (Massive Open Online Course). The e-Learning platform can also be used to complement in-person events, allowing for “flipped learning”-approaches. This is where students are asked to work through material independently to facilitate discussion during in-person sessions.

Alongside the standard Moodle interface, the PaN-training e-Learning platform provides access to a JupyterHub server. This makes the platform perfect for use at neutron or photon training courses focused on topics such as data reduction, analysis or management (where Python programming is frequently used).

If you are interested in using the PaN-Training e-Learning platform for your training course or preparing a MOOC, please get in touch: [admin@pan-learning.org](mailto:admin@pan-learning.org).

**Email address of presenting author:**

[andrew.mccluskey@ess.eu](mailto:andrew.mccluskey@ess.eu)

58

## **Data streaming at SINQ**

**Author:** Michele Brambilla<sup>1</sup>

**Co-author:** Mark Koennecke<sup>1</sup>

<sup>1</sup> *PSI - Paul Scherrer Institut*

**Corresponding Author:** michele.brambilla@psi.ch

In the last years SINQ has undergone a major upgrade both from the hardware and the software point of view. The neutron flux has been increased thanks to new neutron guides, instruments have been upgraded (AMOR, DMC) and new instruments will be installed (SANS-LLB, Falcon). NICOS and EPICS replaced the old control software (SICS).

At the same time, we are undergoing a paradigm shift in the data acquisition sector. From the old “histogramming” approach we are transitioning (where possible) to “event streaming” mode. This is the first step toward time-resolved experiments.

In this talk we will present the status of data streaming at SINQ and the current limitations.

**Email address of presenting author:**

michele.brambilla@psi.ch

**I agree to recordings of my presentation being made at NOBUGS 2022:**

Yes

59

## Ewoks –a meta-workflow system

**Authors:** Wout De Nolf<sup>1</sup>; Giannis Koumoutsos<sup>2</sup>; Olof Svensson<sup>1</sup>; Henri Payno<sup>1</sup>; Julia Garriga Ferrer<sup>1</sup>; Andrew GOETZ<sup>1</sup>

<sup>1</sup> Engineer @ ESRF

<sup>2</sup> Engineer at ESRF

**Corresponding Authors:** wout.de\_nolf@esrf.fr, giannis.koumoutsos@esrf.fr

There are many workflow management systems which provide standalone frameworks for implementing tasks, creating graphs of tasks and executing these graphs. Instead of choosing one of these and taking the risk that it gets outdated after a few years the ESRF has opted for a meta approach to workflows. The ESRF's Workflow Management System (ewoks) *was developed to provide an abstraction layer between graph representation and execution. This allows using the same tasks' and graphs' definitions in different workflow management systems. It is focused on interoperability and binds together several existing solutions into a flexible meta-framework able to deal with acyclic and cyclic directed graphs as well as sub-graphs i.e. graphs within graphs. Currently, bindings have been developed for: Orange - a desktop graphical interface, pypushflow - a task scheduler of cyclic graphs, Dask - a parallel computing library for task scheduling and Ewoks - a simple job scheduler. A web application is provided to create, visualize, persist, execute and monitor the execution of ewoks workflows in the web. The talk will present the architecture of ewoks, examples of how it is being used to automate processing at the ESRF and demonstrate creating and running a workflow with ewoks.* <https://gitlab.esrf.fr/workflow/ewoks>

**Email address of presenting author:**

wout.de\_nolf@esrf.fr

60

## Mantid Imaging

**Authors:** Dolica Akello-Egwel<sup>1</sup>; Sam Tygier<sup>1</sup>

<sup>1</sup> STFC**Corresponding Author:** sam.tygier@stfc.ac.uk

Neutron imaging instruments, such as IMAT at ISIS in the UK, require dedicated software for pre-processing projection data and reconstructing it into 3D volumes using filtered back projection or iterative methods. Mantid Imaging puts powerful tools for noise reduction, artefact removal, alignment, and advanced iterative reconstruction methods in the hands of scientific users without requiring knowledge of programming. Mantid Imaging builds on algorithms provided by libraries including ASTRA Toolbox, Core Imaging Library and Tomopy in a cross platform Qt GUI. It can be installed locally or deployed via remote desktop systems such as the ISIS Data Analysis as a Service to give users access to sufficient resources to handle large datasets. Mantid Imaging has allowed IMAT to migrate away from proprietary software that was no longer supported. We present the software, show examples of its use at IMAT, and discuss planned extensions of Mantid Imaging for energy-resolved neutron imaging.

**Email address of presenting author:**

sam.tygier@stfc.ac.uk

61

## Environmental sustainability for scientific software

**Author:** Sam Tygier<sup>1</sup><sup>1</sup> STFC**Corresponding Author:** sam.tygier@stfc.ac.uk

UN Secretary-General António Guterres described the Intergovernmental Panel on Climate Change's (IPCC) Sixth Assessment Report as a "code red for humanity". Urgent CO2 emission reductions are needed, and it is therefore important to consider the environmental sustainability of everything that we do including our actions as researchers and software engineers. ICT is responsible for between 2 and 4 % of global greenhouse gas emissions when full life cycles are considered. The high rate of growth in computing makes it hard to predict if expected efficiency improvements will be sufficient to bring down emissions. This talk will explore the sources of GHG emissions from computing, and how changes in what we write and how we run it can make a difference. We also consider the context of computing, how does computing effect the operation of facilities and rebound effects.

**Email address of presenting author:**

sam.tygier@stfc.ac.uk

62

## Fast analysis feedback with automated data processing pipelines at PETRA-III

**Authors:** Vijay Kartik<sup>1</sup>; Michael Sprung<sup>1</sup>; Anton Barty<sup>2</sup><sup>1</sup> DESY<sup>2</sup> Coherent Imaging Division, CFEL, DESY

**Corresponding Author:** vijay.kartik@desy.de

The high data rates of next-generation X-ray detectors coming into use at PETRA-III beamlines have triggered intense activity on the topics of 'live' data processing and fast feedback to beamline users.

