NOBUGS 2024

Monday, September 23, 2024 - Friday, September 27, 2024

ESRF Auditorium

Book of Abstracts

Version of 15/07/2024

Contents

Satellite Meeting / 1

Controls and Acquisition GUI Strategies

ESRF - BEL-1-01

Satellite Meeting / 19

Mantid developers meeting

MD-1-21

Satellite Meeting / 160

Workflow engines for FAIR analysis of photon, neutron and muon data

ILL50 110

Satellite Meeting / 18

pyFAI user meeting

ESRF LOB1-45

Satellite Meeting / 159

Bluesky Community Meeting (Data Acquisition)

LOB 1-45

Satellite Meeting / 100

Mantid developers meeting

Satellite Meeting / 20

Introduction to data acquisition and on-line data analysis with BLISS and EWOKS

https://indico.esrf.fr/event/137/
Please follow the link for a detailed description and the program for the meeting!

Satellite Meeting / 21

Roadmap for PaN open data to become FAIR AI/ML ready

ESRF BEL-1-01

Invited Speakers / 164

The state of AI/ML in photon and neutron science

Author: Paul Quinn[1]

[1] *Ada Lovelace Centre*

TBD

AI/ML applications / 151

Large language model-based AI agent as an intelligent assistant system for experimental data acquisition at fourth-generation synchrotron radiation light source

Authors: Xiaoxue Bi[1]; Zhibang Shen[None]; Xinyu Pan[2]; Yi Zhang[None]

[1] *Institute of High Energy Physics, Chinese Academy of Sciences*

[2] *The Institute of High Energy Physics of the Chinese Academy of Sciences*

Due to the diverse data acquisition modes and complex online analysis methods conducted at various beamlines of synchrotron radiation light sources, beamline users are often required to get acquainted with the interface, functionality and workflow of the data acquisition software before the experiment starts. Such process highly relies on the on-site guidance from the beamline staff themselves with routine explanation on instructional documentation to the users, which can be laborious and time-consuming. Here, based on a large language model, we propose an AI agent as an intelligent assistant system of the data acquisition software for synchrotron light source beamlines. The system utilizes large language model as a command hub to parse the beamline users'linguistic description of the experimental process, assisting users to customize experimental process and parameters setting through its knowledge base and comprehension capabilities. The system interactively exchanges information and data with Mamba, the data acquisition software framework for fourth-generation synchrotron radiation source (HEPS), to facilitate experimental process control in an automated and intelligent manner that significantly lowers the learning curve of the data acquisition software. Additionally, the system provides an intelligent Q&A functionality based on large language model to assist beamline experiments. Through interactive Q&A sessions, the required material, structural and physical information can be obtained from the pre-trained large language

model and the self-built knowledge base. Also, together with the experimental data, the system treats the queried information as metadata and deposits it into files through information exchange with Mamba, further enhancing the data's completeness and usability.

Abstract publication:

I agree that the abstract will be published on the web site

AI/ML applications / 81

Tiled at Beamlines: Enhancing Data Access for Machine Learning-Driven Scientific Workflows

Authors: Dylan McReynolds[1]; Daniel Allan[2]

Co-authors: Wiebke Koepp [3]; Guanhua Hao [4]; Petrus Zwart [4]; Dilworth Parkinson [4]; Tanny Chavez [4]; Xiaoya Chong [4]; Jacob Filik [5]; Sharif Ahmed [5]; Tim Snow [5]; Alexander Hexemer [4]

[1] *Lawrence Berkeley National Lab*

[2] *BNL*

[3] *Advanced Light Source (ALS), Lawrence Berkeley National Lab (LBNL)*

[4] *LBNL*

[5] *Diamond Light Source*

From the Bluesky project, Tiled [1] is a data service that removes several barriers by providing secure, authenticated remote access to data. Tiled abstracts variations in file formats and other data storage details across different beamline instruments, making data analysis and visualization code portable. It enables fast, targeted access to specific data regions and offers search and filtering capabilities over metadata. Designed for a wide range of use cases—from analysis notebooks to web-based visualization applications—Tiled is ideally suited for AI/ML workflows at facilities. It provides a unified interface across datasets with varied formats and access control requirements.

The MLExchange project [2], co-developed by LBNL, ORNL, ANL, BNL, and SLAC, builds machine learning tools that facilitate ML analysis at user facilities, focusing on software that supports beam-time experiments. This framework includes browser-based user interfaces along with ML training and inference workflows.

MLExchange applications were recently enhanced to use Tiled for centralized data management at several workflow stages. This updated framework has been installed locally at the DIAD beamline at Diamond Light Source and tested during a tomography beamtime. As data was acquired, tomographic reconstructions were performed and automatically integrated into MLExchange to be used in the Segmentation Application [3]. Once introduced, Tiled fed the reconstructed frames to a browser-based user application, where users can create segmentation classes using several convenient drawing tools. Classes created in the Segmentation Application were then stored in Tiled and made available to the DLSIA[4] framework, which provides access to various segmentation and denoising neural networks. DLSIA conducted training and segmentation, drawing datasets from Tiled and writing results back to Tiled. These results were then available in the Segmentation Application for users to browse. The MLExchange Segmentation application was well received and will soon be installed and used at the Advanced Light Source during a tomography beamtime, where we will scan the same samples and use the models trained at Diamond to segment data reconstructions at ALS.

[1] Bluesky Collaboration. Tiled https://github.com//bluesky/tiled.

[2] Z. Zhao et al., "MLExchange: A web-based platform enabling exchangeable machine learning workflows for scientific studies," 2022 4th Annual Workshop on Extreme-scale Experiment-in-the-Loop Computing (XLOOP), Dallas, TX, USA, 2022, pp. 10-15, doi: 10.1109/XLOOP56614.2022.00007.

[3] Hao G, Roberts EJ, Chavez T, et al. Deploying Machine Learning Based Segmentation for Scientific Imaging Analysis at Synchrotron Facilities. IS&T Int Symp Electron Imaging. 2023;35:IPAS-290. doi:10.2352/ei.2023.35.9.ipas-290

[4] Roberts, E. J., Chavez, T., Hexemer, A., & Zwart, P. H. (2024). Dlsia: Deep learning for scientific image analysis. Journal of Applied Crystallography, 57(2).