The vision is to provide users analyzed and reduced data in near-real time, which can be used to judge whether their data acquisition is producing meaningful output or not. This can in turn help them to tune their data acquisition or to stop and modify their current experiment/sample setup before continuing the experiment further. Fast analysis feedback is essential to keep the experimenter in the loop, able to make agile decisions based on the science without being overwhelmed by the large amounts of raw data being produced by the detectors.

We present MENTO, a data processing toolkit that is automatically triggered during data acquisition, and which remotely runs external data analysis software on-demand using the DESY high-performance computing (HPC) cluster, 'Maxwell'. The processed results are transparently made available at the beamline so that users can immediately evaluate the experiment without having to manually handle any raw data. MENTO is set up to require no input from the users except to point to the desired analysis, and the entire processing pipeline is then managed automatically, including data input, access to the HPC cluster, job submissions to a batch processing scheduler, and result writing. MENTO integrates easily with the experiment control systems currently used at DESY, and is in production at a few PETRA-III beamlines already. At the coherence applications beamline P10 in particular, it is augmented with a graphical user interface for visualizing reduced and pre-processed data, giving the user direct real-time visual feedback about the integrity of the acquired data, potentially flagging beam damage to samples, and thus pre-empting acquiring unusable data, allowing users to make the most of their limited time at the beamline.

**Email address of presenting author:**

vijay.kartik@desy.de

63

## Towards real-time data reduction in serial-crystallography

**Author:** Jerome Kieffer<sup>1</sup>

**Co-authors:** Alejandro Homs <sup>1</sup>; Daniele de Sanctis <sup>1</sup>; Gianluca Santoni <sup>1</sup>; Nicolas Coquelle <sup>1</sup>; Samuel Debionne <sup>1</sup>; Shibom Basu <sup>2</sup>

<sup>1</sup> ESRF

<sup>2</sup> EMBL

**Corresponding Author:** jerome.kieffer@esrf.fr

We present a novel image analysis of diffraction frame which is applied to macro-molecular serial crystallography.

This new signal separation algorithm is able to distinguish the amorphous (or powder diffraction) component from the diffraction signal originating from single crystals. It relies on the ability to work efficiently in azimuthal space and derives from the work performed on pyFAI, the fast azimuthal integration library.

Two applications are presented: a lossy compression algorithm and a peak-picking algorithm; the performances of both is assessed and compared to state of the art reference implementations: XDS and CrystFEL.

**Email address of presenting author:**

Jerome.kieffer@esrf.fr

64

## Experiment data streaming and aggregation at the European Spallation Source

**Author:** Daniel Cacabelos<sup>1</sup>

**Co-authors:** Afonso Haruo Carnielli Mukai <sup>1</sup>; Morten Jagd Christensen <sup>1</sup>; Tobias Richter <sup>1</sup>

<sup>1</sup> *European Spallation Source*

**Corresponding Author:** daniel.cacabelos@ess.eu

The European Spallation Source will provide a collection of instruments for state-of-the-art neutron science. To support the scientific activities, high-throughput instrument data, environment parameters and other metadata must be timely collected, aggregated and made available for real-time and offline analysis. The data aggregation and streaming system comprises an Apache Kafka stream processing data store, collection and aggregation adapters for the different data sources, persistence adapters for the NeXus format, and integrations with the experiment control user interfaces. We present the architecture of this system, its current integration status and our latest performance results.

**Email address of presenting author:**

daniel.cacabelos@ess.eu

**I agree to recordings of my presentation being made at NOBUGS 2022:**

65

## The new tomography data processing software of the ESRF

**Authors:** Henri Payno<sup>1</sup>; Pierre Paleo<sup>1</sup>

**Co-authors:** Christian Nemoz <sup>1</sup>; Peter Cloetens <sup>1</sup>; Marco Di Michiel <sup>1</sup>; Alexander Rack <sup>1</sup>; Paul Tafforeau <sup>1</sup>; V. Armando Solé <sup>2</sup>; Nicola Viganò <sup>1</sup>

<sup>1</sup> *ESRF - The European Synchrotron*

<sup>2</sup> *ESRF*

**Corresponding Author:** henri.payno@esrf.fr

The ESRF (European Synchrotron) hosts several tomography beamlines, which use different types of contrast and offer vastly different resolutions and performance. Their acquisition and processing software are exceedingly fragmented. There exist many different codes and solutions for similar problems. This hampers the exchange and processing of data collected on different beamlines, while its associated maintenance / development costs hindered its progress.

The tomography acquisition system and processing workflows are being re-developed from scratch. The goal is to modernize them, both in terms of offered features and used technologies, while rendering their maintenance easier and less resource demanding.

In particular, we aim at delivering both a homogeneous experience across all the beamlines, and robust high-performance processing software.

In this contribution, we present the new tomographic data processing software of the ESRF. The new software suit consists of mainly three projects, which address three different aspects of the data treatment: data conversion & management, data processing & reconstruction, and graphical user interface. The three corresponding software packages are: NXtomomill, Nabu and TomWer.

NXtomomill guarantees an identical output data format for each ingested raw data format and data type. It decouples data handling from data reconstruction, resulting in uniform user experience, easier development, reduced maintenance costs.

Nabu is a high performance tomographic reconstruction software. It is derived from the popular PyHST, but it is built on modern software technologies, design patterns, and development strategies. It is modular, and it has a low deployment burden.

TomWer is a workflow based GUI for building and automatizing tomographic reconstructions. It greatly decreases the steepness of the learning curve for performing tomographic reconstructions (from raw data to volumes). It also allows to define data processing and reconstruction workflows, which can be later deployed on a large number of datasets.

**Email address of presenting author:**

henri.payno@esrf.fr

66

## Scientific Data Management at European XFEL.

**Authors:** Bharathi Vanganuru<sup>1</sup>; Carsten Patzke<sup>2</sup>; Christian Voss<sup>2</sup>; Christopher Youngman<sup>1</sup>; Djelloul Boukhelef<sup>1</sup>; Fabio Dall'Antonia<sup>1</sup>; Gianpietro Previtali<sup>1</sup>; Illia Derevianko<sup>1</sup>; Janusz Malka<sup>1</sup>; Janusz Szuba<sup>1</sup>; Juergen Hannappel<sup>2</sup>; Kars Ohrenberg<sup>2</sup>; Kilian Schwarz<sup>2</sup>; Krzysztof Wrona<sup>1</sup>; Luis Maia<sup>1</sup>; Martin Gasthuber<sup>2</sup>; Maurizio Manetti<sup>1</sup>; Mwai Karimi<sup>2</sup>; Nasser Alqudami<sup>1</sup>; Philipp Schmidt<sup>1</sup>; Steve Aplin<sup>1</sup>; Thomas Kluyver<sup>1</sup>; Tigran Mkrtchyan<sup>2</sup>; Uwe Ensslin<sup>2</sup>

<sup>1</sup> *European XFEL GmbH*

<sup>2</sup> *Deutsches Elektronen-Synchrotron DESY*

**Corresponding Author:** janusz.malka@xfel.eu

Data Management is essential to make research data easily accessible and usable. Important ingredients of data management include data policies and data workflows.

The data workflows are based on the policies which are implemented by defining a set of parameters stored in the metadata catalogue. The role of the metadata catalogue in relation to the data management services and underlying hardware solutions for the data storage systems will be presented. The architecture of the storage system consists of four layers, each addressing a different set of challenges. The first –online - is designed as a fast cache for the data generated directly at the scientific instruments during experiments. The second layer –offline - provides the performance for data processing during and after the beamtimes. The third layer - dCache disk pool - delivers the capacity to the system for long-term storage and the last one - tape archive - provides data safety and long-term archive. The storage system is able to accept 2PB/day of raw data, demonstrating the real capabilities with all sub-services being involved in this process. The storage system is connected to the high-performance computing cluster supporting remote data analysis and alternatively allows external users to export data outside of the European XFEL facility.

**Email address of presenting author:**

janusz.malka@xfel.eu

67

## High-Rate Data Acquisition, Streaming, and Processing at the EMBL PETRA III Beamlines



**Authors:** Gleb Bourenkov<sup>1</sup>; Marina Nikolova<sup>1</sup>

**Co-authors:** Andrea d'Amato <sup>1</sup>; Elizabeth Duke <sup>1</sup>; Ivars Karpics <sup>2</sup>; Jonas Albers <sup>1</sup>; Kirill Kovalev <sup>1</sup>; Stefan Fiedler <sup>1</sup>; Thomas R. Schneider <sup>3</sup>

<sup>1</sup> EMBL

<sup>2</sup> European XFEL

<sup>3</sup> EMBL c/o DESY

**Corresponding Author:** nikolova@embl-hamburg.de

The EMBL beamlines, P14, TREXX, P13, P12, support macromolecular crystallography, SAXS and X-ray imaging experiments. A common denominator between these techniques is the usage of high-frame-rate megapixel detectors and cameras. In an optimal scenario the data that they produce would immediately undergo some preliminary analysis and provide feedback to the users, allowing them to make informed decisions about the sample, the beamline environment and the further progress of the experiment. We have designed our software system and computing and hardware infrastructure with these goals in mind placing a particular emphasis on robustness at high loads. In recent developments we introduced a multi-node data stream receiver structure enabling us to transfer, process and store data at the maximum frame rate of the DECTRIS EIGER 2 X 4M detector of 1120 frames per second in 8-bit mode. The consolidated results of the crystallographic analysis using DOZOR 1 are displayed in real-time in the GUI, MXCuBE 2, and are additionally used to generate live highlights around the reflections in the diffraction viewer, ADXV 3.

The backbone of the system is a 40Gbit InfiniBand, a petabyte parallel cluster file system BeeGFS storage and a collection of servers that take care of acquisition and experiment control, stream receiving, data processing, image tracking, etc. A concerted effort has been made to harmonize workflows and re-use computing infrastructure across experiment types. To this end, acquisition, data stream sending and receiving software has been developed for an X-ray imaging camera, PCO –emulating the architecture of our EIGER setup. The streamed frames are flat-field corrected and subsequently displayed to the user in real-time. Improvements to the system to achieve also a live tomographic reconstruction are in progress.

1 Popov, A. N. & Bourenkov, G. (2016). DOZOR. ESRF, Grenoble, France

2 Oscarsson, et al. J. Synchrotron Rad. (2019). 26, 393-405

3 Arvai, A. (2015). The ADXV User Manual, <https://www.scripps.edu/tainer/arvai/adxv/AdxvUserManual.pdf>

**Email address of presenting author:**

nikolova@embl-hamburg.de

68

## A common control and readout software for different X-ray detector systems

**Author:** Dhanya Thattil<sup>1</sup>

**Co-authors:** Aldo Mozzanica <sup>1</sup>; Anna Bergamaschi <sup>1</sup>; Bernd Schmitt <sup>1</sup>; Carlos Lopez Cuenca <sup>2</sup>; Christian Ruder ; Davide Mezza ; Dominic Greiffenberg <sup>3</sup>; Erik Fröjdh <sup>1</sup>; Filippo Baruffaldi ; Jiaguo Zhang <sup>1</sup>; Julian Heymes ; Konstantinos Moustakas <sup>1</sup>; Mar Carulla <sup>3</sup>; Martin Brückner <sup>4</sup>; Pawel Kozlowski ; Rebecca Barten ; Roberto Dinapoli <sup>3</sup>; Shqipe Hasanaj ; Simon Ebner ; Sudip Shrestha ; Thomas King ; Viktoria Hinger <sup>1</sup>

<sup>1</sup> PSI - Paul Scherrer Institut

<sup>2</sup> SLS Detectors

<sup>3</sup> Paul Scherrer Institut

<sup>4</sup> Paul-Scherrer-Institut

**Corresponding Author:** dhanya.thattil@psi.ch

The slsDetectorPackage is a control and readout software for the high performance X-ray detectors developed at the Paul Scherrer Institute. It is an attempt to provide a common and flexible interface for a family of detectors ranging from small 1D detectors (1280 channels) to large pixel detectors (16M). The core is written in C++ with a custom command line interface and Python binding for scripting. The software is in use at several facilities worldwide that use the detectors developed at PSI. Since our software is open-source, drivers for EPICS, TANGO and Karabo have already been developed by collaborators.