Abstract publication:

I agree that the abstract will be published on the web site

AI/ML applications / 63

Closed-loop experiments using ML-based online analysis at synchrotron beamlines: A case study in x-ray reflectometry and perspectives

Author: Linus Pithan[1]

[1] *DESY*

Modern synchrotron beamlines and neutron instruments have undergone significant changes due to technological advances and newly deployed infrastructure. Thus, experiments are becoming more data-intense and data-driven and increasingly relying on online data analysis for efficient use of experimental resources. In this regard, machine-learning (ML) based approaches of specific importance for real-time decision-making based on online data analysis and connected closed loop feedback applications.

Here we focus on a case study in x-ray reflectometry performed at ESRF using BLISS and TANGO to operate an autonomous experiment in closed-loop operation with an underlying ML model. We discuss infrastructure aspects as well as the use of ML-models in real time data analysis, essentially allowing to transfer the time spend on data analysis to a point in time prior the actual experiment.

Looking ahead, specifically in view of planed upgrade to Petra IV at DESY and the RockIT project, we also try to give some more general perspectives on the interplay of ML and autonomous experiments in beamline control environments.

Pithan et al., J. Synchrotron Rad. (2023). 30, 1064-1075
https://doi.org/10.1107/S160057752300749X

Munteanu et al., J. Appl. Cryst. (2024). 57, 456–469
https://doi.org/10.1107/S1600576724002115

Abstract publication:

I agree that the abstract will be published on the web site

AI/ML applications / 43

Planning on using AI to help neutron scattering experiments in ANSTO

Authors: Nick Hauser[1]; Norman Xiong[1]

[1] *ANSTO*

The Australian Centre for Neutron Scattering (ACNS) at ANSTO aims to incorporate advanced artificial intelligence (AI) methodologies into neutron scattering experiments in a novel approach, with the goal of enhancing overall experimental efficiency. This integration will assist users throughout each stage of their experiment, including planning, data collection, analysis, and logging.

The proposed strategy is organized into three stages:

1. In the first stage, an AI model will be integrated within our instrument control software to provide guidance on running neutron scattering experiments through a console user interface (UI) component for seamless communication. The AI model will offer advice in both scientific and engineering aspects of the experiment.
2. During the second stage, we aim to enhance the capabilities of the AI model so that it can directly issue commands to our control system and generate codes for data collection.
3. In the third stage, a more interactive approach will be introduced by incorporating voice and image inputs into the user experience.

This project is currently in its early stages, with efforts focused on identifying suitable models and hardware infrastructure for this initiative.

Abstract publication:

I agree that the abstract will be published on the web site

AI/ML applications / 123

Appliction of Google TPU-fined Adam Algorithm and Huawei NPU CANN Mindspore Toolkit in Physics-Informed Neural Network Training for Ptychography

Author: LEI WANG[1]

Co-authors: Jianli Liu [2]; Rui Liu [3]; Shiyuan Fu [3]; Shuang Wang [3]; Yu Hu [4]

[1] *Institute of High Energy Physics*

[2] *Computing Center, Institute of High Energy Physics, Chinese Academy of Sciences*

[3] *IHEP*

[4] *IHEP, CAS*

Google team showed a new Adam-tuned optimization solver for deep neural network training called LION (EvoLved Sign Momentum) by thousands of hours of training at TPU cluster, 2023. It is more memory-efficient than Adam as it only keeps track of the momentum and cuts the epsilon and a group of momentum parameters off. We applied "LION"solver to the physics-informed neural network-PtychoPINN which advanced one more step than PtychoNN from APS with physics constrained (probe matrix computed by ePIE first) in the loss function. The PtychoPINN algorithm is now used in High Energy Photon Source(HEPS), China, the fourth generation source with less than 20% GPU-memory occupied and 3X less time consumed.
Apart from the new LION solver, we moved this training process from Nvidia A100 to 8*Huawei Ascend 910A GPU which is called NPU(Neural Processing Unit). The CANN Mindspore toolkit were used and it could achieve 70% performance of A100. The training process would be revealed in the meeting.

Abstract publication:

I agree that the abstract will be published on the web site

AI/ML applications / 61

A Systematic Deep Learning-empowered Denoising Solution for Synchrotron X-ray Diffraction and Scattering Images

Authors: Chun Li[1]; Zhongzheng Zhou[2]

[1] *Institute of High Energy Physics*

[2] *Institute of High Energy Physics, University of Chinese Academy of Sciences*

Synchrotron light source facilities are evolving into the fourth generation with extreme spatial, temporal and energy resolving capabilities, which pushes the transition of experiment modes into high resolution, multiscale, ultra-fast, and in-situ characterization with dynamic loading or under operando conditions. Such transition raises challenges to balance acquisition efficiency and data quality, where denoising algorithms play a crucial role in signal-to-noise improvement and physical information retrieval. Herein, we develop two efficient denoising algorithms by integrating physical prior knowledge of diffraction/scattering images into deep learning methods.

The first algorithm [1] is based on a small, yet efficient machine learning model designed specifically for SAXS/WAXD experimental image denoising. This model allows for the preservation of physical information and signal-to-noise ratio even when exposure time or dose is significantly reduced, providing a customized solution compared to traditional denoising models designed on natural images. In particular, the proposed model demonstrates superior performance in processing highly textured SAXS/WAXD images compared to the mainstream denoising algorithms. Additionally, the versatility of the proposed model enables wide application in other synchrotron imaging experiments, particularly when data volume and image complexity is concerned.

The second algorithm [2] leverages the intrinsic physical symmetry of X-ray patterns, achieving excellent blind denoising and physical information recovery capabilities without high signal-to-noise ratio reference data. This method is more efficient and effective than the deep learning approaches without physical symmetry considered. It can effectively recover physical information from spatially and temporally resolved data acquired in X-ray diffraction/scattering and Pair Distribution Function experiments, while also maintaining high tolerance on asymmetric distribution of experimental patterns. As a self-supervised denoising approach, the proposed algorithm benefits and facilitates photon-hungry as well as time-resolved in-situ experiments with dynamic loading.

This talk will introduce our most recent discoveries on a systematic denoising solution comprising both supervised and self-supervised denoising methods. The presented material is suitable and contributes to one of the covered topics of NOBUGS 2024: AI/ML applications. We are looking forward to having further discussions with the audience interested in such an intriguing topic.

[1] Zhou, Z., Li, C., Bi, X. et al. A machine learning model for textured X-ray scattering and diffraction image denoising. npj Comput Mater 9, 58 (2023).
[2] Zhou, Z., Li, C., Fan, L. et al. Denoise X-ray Image by Exploring the Power of Its Physical Symmetry. J. Appl. Cryst., Accepted (2024).

Abstract publication:

I agree that the abstract will be published on the web site

Data formats / 70

Crashproofing the HDF5 library

Authors: Dana Robinson[1]; Glenn Song[1]; Matt Larson[1]; Neil Fortner[1]; Vailin Choi[1]

[1] *The HDF Group*

In this presentation, we review several approaches to crashproofing the HDF5 library. We describe an implementation based on a Write-Ahead Log (WAL) for metadata within the HDF5 library. In the event of a crash during the lifetime of an application using HDF5, this WAL can be used by the library or an external tool to restore the metadata within an HDF5 file to a self-consistent state, preventing file corruption and limiting the loss of application data with minimal I/O and memory overhead.

Abstract publication:

I agree that the abstract will be published on the web site

Data formats / 90

Viewing HDF5/NeXus files in the web with H5Web

Authors: Axel Bocciarelli[1]; Loic Huder[1]; Thomas Vincent[1]

[1] *ESRF*

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). Having such a viewer on the web is especially interesting since it allows users to browse files remotely without having to install anything locally.

To answer these needs, we developed H5Web at the ESRF, in open source, with the aim of providing easy access to HDF5/NeXus files and high interactivity with performant WebGL visualisations.

Thanks to its high modularity, H5Web can now be used to view HDF5 files:

- in JupyterLab
- Visual Studio Code
- in the browser via myHDF5
- internally, in the ESRF data portal
- and in many other software at other research institutes and private organisations.

In this presentation, I'll report on the recent progress of the H5Web project as a whole: how we worked to remove the size limit for local files in the browser, to bring the support of compression filters in the browser and how we try to answer general requests on the viewer itself.

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

Abstract publication:

I agree that the abstract will be published on the web site

Beamline Control Systems / 4

Scientific Computing Strategy Developments at the Upgraded Advanced Photon Source

Author: Nicholas Schwarz[1]

[1] *Argonne National Laboratory*

The Advanced Photon Source (APS) at Argonne National Laboratory (ANL) is in the midst of an upgrade project that includes the replacement of the entire storage ring with a ring based on a multi-bend achromat lattice design. This new storage ring will increase the APS's brightness by factors of 500, 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, automation, computing, workflow, data reduction and analysis tools, including AI/ML approaches, and data management to operate effectively.

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 are required to keep up with anticipated data rates and volumes and to provide real-time experiment steering capabilities.

The APS has made great strides developing key elements of its scientific computing strategy. These strides include upgrades to networking infrastructure within the APS and between the APS and the Argonne Leadership Computing Facility (ALCF), deployment of state-of-the-art experiment control software at beamline instruments, expanded capabilities and use of common data management and workflow tools and science portals, utilization of new supercomputers at the ALCF for large on-demand data processing and analysis tasks, development of high-speed, highly parallel data processing and analysis software, and the application of novel mathematical and AI/ML methods to solve challenging data reduction and analysis problems. Additionally, the APS continues to collaborate 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.

Abstract publication:

I agree that the abstract will be published on the web site

Beamline Control Systems / 155

Software Developments for the CERN Fixed Target Experimental Areas

Author: Chris Roderick[1]

Co-authors: Anastasiia Moshenska [1]; Anti Asko [1]; Emanuele Matli [1]; Lukasz Burdzanowksi [1]; Maciej Pert [1]; Stephane Deghaye [1]; Tiago Oliveira [1]

[1] *CERN*

With around 800 devices acting on more than 8km of beam lines, the CERN fixed target experimental areas have a dedicated beam line control system used by a few CERN experts and individuals from more than 140 groups of visiting users to perform experiments. In late 2021, an initiative was triggered to try to unify the accelerator and beam line controls, with the aim of containing maintenance costs and sharing common functionalities. Other software developments beyond pure Controls, have been taking place to facilitate the exploitation of the experimental areas. One such example is the extension of the Accelerator Schedule Management (ASM) software to capture user requests and plan subsequent beam time and associated set-up work. Another example is the use of the Accelerator Fault Tracking (AFT) system to accurately record downtimes of the beam lines and better understand both availability for users and the needs for / impact of consolidation activities. Lastly, data-driven generation of beam line schematics are aiming to help validate layout configurations and eventually evolve into intuitive control synoptics. This paper will give an overview of the developments, including challenges, solutions, and plans.

Abstract publication:

I agree that the abstract will be published on the web site

Beamline Control Systems / 119

BEC –A Beamline and Experiment Control System for the SLS

Author: Klaus Wakonig[1]

Co-authors: Christian Appel ; Alun Ashton [1]; Sven Augustin [1]; Matias Guijarro [2]; Mirko Holler ; Ivan Usov [3]; Jan Wyzula [1]; Xingxing Yao [1]

[1] *Paul Scherrer Institute*

[2] *Paul Scherrer Institute (PSI)*

[3] *PSI*

The Swiss Light Source (SLS) is undergoing a significant hardware upgrade with the SLS 2.0 program, presenting an opportunity to address software challenges, particularly in beamline and experiment control systems. With official endorsement for deployment at the SLS, the new Beamline and Experiment Control system (BEC) provides a unified solution for beamlines, overcoming past challenges of incompatible interfaces and limited collaboration opportunities.
Leveraging community tools like NSLS-II's ophyd hardware abstraction layer, BEC adopts a client-server architecture. The server, composed of multiple smaller services, communicates through a Redis-based message broker and in-memory database. This architecture not only enables a separation of concerns but also accelerates development cycles of internal services. Moreover, it simplifies the integration of new services, data analysis routines, as well as the integration with other systems, owing to shared memory access. This contribution will provide an overview on the architecture of BEC, its components, derived projects, as well as rollout plans for the SLS 2.0.