Our main challenge is the diversity of detectors and computer environments that the software has to support: dimensions (both 1d and 2d), scalability (1 up to 36 modules), Ethernet interfaces (1g and 10g), dynamic range (4, 8, 16 and 32 bits per pixel), synchronized with master and slave architecture or independent modules, file formats (binary and hdf5), different packet sizes, and top it off with different detector specific features.

To a large extent, we encapsulate the detector specific behavior to a part of the software that runs on the detector readout board CPU, but not all aspects can be tackled in the same manner.

Furthermore, multiple ways exist to control the detector system via the C++ or Python API, the command line and the Qt based GUI. There are various ways to pick up the data via files, zmq streams or call back routines. The system is designed to run on a variety of data backend platforms and the data receiver part of the software could be replaced with a custom process. We are currently examining different avenues into easier online data analysis, despite the different detector types.

Such a multifaceted system for constantly evolving research applications comes with obvious challenges demanding constant improvements. We will present an overview of the current architecture, challenges and future improvements.

**Email address of presenting author:**

dhanya.thattil@psi.ch

**I agree to recordings of my presentation being made at NOBUGS 2022:**

69

## Bluesky at the Advanced Photon Source

**Author:** Peter Jemian<sup>1</sup>

<sup>1</sup> Argonne National Laboratory, Advanced Photon Source

**Corresponding Author:** jemian@anl.gov

In a few short months from now, the Advanced Photon Source Upgrade will begin a complete reconstruction and upgrade of the storage ring. The upgrade involves a long dark period where no experiments will be possible. Significant changes will be made to many beam lines. This dark time offers the APS a major interval in which to retool the beam experiment controls with advancements not possible during regular operations and maintenance periods.

Developed to coordinate the complete scientific data life cycle, from measurement through analysis, Bluesky has been adopted at the Advanced Photon Source as the software framework for user interface. Bluesky is written with Python, leveraging the existing collection of packages and shared experience. Bluesky was created for initial operations at NSLS-II and has since been adopted by other scientific user facilities across the US and now enjoys an international collaboration.

This presentation will describe how the APS will be using Bluesky, on top of its existing EPICS controls, to orchestrate the next generation of data science.

**Email address of presenting author:**

jemian@anl.gov

70

## Muon Galaxy –an open web platform for computational muon science

**Author:** Eli Chadwick<sup>1</sup>

**Co-authors:** Alejandra Gonzalez-Beltran ; Anish Mudaraddi <sup>1</sup>; Jyothish Thomas <sup>1</sup>; Leandro Liborio <sup>1</sup>; Patrick Austin ; Simone Sturniolo <sup>2</sup>

<sup>1</sup> *Science and Technology Facilities Council, UK Research and Innovation*

<sup>2</sup> *CoMind (formerly Science and Technology Facilities Council, UK Research and Innovation)*

**Corresponding Author:** eli.chadwick@stfc.ac.uk

The Muon Spectroscopy Computational Project (MSCP) is an initiative that currently includes members of the Theoretical and Computational Physics Group and the Data and Software Engineering Group in the Scientific Computing Department, STFC and members of the Muon Group at ISIS, STFC. The main objective of the MSCP is to support users of muon sources via the development of a sustainable and user-friendly set of software tools and a software platform that can be used for interpreting muon experiments. We are relying on the Galaxy platform to achieve some of these goals.

Galaxy is an open, web-based platform for accessible, reproducible, and transparent computational research. It originated in the bioinformatics community but now spans many research domains. The Galaxy interface allows users to run analysis workflows, preserve them in a reproducible way, and share or publish them, all without the need to know programming or the command line.

Muon Galaxy is where these two projects meet. The MSCP develops several command-line software tools for muon science, and the Galaxy platform is ideal for providing a graphical interface to these tools. We (members of the MSCP) will present our tools, our work integrating those tools into the Galaxy platform, and our progress launching Muon Galaxy as an STFC service available to all.

We will also demonstrate the Muon Galaxy interface and show how the platform's features help us to reliably reproduce published results. We'll discuss our ideas and plans for connecting Muon Galaxy up to other infrastructure such as STFC's computational resources and public repositories for materials science data.

Finally, we'll promote a new materials science Galaxy subcommunity to connect with others with an interest in applying Galaxy to X-ray, neutron, and muon science and materials science in general.

**Email address of presenting author:**

eli.chadwick@stfc.ac.uk

71

## Muon Galaxy –an open web platform for computational muon science

**Author:** Eli Chadwick<sup>1</sup>

**Co-authors:** Alejandra Gonzalez-Beltran ; Anish Mudaraddi <sup>1</sup>; Jyothish Thomas <sup>1</sup>; Leandro Liborio <sup>1</sup>; Patrick Austin <sup>1</sup>; Simone Sturniolo <sup>2</sup>

<sup>1</sup> *Science and Technology Facilities Council, UK Research and Innovation*

<sup>2</sup> *CoMind (formerly Science and Technology Facilities Council, UK Research and Innovation)*

**Corresponding Author:** eli.chadwick@stfc.ac.uk

The Muon Spectroscopy Computational Project (MSCP) is an initiative that currently includes members of the Theoretical and Computational Physics Group and the Data and Software Engineering Group in the Scientific Computing Department, STFC and members of the Muon Group at ISIS, STFC. The main objective of the MSCP is to support users of muon sources via the development of a sustainable and user-friendly set of software tools and a software platform that can be used for interpreting muon experiments. We are relying on the Galaxy platform to achieve some of these goals.