Abstract publication:

I agree that the abstract will be published on the web site

Open Source Collaborations / 104

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

Author: Michael Schuh[1]

Co-authors: Andrea Parenti [1]; Fabio Dall'Antonia [1]; Luca Gelisio [1]; Oleksii Turkot [1]; Steffen Hauf [1]

[1] *European XFEL*

The VISA platform has been proposed as a common and local portal for data analysis services for the photon and neutron community. It seeks to abstract away site and experiment specific configuration with ready-to-use virtual data analysis environments and it keeps track of interactive sessions.
At the European XFEL, we have gained first experience with using the VISA platform for schools and trainings with internal and external audience. We adopted a Gitlab pipeline to prepare dedicated virtual machine images for each event and use-case, and provided all participants with access to VISA, enabling self-service provisioning of personalized learning environments.
Use-cases leveraged both interfaces that VISA combines, the remote desktop for graphical user interfaces, and Jupyter Lab for notebooks. We provide an environment accompanying a publication on EXtra-Xwiz, a processing pipeline for Serial Crystallography, and we conducted trainings on the Karabo GUI, the main operator interface of the control software at the European XFEL. These, amongst other use-cases, were useful incentives to iterate on our deployment strategy and to test VISA's multi-cloud capabilities.
In this talk, we summarize on lessons learned, and we provide an outlook on future activities.

Abstract publication:

I agree that the abstract will be published on the web site

Posters / 11

Blissterm: a web shell for BLISS

Author: Stuart Fisher[None]

Co-authors: Jens Meyer [1]; Matias Guijarro ; Valentin Valls [1]

[1] *ESRF*

Blissterm is a web shell for the BLISS [1] beamline control framework. In addition to providing the BLISS shell via a web interface, blissterm also adds the possibility of creating simple User Interfaces (UIs) with minimal configuration. The shell can be adorned with for example motors, multi-position, and shutter objects, and there is also the possibility of creating separate monitoring pages. The aim is to provide a coherent User Experience (UX) with the acquisition UI daiquiri [2] so that users can move seamlessly between the two applications. To this extent daiquiri-lib [3] has been created, a shared javascript UI component library published publicly on npm. In addition to providing a web based terminal, blissterm also provides a REST API which allows BLISS to be controlled remotely. This will simplify the interaction with other applications and allow daiquiri to be completely decoupled from the controls system. A client library on top of the REST API, blissclient is also provided. This is a small python library that can be used to remotely control bliss and provide meaningful feedback in case of errors. It also integrates with blissdata to retrieve scan data from redis.

Further information can be found at https://bliss.gitlab-pages.esrf.fr/bliss/master/blissterm.html

[1] Bliss: https://bliss.gitlab-pages.esrf.fr/bliss/master/
[2] Daiquiri: a web-based user interface framework for beamline control and data acquisition, Fisher et al., J. Synchrotron Rad. (2021). 28, 1996-2002
[3] daiquiri-lib: https://www.npmjs.com/package/@esrf/daiquiri-lib

Abstract publication:

I agree that the abstract will be published on the web site

Posters / 94

NexusCreator & ICAT - Helmholtz-Zentrum Berlin applying FAIR data management.

Authors: Hector Perez Ponce[1]; Heike Görzig[1]; Rolf Krahl[1]

[1] *Helmholtz Zentrum Berlin*

The research data management group at Helmholtz-Zentrum Berlin is applying FAIR data management. Data starts to be moved from specific file formats into NeXus/HDF5 files. The standardization program involves the conversion of already generated data, and the automation for the creation of NeXus files from new experiments (example: Bluesky). Our tool, NexusCreator, allows to separate the standarization process in two parts: 1) defining instruments and application definitions via NeXus standard, and 2) creation of file converters or automated generation processes. NexusCreator comes in two flavours, python and javascript.

Abstract publication:

I agree that the abstract will be published on the web site

Posters / 60

ROCK-IT Beamline and Experiment Control

Authors: Andrey Sapronov[1]; Devin Burke[2]; Jonas Gorgis[3]; Linus Pithan[2]; Nicole Wagner[3]; Simone Vadilonga[4]; Udai Singh[2]; William Smith[4]

[1] *IKFT*

[2] *DESY*

[3] *HZDR - Helmholtz-Zentrum Dresden Rossendorf e.V.*

[4] *Helmholtz-Zentrum Berlin*

ROCK-IT is a collaboration which aims to demonstrate the ability to perform complex operando catalysis experiments in a highly automated way, enabling remote operation. The project finds common solutions between different facilities which have various control systems and infrastructure. Ophyd provides a common abstraction layer to Tango, EPICS and SECoP. In the demonstrators of this project, Bluesky is used to orchestrate experiments which involve simultaneous control of a sample environment and various measurement techniques. In this presentation the challenges associated with controlling such an experiment will be presented and the proposed solutions explored.

Abstract publication:

I agree that the abstract will be published on the web site

Posters / 153

The construction of a virtual beamline modeled after a HEPS beamline

Author: Zhibang Shen[1]

[1] *Institute of High Energy Physics, Chinese Academy of Sciences*

High Energy Photon Source (HEPS) is currently under construction. Prior to the installation of all equipment, tasks such as beam tuning can only be carried out based on the experience of the beamline staff and some mathematical estimations, in an imaginary space. While effective, this approach is not convenient. Here, taking the Low-dimension Structure Probe beamline of HEPS as a prototype, we have combined 3D modeling technology and optical ray tracing calculations, based on Unity engine and EPICS, to construct a three-dimensional virtual beamline. The virtual beamline allows users to have a more intuitive perception of the experimental equipment in conditions closer to real scenarios.
The virtual beamline model is linked to optical element modules in the optical x-ray tracing library. The movable parameters of the optical elements are mapped to EPICS process variables (PVs)\. By changing the values of the PVs, the state of the optical elements can be modified, thereby utilizing the x-ray tracing library to simulate the beam.. Other experimental equipment, such as a 6-circle diffractometer, the motor displacements can also be mapped to PVs. By changing the values of the PVs, the movement of the equipments can be observed. In conjunction with Unity3D collision detection functionality, this can be used to anticipate the feasibility of experimental plans.
The virtual beamline provides users, beamline scientists, and beamline engineers with a testing environment where experiments can be conducted freely without concerns about damaging instruments or wasting beam time. In the future, more simulation functions for experimental methods can be added to the virtual beamline, enabling simulation and rehearsal of the experimental process. As an effective tool, the communication efficiency between beamline staff and users, as well as the efficiency of experiments, can be greatly improved. If the mapping between the virtual beamline and the real beamline state can be achieved, a digital twin of the synchrotron radiation beamline will also be constructed.

Abstract publication:

I agree that the abstract will be published on the web site

Posters / 87

The Daisy Workflow Management System for High Energy Photon Source

Authors: Hao-Kai Sun[1]; Yu Hu[2]

Co-authors: Fazhi Qi [3]; Jianli Liu [2]; Lei Wang [4]; Rui Liu [4]; Shiyuan Fu [4]

[1] *Computing Center, Institute of High Energy Physics, Chinese Academy of Sciences*

[2] *IHEP, CAS*

[3] *Institute of High Energy Physics, CAS*

[4] *IHEP*

The High Energy Photon Source (HEPS), currently under construction, represents an advanced experimental platform facilitating breakthroughs in fundamental scientific research. Boasting over fourteen experimental beamlines, HEPS offers a rich array of research domains and employs complex analytical methodologies. Consequently, it faces formidable data processing challenges, including managing high-throughput multi-modal data and accommodating diverse scientific methodologies. Historically, a dearth of comprehensive, user-friendly data processing systems has plagued both China and international research communities. As part of the Daisy project, we endeavor to address these challenges by leveraging the concept of "workflow" to develop a suite of graphical, general-purpose management systems. Our aims include facilitating rapid data processing method dissemination among beam-line scientists, experiment users, and methodology developers, empowering researchers to customize and monitor complex data processing workflows, and enabling batch application of experimental analyses to similar experiments for reproducible results. Through these efforts, we aspire to establish a robust framework for streamlined data processing in the HEPS research ecosystem.

Abstract publication:

I agree that the abstract will be published on the web site

Posters / 101

One year ahead of full deployment at HEPS: the status of Mamba project

Author: Yi Zhang[None]

The advancements on instrumentations of advanced light sources and beamline technologies are making revolutionary changes to the modes of scientific experiments. The Mamba project was launched from 2020, aim to develop a systematic modern software system on top of Bluesky (NSLS II) to address the control and data acquisition task for HEPS experiments. Over the past years, Mamba has formed its own ecosystem. A series of software applications based on Mamba have been deployed on beamlines at Beijing Synchrotron Radiation Facility (BSRF), and the team is now focusing on HEPS applications with the goal to cover most HEPS experiment methods at Day I of user operation. The talk will also cover our vision to integrate latest AI techniques into the full experiment process and how experiment will be carried out in an intelligent synchrotron source.

Abstract publication:

I agree that the abstract will be published on the web site

Posters / 105

Diagnostics of Human Breast Cancer with Tabletop Compact X-ray Setup

Authors: A. Ajeer[None]; A. Alexeev[None]; O. Avdieiev[None]; B. Blinchevsky[None]; S. Denisov[None]; S. Galochkin[None]; R. Keith[None]; V. Kubytskyi[None]; L. Mourokh[None]; G. Sandison[None]; S. Shcherbakov[None]; A. Lazarev[None]; P. Lazarev[None]

Our research demonstrates the feasibility of utilizing compact X-ray setup coupled with advanced computational techniques for accurate and efficient breast cancer diagnostics. We present study of Small-Angle X-ray Scattering (SAXS) in human tissues through Geant4 Monte Carlo simulations. Scattering events from 1 mm thick tissue is recorded with sensitive detector placed at 20 mm (WAXS) and 160 mm (SAXS) from the sample. Materials composed of four components are used to artificially generate tissues which are then used to calculate 2D scattering images. With pyFAI azimutal integration 2D image is transformed to the 1D scattering vector in reciprocal space and appended in dataset. We then use obtained dataset to train a classifier capable of distinguishing between different material compositions (data labels) by providing as input 1d scattering data (features vector).

Abstract publication:

I agree that the abstract will be published on the web site

Posters / 49

ICAT Metadata Ingest using python-icat

Author: Rolf Krahl[1]

[1] *Helmholtz-Zentrum Dresden - Rossendorf (HZDR)*

The ICAT server is a metadata catalogue to support large facility experimental data. python-icat is a Python client library for ICAT. The package provides a collection of modules for writing programs that access an ICAT service. The most important features include the dynamic generation of Python classes to represent the entity object types from the ICAT schema that is automatically adapted to the respective schema version, an extended configuration management, a JPQL query builder, and tools to dump and restore ICAT content to and from a flat file.

Recent versions added a module to read metadata ingest files to add their content to ICAT. This is designed for the use case of ingesting metadata for datasets created during experiments.

This talk will briefly introduce python-icat and discuss the new ingest feature in detail. It will explain how the metadata reader is supposed to be integrated into the ingest pipeline for the data from the experiment.

Abstract publication:

I agree that the abstract will be published on the web site

Posters / 148

Mantid Imaging –A Graphical Interface for Neutron Imaging and Tomography

Authors: Jack Allen[1]; Sam Tygier[1]

Co-authors: Ashley Meigh [1]; Daniel Nixon [1]; Dimitar Tasev [1]; Dolica Akello-Egwel [1]; Martyn Gigg [1]; Michael Sullivan [1]; Rachel Baust [1]; Samuel Jones [1]; Samuel Stock [1]; Will Taylor [1]

[1] *ISIS Neutron and Muon Source*

Mantid Imaging is a user friendly, interactive, open source and free to download GUI application for Linux and Windows. Mantid Imaging is used by the ISIS Neutron and Muon Source Instrument: IMAT, scientists and visiting users for data reduction, reconstruction, and live viewing of 2D and 3D data. The software application is designed to be intuitive such that users of varying technical ability can work with neutron imaging data. While early development of Mantid Imaging has focused on white beam imaging and tomography. To enable energy resolved imaging for Bragg-edge and resonance profile studies, Mantid Imaging has gained ToF (Time of Flight) data loading, a spectrum viewer, and integration with other analysis packages.