Galaxy is an open, web-based platform for accessible, reproducible, and transparent computational research. It originated in the bioinformatics community but now spans many research domains. The Galaxy interface allows users to run analysis workflows, preserve them in a reproducible way, and share or publish them, all without the need to know programming or the command line.

Muon Galaxy is where these two projects meet. The MSCP develops several command-line software tools for muon science, and the Galaxy platform is ideal for providing a graphical interface to these tools. We (members of the MSCP) will present our tools, our work integrating those tools into the Galaxy platform, and our progress launching Muon Galaxy as an STFC service available to all.

We'll also show examples of published results we've reproduced using Muon Galaxy and explain how the platform's features help this to be accomplished reliably. We'll discuss our ideas and plans for connecting Muon Galaxy up to other infrastructure such as STFC's computational resources and public repositories for materials science data.

Finally, we'll promote a new materials science Galaxy subcommunity to connect with others with an interest in applying Galaxy to X-ray, neutron, and muon science and materials science in general.

**Email address of presenting author:**

eli.chadwick@stfc.ac.uk

72

## Machine-learning driven beamline alignment at EuXFEL

**Authors:** Felix Brauße<sup>1</sup>; Liubov Samoylova<sup>2</sup>

**Co-authors:** Alexey Buzmakov<sup>3</sup>; Carsten Fortmann-Grote<sup>4</sup>; Maurizio Vannoni<sup>1</sup>

<sup>1</sup> *European XFEL*

<sup>2</sup> *Europaen XFEL*

<sup>3</sup> *FSRC "Crystallography and Photonics"*

<sup>4</sup> *Max Planck Institute for Evolutionary Biology*

**Corresponding Author:** liubov.samoylova@xfel.eu

EuXFEL is a large scale laser facility which operates seven different instruments and the eighth is coming into operation now. All the instruments have a few hundred meters multi-component optical setups, which includes grazing incidence offset mirrors and focusing elements, such as CRL or KB mirrors. To increase efficiency of operation, automation of the beamline alignment procedure has a great importance to deliver the XFEL radiation with its unique properties to the experiment.

At EuXFEL the SiMEX platform 1 for simulating FEL experiment was developed and operated. Here we present an extended approach, in which a Convolutional Neural Network (CNN) model is trained with beam-profile simulations extracted from FEL simulations 2 in combination with the wavefront propagation package 3. The CNN model is used to estimate and optimize the alignment parameters of the optical components, in order to implement digital shadowing and drastically facilitate the alignment of the XFEL beamlines. The first results of application of the technique to one of the EuXFEL beamlines will also be presented.

1. <https://github.com/PaNOSC-ViNYL/SimEx>
2. AIP Conference Proceedings 2054, 030019 (2019)
3. Journal of Applied Crystallography 08/2016; 49(4) pp.1347-1355. doi:10.1107/S160057671600995X, <https://github.com/samoylv/WPG>

**Email address of presenting author:**

liubov.samoylova@xfel.eu

**I agree to recordings of my presentation being made at NOBUGS 2022:**

73

## Managing Experiment Data with Ease at the Advanced Photon Source

**Author:** Hannah Parraga<sup>None</sup>

**Co-authors:** John Hammonds<sup>1</sup>; Nicholas Schwarz<sup>1</sup>; Sinisa Veseli<sup>1</sup>; Steven Henke<sup>1</sup>

<sup>1</sup> *Argonne National Laboratory*

**Corresponding Author:** hparraga@anl.gov

### Managing Experiment Data with Ease at the Advanced Photon Source

Hannah Parraga, Sinisa Veseli, John Hammonds, Steven Henke, Nicholas Schwarz

*Advanced Photon Source, Argonne National Laboratory, 9700 S Cass Ave, Lemont, IL 60439, USA*

Data are essential to the scientific discoveries enabled by experiments performed at the Advanced Photon Source (APS). At present, the APS generates approximately 10 PB of raw experimental data per year from its sixty-eight operating beamlines that house over 100 unique instruments. This data is generated as a part of over 6,000 annual experiments performed by over 5,500 facility users each year. Similar to other synchrotrons, the amount of data generated at the APS continues to quickly increase due to beamline advances, such as new measurement techniques, technological advances in detectors and instrumentation, multi-modal instruments that can acquire several measurements in a single experiment, and advanced data processing algorithms. This trend is expected to continue in the future.

As a scientific user facility, the APS presents several unique challenges for data management. Beamlines can perform multiple experiment techniques, use different types of detectors, produce data at different rates, use multiple data formats, use machines with different operating systems, and execute various processing workflows. Additionally, the users themselves vary. They come from different research institutions, universities, and industries, but all must be able to access their data after leaving the facility. They may want their data immediately or several years after it is created. They may be conducting experiments independently and remotely, or in person with close involvement with beamline staff. Also, beamline scientists have different levels of technical expertise. Some desire a hands-off approach to data management and some want the flexibility to program custom tools. The APS must have a data management solution that addresses these unique challenges.

This presentation covers the features of the APS Data Management System, which provides tools that beamline staff can use to support their users. A command line interface and graphical user interface give users the ability to upload data to long-term storage. The built-in workflow engine allows data to be processed with any given set of shell commands. Using Globus, data is secure but also accessible to users from their home institution. Data is secured on local beamline machines with tools for managing file system permissions. Furthermore, users are able to catalog metadata to include additional information about the experiments alongside their results.

Although the APS Data Management System addresses many needs, further development is underway of additional features to provide users with an improved data management experience. This includes streaming data directly from detectors to storage to decrease the transfer time as data rates and volumes continue to increase. Workflows are being developed which publish to common data portals for visualizing results. Interfacing with the tape archives of the Argonne Leadership Computing Facility will allow more data to be stored for longer periods.

\*Work supported by U.S. Department of Energy, Office of Science, under Contract No. DE-AC02-06CH11357.

**Email address of presenting author:**

hparraga@anl.gov

74

## **Unified recording of Channel Access and beam-synchronous data for PSI**

**Author:** Dominik Werder<sup>1</sup>

<sup>1</sup> *PSI*

**Corresponding Author:** dominik.werder@psi.ch

Data from Channel Access and beam-synchronous sources is currently archived to storage by several different software packages, each with their own on-disk file formats. In order to improve reliability, reduce complexity, ensure maintainability and pave the way for new features, a replacement is under development which uses ScyllaDB as the data store. A prototype of this is being tested at SwissFEL.

**Email address of presenting author:**

dominik.werder@psi.ch

**I agree to recordings of my presentation being made at NOBUGS 2022:**

75

## **EasyDiffraction: A user-friendly interface for diffraction data analysis**

**Author:** Andrew Sazonov<sup>1</sup>

**Co-authors:** Simon Ward<sup>1</sup>; Piotr Rozyczko<sup>1</sup>

<sup>1</sup> *European Spallation Source ERIC*

**Corresponding Author:** andrew.sazonov@ess.eu

Diffraction is a key tool for structure analysis. However, currently available software for modelling and analysis of diffraction data may be, on the one hand, difficult for new users looking to apply diffraction to their field of expertise and, on the other hand, not flexible enough for domain experts.

EasyDiffraction 1 aims to lower the barrier of entry to diffraction data analysis by providing an intuitive and user-friendly graphical interface, which is distributed as an all-in-one package that includes all dependencies and can be installed with just a few clicks on different operating systems. For more complex problems and increased flexibility the Python library behind EasyDiffraction can be used through Jupyter notebooks and scripting.

Simple interface of EasyDiffraction can help improve the user experience and thereby make it easier to train users and students, as well as be better prepared for experiments. We plan to integrate EasyDiffraction into the full data processing workflow to increase experiment automation and make better use of beam time.

EasyDiffraction is built on the EasyScience framework 2, a platform aimed at unifying neutron scattering analysis software. In addition to diffraction, this framework has been successfully applied to reflectometry. Quasielastic neutron scattering will also be targeted in the future.

Currently EasyDiffraction has the basic features of CrysPy 3 and CrysFML 4 crystallographic libraries. We are collaborating with LLB and ILL regarding the CrysPy and CrysFML, respectively, and more functionality will become available as the project matures.

EasyDiffraction is being developed free and open source, keeping the idea of FAIR and sustainable software in mind. We hope to attract interested people to jointly contribute to this project and help us, for the benefit of everyone, in making diffraction data analysis and modelling easier.

1 <https://easydiffraction.org>

2 <https://easyscience.software>

3 <https://github.com/ikibalin/cryspy>

4 <https://code.ill.fr/scientific-software/crysfml>

**Email address of presenting author:**

Andrew.Sazonov@ess.eu

76

## Scaling Diffuse Scattering Workflows with Hybrid HPC Workflows

**Authors:** Guy Jennings<sup>1</sup>; Justin Wozniak<sup>2</sup>; Matthew Krogstad<sup>1</sup>; Ray Osborn<sup>1</sup>; Stephan Rosenkranz<sup>1</sup>

<sup>1</sup> ANL

<sup>2</sup> Argonne National Laboratory

**Corresponding Author:** woz@anl.gov

Data analysis pipelines for diffuse scattering workflows consist of various steps with differing requirements for computation time and user interaction. The NXrefine workflow system is a semi-automated Python GUI toolkit based around NeXpy and the NeXus data format for diffuse x-ray scattering and other applications. The user is able to orchestrate many analysis pipelines on different datasets concurrently, distributing work to typical Linux clusters. Its most computationally intensive component is a coordinate transform implemented by the previously developed CCTW application. In practical, live data collection/analysis use cases, a user team may end up with a large backlog of CCTW work capable of exploiting 32K threads for tens of minutes. Executing this

workload immediately is important to inspect data quality and other application goals during data collection, but it is difficult to gain access to adequate computational resources to perform this analysis. In this presentation, we describe a new hybrid HPC component which distributes this work to an MPI-enabled workflow system running on Theta or other HPC systems. We describe the portability and scalability of this component with respect to the diffuse scattering application as well as more general workloads.

This work was supported by the U.S. Department of Energy, Office of Science, Advanced Scientific Computing Research, under contract number DE-AC02-06CH11357 and the Office of Basic Energy Sciences, Division of Materials Sciences and Engineering.

The following text will be removed in the final submission:

The submitted manuscript has been created by UChicago Argonne, LLC, Operator of Argonne National Laboratory ("Argonne"). Argonne, a U.S. Department of Energy Office of Science laboratory, is operated under Contract No. DE-AC02-06CH11357. The U.S. Government retains for itself, and others acting on its behalf, a paid-up nonexclusive, irrevocable worldwide license in said article to reproduce, prepare derivative works, distribute copies to the public, and perform publicly and display publicly, by or on behalf of the Government. The Department of Energy will provide public access to these results of federally sponsored research in accordance with the DOE Public Access Plan. <http://energy.gov/downloads/doe-public-accessplan>

**Email address of presenting author:**

woz@anl.gov

77

## The UK Ada Lovelace Centre

**Author:** Robert McGreevy<sup>1</sup>

<sup>1</sup> UKRI/STFC

**Corresponding Author:** robert.mcgreevy@stfc.ac.uk

The Ada Lovelace Centre (ALC) will be a centre of expertise in scientific software engineering and data management, helping UK national facilities (Diamond Light Source, ISIS Neutron and Muon Source, Central Laser Facility, Scientific Computing Department) to maximise scientific and economic impact along all parts of the data chain e.g. machine learning techniques to efficiently filter high value data at the beamline, advanced simulation techniques to interpret analysed data, and long term curation and re-use of aggregated data. At the same time ALC will help to reduce the related environmental (energy) impact, and associated cost, of facility and computing resource operation. A number of collaborative projects have been carried out under the ALC 'brand' over the last few years, but funding is now available to expand the activity and properly establish the centre. This contribution will outline the current status and future plans and prospects.