Imaging datasets are recorded as 2D radiograms/projections with additional dimensions for angle of rotation and ToF. With the provision of pre- and post-processing operations, Mantid Imaging can use a variety of processes to remove artifacts, improving the quality of data before choosing from a selection of reconstruction techniques. Mantid Imaging includes several GPU accelerated reconstruction algorithms including Filtered Back Projection as well as advanced iterative methods with a choice of regularisation methods.

We report on recent enhancements in Mantid Imaging for energy resolved imaging and improved reconstruction. We present our roadmap for future development to enable new scientific workflows on current and future instruments.

Abstract publication:

I agree that the abstract will be published on the web site

Posters / 69

Not-invented-here: Building a DAQ Platform with off-the-shelf components

Author: Joseph Ware[None]

Diamond has built up a portfolio of internally produced software with analogues in other facilities' products and off-the-shelf solutions in both the commercial and free-open-source spheres. Off-the-shelf software often fulfil most use cases for a fraction of the investment, allowing development resources to be allocated to areas which are unique to the requirements of a user-facing facility, and provide a support lifecycle beyond the contract of an individual.

We present Diamond's work to replace custom-rolled software with off-the-shelf solutions, both replacing existing solutions and adding new functionality. Including:
- Utilising open source libraries for service development and to expose standardised APIs
- Implementing security through standard OIDC Authentication and OPA Authorization flows
- Providing observability & metrics with OpenTelemetry and Prometheus to assist in debugging and identifying improvement requirements

Abstract publication:

I agree that the abstract will be published on the web site

Posters / 103

Novel and flexible data analysis framework combining real-time interaction and remote computing resources, tomography reconstruction software example (STP3)

Author: Aljoša Hafner[1]

Co-authors: Elena Longo [1]; George Kourousias [1]; Giuliana Tromba [1]; Marco De Simone [1]; Marko Kudrna Prašek [1]

[1] *Elettra Sincrotrone Trieste, Basovizza, Trieste, Italy*

The information technology (IT) requirements of complex data analysis have been growing steadily in the last decades. Among the techniques, readily performed at synchrotrons, computational tomography (CT) is one of the most IT resource demanding. This holds both for the computing (CPU and GPU) and storage (I/O) requirements. Taking into account the faster and larger detectors (exceeding 5 megapixels in resolution) and detailed scans (several thousands of projections), the individual datasets readily take several tens of gigabytes of space and thus result in total reconstruction times of tens of minutes in certain cases. Traditionally, tomography workflows have focused most of their attention on speeding up the said reconstruction process. These efforts have resulted in capable and portable open source libraries and packages such as ASTRA and TomoPy. At SYRMEP beamline of Elettra Sincrotrone Trieste, we have recently started to rethink the user interaction with the beamline (specifically the beamline's IT systems and its reconstruction software STP [1, 2]), with a view towards the new Life science beamline, which will be constructed during the ongoing accelerator upgrade (Elettra 2.0 project). Firstly, we have taken into account the actual user's full workflow. We are now considering the full data pipeline, including the interaction of the user with the data itself, i.e. access, loading and saving of the images or when the user is refining the phase retrieval or reconstruction parameters. This not only calls for efficient algorithms, but also for a modern interactive user interface to the remote computing resources, allowing for multiple different CT data processing workflows to be used.
We present here for the first time our work on a novel framework which allows for simultaneous

interactive work and working with HPC-level resources. In brief, the total time of a standard procedure from acquisition to the final full reconstructed volume is reduced from several hours down to tens of minutes. Users are able to interactively refine the parameters using a modern web-based interface which allows for seamless remote work and is able to access the storage in a fast and efficient manner. The actual workflow is not imposed on the user, but a collection of ways to interact with the system exist. For example, a user can choose to refine the reconstruction parameters in-program or obtain them from another source, saving them into the framework's database and then running the jobs in batch mode. Another option is to do each individual full reconstruction directly after obtaining the parameters, bypassing the batch mode and the database entirely. Additional workflows are supported, as the framework's components (backend and frontend) are interoperable and independent of each other.
All this has benefits both on the user, as well as the beamline side. For the user, the whole data analysis process is significantly faster (working on our HPC-grade resources without prior reservation or queue system) and more efficient, where the reconstructed images are obtained already during the beamtime. For the beamline, this opens up opportunities for performing near real-time reconstruction, seamlessly linking the data analysis, acquisition and storage systems.

[1] F. Brun et al., (2015) Fundamenta Informaticae, 141 (2-3), pp. 233-243, DOI: 10.3233/FI-2015-1273
[2] F. Brun et al., (2017) Advanced Structural and Chemical Imaging 3:4, DOI: 10.1186/s40679-016-0036-8

Abstract publication:

I agree that the abstract will be published on the web site

Posters / **82**

Ontology management for the SciCat catalog using LinkML

Authors: Dylan McReynolds[1]; Linus Pithan[2]

Co-authors: Anjali Aggarwal [2]; Paul Millar [2]; Runbo Jiang [3]; Tim Wetzel [2]

[1] *Lawrence Berkeley National Lab*

[2] *DESY*

[3] *LBNL*

The SciCat[1] metadata catalog is in use at several scientific user facilities. SciCat stores metadata about datasets (both raw and derived), proposals, and instruments. When introduced into SciCat, each dataset is given a unique identifier. Datasets can be searched for and browsed in a web portal. Authorization rules can be applied to allow fine-grained access to datasets for staff and users. Datasets can also be made publicly accessible with SciCat, integrating with DOI systems.

When datasets are introduced into SciCat, they contain a "Scientific Metadata" section. SciCat's Scientific Metadata is completely unopinionated, allowing for any fields and values to be extracted from datasets and added to the catalog. This flexibility has enabled adoption by a wide variety of facilities, including X-ray sources, neutron sources, and academic groups. However, this flexibility comes at the expense of standardization, documentation, and machine readability.