**Email address of presenting author:**

robert.mcgreevy@stfc.ac.uk



78

## CamServer: Data Processing Pipelines @ SwissFEL

**Author:** Alexandre Gobbo<sup>1</sup>

**Co-author:** Andrej Babic<sup>2</sup>

<sup>1</sup> PSI

<sup>2</sup> PSI - Paul Scherrer Institut

**Corresponding Author:** alexandre.gobbo@psi.ch

CamServer is a Python package currently used at SwissFEL for running data processing pipelines. The system is deployed in a cluster of servers and handles 100Hz high-resolution camera images and other generic ZMQ streams. It can also align different data streams and images before processing. Standard processing pipelines are available out of the box (e.g. calculation of standard beam metrics), but users can also upload their custom scripts. The output data streams are used for a variety of purposes: transient and permanent data storage, image visualisation, DAQ applications and beamline specific tools. The system is managed through a REST API using either a management GUI or a Python client library. Current developments include adding support for detector data, pushing logs and metrics to Elastic and creating an IDE to simplify pipeline development.

**Email address of presenting author:**

alexandre.gobbo@psi.ch

**I agree to recordings of my presentation being made at NOBUGS 2022:**

79

## smargopolo: A new control system for SmarGon/MCS2 using ROS

**Author:** Wayne Glettig<sup>1</sup>

**Co-authors:** Dominik Buntschu<sup>2</sup>; Ezequiel Panepucci<sup>2</sup>

<sup>1</sup> PSI - Paul Scherrer Institut

<sup>2</sup> Paul Scherrer Institute

**Corresponding Author:** wayne.glettig@psi.ch

A new control system has been developed for the SmarAct GmbH multi-axis goniometer SmarGon. SmarGon is a six degree-of-freedom positioning device, allowing positioning of a sample and orientation around any given point. It was purpose build for protein-crystallography experiments, but, as will be presented here, was also re-purposed for other applications.

Due to limitations in SmarGon's initial control system, which was based around Delta Tau's PPMAC, a new control system "smargopolo" was developed, based on the open source robotics framework ROS (Robot Operating System) for high-level control, in connection with SmarAct's MCS2 controller for low-level control. The internals of the system will be presented.

This architecture allows strong customisation, mainly regarding interfaces, coordinate systems, logging, debugging and visualisation tools. Also, new calibration routines could be realised, tested and tweaked for optimal practical use.

Because SmarGon's predecessor PRIGo was developed at PSI, several concepts from PRIGo could be reused, allowing tight integration into the overall beamline control software, leading in practice to a better overall reliability of the system.

**Email address of presenting author:**

wayne.glettig@psi.ch

80

## Integrated Real-Time Auto-Processing at Diamond Light Source

**Author:** Jacob Filik<sup>1</sup>

**Co-authors:** Dean Keeble <sup>1</sup>; James Hall <sup>1</sup>; Karl Levik <sup>1</sup>; Peter Chang <sup>1</sup>; Richard Gildea <sup>2</sup>; Satheesh Maheswaran <sup>1</sup>; Tim Snow <sup>1</sup>

<sup>1</sup> *Diamond Light Source*

<sup>2</sup> *Diamond Light Source Ltd*

**Corresponding Author:** jacob.filik@diamond.ac.uk

As data rates and experimental complexity increase it is critical that facilities reduce the burden of reduction and processing of raw experimental data for users. While this statement itself is clear and simple, the reality behind implementing generic real-time auto-processing is not.

Broadly the problem can be split into two categories: Infrastructure and User Experience. Infrastructure requirements include things like data and metadata storage and access, cross process communication between different systems, and access to High Performance Computing resources. User Experience deals with how can these separate systems come together to provide a flexible and usable system, but most critically –how do we give facility users the confidence that the data is processed correctly, and with full provenance, so that they will use it for real-time experimental decision making or as the basis for publication?

Here we present the system deployed for the Physical Sciences at the Diamond Light Source and show how technologies like HDF5 (with SWMR), message brokers (like ActiveMQ) and information management systems (like ISPyB), can be used to build a versatile system for generic real-time data processing.

**Email address of presenting author:**

jacob.filik@diamond.ac.uk

81

## DIALS as a Toolkit

**Author:** Graeme Winter<sup>1</sup>

**Co-authors:** Ben Williams <sup>1</sup>; David McDonagh <sup>2</sup>; David Waterman <sup>2</sup>; Elena Pascal <sup>1</sup>; Gwyndaf Evans <sup>1</sup>; James Bielsten-Edmands <sup>1</sup>; Markus Gerstel <sup>1</sup>; Nicholas Devenish <sup>1</sup>; Richard Gildea <sup>1</sup>

<sup>1</sup> *Diamond Light Source*

<sup>2</sup> *CCP4*

**Corresponding Author:** graeme.winter@diamond.ac.uk

The DIALS software for the processing of X-ray diffraction data is presented, with an emphasis on how the suite may be used as a toolkit for data processing. The description starts with an overview of the history and intent of the toolkit, usage as an automated system, command-line use, and ultimately how new tools can be written using the API to perform bespoke analysis. Consideration is also made to the application of DIALS to techniques outside of macromolecular X-ray crystallography.

**Email address of presenting author:**

graeme.winter@diamond.ac.uk

82

**PyBluIce: Modular data acquisition software for long duration, scripted automation, and high speed protein crystallography experiments****Authors:** Mark Hilgart<sup>1</sup>; Qingping Xu<sup>1</sup>; Oleg Makarov<sup>1</sup>; Sergey Stepanov<sup>1</sup>; Michael Becker<sup>1</sup>; Nagarajan Venugopalan<sup>1</sup>; Craig Ogata<sup>1</sup>; David Kissick<sup>1</sup>; Robert Fischetti<sup>1</sup><sup>1</sup> Argonne National Laboratory**Corresponding Author:** mhilgart@anl.gov

In order to make new methods for protein crystallography experiments widely accessible at GM/CA@APS, the data acquisition program JBluIce was rewritten with improved modularity as the primary design goal. A comprehensive HTTP and Redis-based interface covering all levels of beamline functions enables a range of applications to control the beamline. This includes single use scripts for trying new methods, as well as a permanent GUI with decades of refinement built in which will be extended with new features including serial crystallography and automated planning. The API and plug-in architecture allow for more complex data collection modes by creating a shared framework to remove complexity. Automation is designed in a way to make it highly visible to the user and allow them to step in at any time. Viewing and analysis of data is decoupled from the data collection engine and adapts to any collection rate. And finally, a realistic beamline simulation with genuine EPICS, HTTP, SQL and Redis servers allows for most development to be done without a beamline.

**Email address of presenting author:**

mhilgart@anl.gov

**I agree to recordings of my presentation being made at NOBUGS 2022:**

83

**SOLEIL digital transformation toward the upgrade of the facility****Author:** Yves-Marie ABIVEN<sup>1</sup><sup>1</sup> synchrotron SOLEIL**Corresponding Author:** yves-marie.abiven@synchrotron-soleil.fr

SOLEIL offers with its 29 beamlines a wide energy range (from THz to hard X-ray), a large variety of experimental techniques, and sample environments to its users. SOLEIL's teams started detail studies for an upgrade of the facility, which will offer new performances for the accelerators increased in brilliance, coherence and flux. These changes will be accompanied by new access modes in the multidisciplinary environment proposed to our Users communities. Toward this upgrade, the information system teams are building continuously the digital transformation required to handle the new operation modes and the increasing data volume. To address the future scientific challenges in an open data perspective, automated and data-driven processes are under development. The poster will depict the information system strategy choose to transform our enterprise architecture and the ongoing tasks among which we can find :

- improvement of operational and organizational working approach,
- integration of service-oriented technology and cloud,
- upgrade of users interface supervision and monitoring technology,
- and integration of artificial intelligence for control, maintenance and data reduction and analysis.

**Email address of presenting author:**

yves-marie.abiven@synchrotron-soleil.fr

**I agree to recordings of my presentation being made at NOBUGS 2022:**

84

## SciLog –An Electronic Logbook for User Experiments

**Authors:** Klaus Wakonig<sup>1</sup>; Stephan Egli<sup>1</sup>

<sup>1</sup> *PSI - Paul Scherrer Institut*

**Corresponding Author:** klaus.wakonig@psi.ch

Properly capturing raw and meta-data during an experiment is rightfully given a high priority. Yet, it is the logbook that aids in putting the decisions made during the experiment and thus also the acquisition strategy itself into context. However, logbooks are frequently lacking a good integration into facility-specific services such as authentication and data acquisition systems and often end up as a burden, especially in stressful situations during an experiment.

SciLog, a logbook system based on MongoDB, Loopback and Angular, aims to alleviate these constraints by providing a flexible and extensible environment as well as a simple and intuitive user interface. At its base it relies on atomic entries in a NoSQL database that can be queried, sorted and displayed to the user's requirements. An integration with facility-specific authorisation systems and the automatic import of new experiment proposals allow for a user experience that is specifically targeted for the challenging environment of experiments at large research facilities.

**Email address of presenting author:**

klaus.wakonig@psi.ch

**I agree to recordings of my presentation being made at NOBUGS 2022:**

85

## SLS 2.0 Controls and Science IT Sub Project Status Update

**Authors:** Alun Ashton<sup>1</sup>; Elke Zimoch<sup>1</sup>; Marie Xingxing Yao<sup>1</sup>; Stefan Fries<sup>1</sup>

<sup>1</sup> *PSI - Paul Scherrer Institut*

**Corresponding Authors:** alun.ashton@psi.ch, xingxing.yao@psi.ch

The Swiss Light Source upgrade project (called SLS 2.0) started with the focus on the diffraction-limited storage-rings (DLSRs) for a fourth generation synchrotron light source. The main expected improvements from the machine upgrade is significantly reduced electron beam's horizontal emittance, thus higher proton beam brilliance for hard x-ray beamlines and better photon beam coherent fraction for insertion-devices. Along with recent research and development on photon detectors and data acquisition, our flagship beamlines are looking forward to potentially higher resolution and faster experiment throughput. This talk will outline the scope of the SLS 2.0 Controls and Science

IT Sub Project; share our challenges and opportunities; as well as our plans and early progresses toward meeting the exponentially increased demands, which include beamline experiment control, data and metadata collection and processing, and computing and storage infrastructure.

**Email address of presenting author:**

alun.ashton@psi.ch; xingxing.yao@psi.ch

86

## **#29 Daiquiri: a web based user interface for beamline control and data acquisition**

**Corresponding Author:** stuart.fisher@esrf.fr

89

## **#22, The Mamba Data Acquisition Software Project for HEPS**

**Corresponding Author:** zhangyi88@ihep.ac.cn

97

## **Welcome by Thierry Straessle, PSI Deputy Director**

129

## **Discussion and NOBUGS 2022 - Closing**

131

## **C2 Data Viewer: Visualization tool for EPICS7 Data Streaming**

**Authors:** Elaine Chandler<sup>1</sup>; Guobao Shen<sup>1</sup>; Sinisa Veseli<sup>1</sup>; Thomas Fors<sup>1</sup>

<sup>1</sup> Argonne National Laboratory

**Corresponding Author:** echandler@anl.gov

A high-performance data acquisition system (DAQ) has been under active development to meet APS-U needs. It takes data from underneath FPGA (Field Programmable Gate Array) and streams it to its downstream users. The APS-U DAQ system software framework is implemented as a major portion of APS-U new control system software infrastructure, which is called C2. To visualize the DAQ data on the fly, C2 Data Viewer (C2DV) is implemented using Python, which can be used for displaying live PV data streams for monitoring, troubleshooting and diagnostics purposes. It is now capable of handling both EPICS pvAccess (PVA) and Channel Access (CA) data and includes several different applications: a scope viewer for plotting PVA waveforms, an image viewer for displaying

Area Detector image data, and a striptool for monitoring PVA as well as CA scalar PVs. In this presentation, we discuss various C2DV features, its usage at the Advanced Photon Source, as well as plans for future development.

**Email address of presenting author:**

[echandler@anl.gov](mailto:echandler@anl.gov)