The LinkML[2] project provides a set of tools facilitating the definition and publication of "schemas," and allows these to be integrated into a workflow. It enables the definition, maintenance, and interlinking of domain-specific ontologies, and expresses these in a variety of standard definition languages such as JSON Schema, JSON-LD, RDF, and OWL.

Defining schemas for scientific metadata outside of a catalog like SciCat opens the possibility to use higher-level abstractions in the metadata definition than what would be possible, for example, with pure JSON schema definitions. For instance, it is possible to specify hierarchical dependencies and ontological mappings in dedicated data-model frameworks such as LinkML.

DESY and the Advanced Light Source have begun working with LinkML, investigating its use as part of the workflow with SciCat. The aim is to develop a comprehensive solution to enhance the management and validation of scientific metadata within the SciCat framework. The primary objectives include the preparation of a robust data model for experiment metadata, the establishment of a validation layer to ensure the accuracy and integrity of ingested data, the generation of detailed documentation for metadata classes and attributes, the building of schema-based GUIs to insert datasets into SciCat, and the creation of a flexible spreadsheet for efficient metadata list management.

[1] https://scicatproject.github.io/
[2] https://linkml.io/

Abstract publication:

I agree that the abstract will be published on the web site

Posters / **146**

Accelerating Neutron Tomography Ring Artifact Removal Using BM3DORNL

Author: Chen Zhang[1]

Co-authors: Dmitry Ganyushin [1]; Jose Borreguero-Calvo [1]; Pete Peterson [1]

[1] *Oak Ridge National Laboratory*

Neutron tomography is a crucial tool for material examination, but ring artifacts can significantly decrease data quality and complicate tasks like segmentation and morphological analysis. The Block-Matching and 3D filtering (BM3D) algorithm, known for mitigating vertical streaks in sinograms and addressing the root cause of ring artifacts, is unfortunately slow and CPU-intensive. We introduce a unique, open-source software solution that eliminates ring artifacts in neutron tomography using the BM3D algorithm. By leveraging both CPU acceleration through Numba and GPU acceleration through CuPy, our approach significantly improves computational efficiency while maintaining data integrity. This dual-acceleration framework drastically speeds up BM3D processing, allowing researchers to quickly obtain refined results and streamline segmentation and morphological analysis.

Abstract publication:

I agree that the abstract will be published on the web site

Posters / **107**

Improving the Experiment Control System Reliability at MAX IV Laboratory

Author: Vincent Hardion[1]

[1] *MAX IV Laboratory*

The Experiment Control System Reliability project aimed at enhancing the reliability of the beamlines operation at MAX IV. The project focused on the scanning software and the support process, identifying improvements to reduce downtime and increase system reliability. It was structured in innovative organisation to learn from one stable beamline and, in parallel, to improve the reliability

from another beamline. Strategies included developing a recovery strategy, improving configuration management, enhancing diagnostic software, and increasing expert knowledge.
The project resulted in significant reductions in software interventions which will be presented together with the lessons learned on how to target operational efficiency across the MAX IV facility.

Abstract publication:

I agree that the abstract will be published on the web site

Posters / 145

Farewell Mark Koennecke and SICS

Author: Nick Hauser[1]

[1] *ACNS ANSTO*

Mark Koennecke has had a tremendous impact in the NOBUGS community, having significant influence on the development of NeXus and more recently on NICOS for use at ESS.
In 2002, when the software development planning was begun at ACNS (formerly The Bragg Institute), there was a timeline of 2 years to provide experimental controls software for 7 neutron beam instruments. The budget allowed for one developer. Mark was willing to work with ACNS and PSI agreed to provide SICS as shared-source. The project was successful, and SICS is used on 11 instruments at ACNS in 2024.
This communication is to say thank you to Mark for his contribution to the success of ACNS, and to flag that it may be time to look for alternatives to SICS, even though SICS is still functional and fit for purpose. The longevity of SICS (first designed in 1996) is a testament to Mark's design and to the timelessness of C.

Abstract publication:

I agree that the abstract will be published on the web site

Posters / 144

When open source doesn't live up to expectation. Gum Tree 2004 to 2024

Author: Nick Hauser[1]

[1] *ACNS ANSTO*

In 2004, Andy Gotz envisaged Gum Tree as a data acquisition / data reduction workbench that provides scientists the tools to do interactive experimentation, improving productivity by determining data quality at the time of the experiment.
The concept worked and has been used routinely at ACNS on 12 instruments since 2007. The Gum Tree suite has received positive feedback from visiting researchers to the facility. The project was made open source early in its development with a view to having features added and quality improved by collaborative development.
Whilst there have been some fruitful collaborative efforts with Soleil on the Common Data Model Access, Diamond on the January suite and NECSA on the use of Gum Tree on their neutron beam instruments, the Gum Tree suite has not seen the wider adoption we anticipated.
In this communication we describe the issues that held back the adoption of Gum Tree; the effort

required to promote and sell the concept, the effort required to maintain collaborations, the differentiation with 'competing'products such as Diamond's GDA, time zone and geographical issues and the choice of programming language.

Abstract publication:

I agree that the abstract will be published on the web site

Posters / **62**

Automatic Export of Data from Catalysis Experiments to NeXus in ROCK-IT

Authors: Daniel Tomecki[1]; Gerrit Guenther[1]; Hector Perez-Ponce[1]; Sonal Patel[1]; William Smith[1]

[1] *Helmholtz-Zentrum Berlin*

Operando Catalysis experiments involve many different devices and processes running simultaneously and on different timescales. Data from sample environment which is changing over time has to be correlated with data from measurement techniques acquired using X-Ray beamlines. Users performing these experiments commonly face a daunting task of collating all their data together from different sources. In this poster we describe how as part of the ROCK-IT collaboration this problem is addressed. Metadata from devices is collected using HAPPI and Bluesky. It is then exported automatically to NeXus files which can contain information from the sample environment and multiple measurement techniques.

Abstract publication:

I agree that the abstract will be published on the web site

Posters / **142**

A Python Package for Bragg Coherent X-ray Diffraction Imaging Processing, Analysis and Visualisation

Author: Clement Atlan[1]

Co-authors: Corentin Chatelier [2]; Ewen Bellec [2]; Kyle Olson [2]; Marie-Ingrid Richard [2]; Micheal Grimes [2]; Steven Leake [1]; Vincent Favre-Nicolin [1]

[1] *ESRF*

[2] *CEA*

Bragg Coherent Diffraction X-ray Imaging (BCDI) is a non-invasive X-ray characterisation technique for probing in three-dimensions (3D) structures of single nano-objects. While traditional BCDI analysis yields 3D maps of electron density, displacement, and heterogeneous strain, this work introduces a methodology and a user-friendly open-source tool for resolving 3D d-spacing maps within the examined object.

In BCDI, the primary output of phase retrieval algorithms is a 3D map of phase with a spatial resolution of approximately 10 nm. However, analysing BCDI data goes beyond this, and essentially revolves around deriving this phase into more physics-related aspects such as atomic displacement

and strain. To address potential ambiguities and differences in strain definitions across different scientific communities, we introduce the concepts of heterogeneous and homogeneous strains. Heterogeneous strain aligns with traditional BCDI strain analysis, while homogeneous strain corresponds to the shift of the Bragg peak in reciprocal space, a fundamental quantity in X-ray analysis under Bragg conditions. To combine both types of strain information, we present a methodology for computing 3D maps of local d-spacing, an absolute and more fundamental quantity than traditional heterogeneous strain maps. Expanding on this concept allows for the generation of 3D maps of global strain, a quantity analogous to d-spacing.

Cdiutils [1] is a python package designed to facilitate pre- and post-processing stages of BCDI data. It meticulously manages the cropping of the Fourier window prior to phasing, with a particular focus on handling the centre of mass of the 3D Bragg peak intensity relative to the cropped Fourier window. This feature enables the subsequent post-processing of the phase ramp, strain shift, and particularly, d-spacing maps. When coupled with PyNX software [2] for the phasing stage, cdiutils integrates all the steps of the entire data analysis process, including pre-processing, phasing and post-processing, all within a single user-friendly notebook. Multiple methods for heterogeneous strain computation are routinely employed and compared to ensure robust data analysis. Cdiutils runs on ESRF machine environment, transitions between CPUs and GPUs when required, can handle several data formats and beamline geometries and is easily adaptable to any other machine environments and beamline geometries.

[1] Atlan, C., Chatelier, C. & Olson, K. A python package to help Coherent Diffraction Imaging (CDI) practitioners in their analysis. (2023) doi:10.5281/zenodo.8013233.

[2] Favre-Nicolin, V. et al. PyNX: high performance computing toolkit for coherent X-ray imaging based on operators. J Appl Crystallogr 53, 1404–1413 (2020).

Abstract publication:

I agree that the abstract will be published on the web site

Posters / 112

Machine Learning for the Automated Analysis of X-Ray Spectroscopic Measurements: Are We There Yet?

Author: Marius Retegan[None]

Co-authors: Federico Zecchi [1]; Jan Pieter Glatzel [1]

[1] *ESRF*

X-ray absorption spectroscopy provides a wealth of information regarding the local structure and electronic properties of materials. However, data analysis is significantly more time-consuming than acquisition and initial data reduction. Decoding the information relies on comparing it with similar compounds for which the spectrum–property mapping is already established, a task that is very often performed by visual inspection.

Machine learning (ML) is revolutionizing many fields with its ability to extract and learn patterns in big data without having to provide additional prior information other than the data itself. ML models give access to instantaneous predictions of properties and observables, which makes them particularly attractive for performing real-time analysis of the measured data or autonomous experimental acquisitions.

In this talk, I will present the different challenges faced when using machine learning models to analyze X-ray spectroscopic data, from building the initial training datasets to evaluating the ML models' robustness to the different sources of errors, such as spectral shift, normalization, noise level, and class imbalance, that limit the quality of the prediction.

Abstract publication:

I agree that the abstract will be published on the web site

Posters / 84

DonkiWeb: a KISS web SCADA

Author: Roberto Borghes[1]

Co-author: Valentina Chenda

[1] *Elettra Sincrotrone Trieste*

Contemporary web applications are exceptionally efficient and often use a complex structure of frameworks and libraries for front-end development. This escalation in technical complexity requires a team with diverse, specialized skills, inevitably increasing the cost of web development. DonkiWeb is a simple, control system oriented, web SCADA that follows the KISS (Keep It Simple, Stupid) design principle. It integrates a REST interface to Tango, a user authentication layer, a websocket communication protocol and a Javascript API for easy GUI development. Writing a web GUI aims to be as simple as writing a few lines script. This work presents and describes the architecture of DonkiWeb, a fully working web SCADA prototype that is currently used by Elettra Synchrotron beamlines.

Abstract publication:

I agree that the abstract will be published on the web site

Posters / 65

Enhancing User Experience on Screen: ROCK-IT GUI Design for Automated and Remote-Accessed Operando Catalysis Experiments

Author: Zeynep Isil Isik Dursun[1]

[1] *DESY*

Graphical User Interfaces (GUIs) play a crucial role in defining the user-friendliness of software applications, enhancing efficiency by structuring and organizing the information presented. The ROCK-IT project aims to develop all necessary tools for the automation and remote access of synchrotron-based in-situ and operando experiments, using operando catalysis experiments as a pilot case. This poster aims to present the GUI designs developed for ROCK-IT from a visual designer's perspective regarding the fundamental elements of UI/UX design for a more enhanced user experience aimed at a scale of professional and non-professional users.

Abstract publication:

I agree that the abstract will be published on the web site

Posters / 76

Comprehensive Python IOC development with queue_iocs

Authors: Pengcheng Li[None]; Qun Zhang[None]; Yu Liu[None]

EPICS IOCs can be quite inefficient to develop and inflexible to use, because of architectural limitations inherent in EPICS itself. Based on the caproto library, we developed the queue_iocs framework, which we expect to be capable of replacing most EPICS IOCs currently used with more maintainable and flexible Python IOCs. Simple examples include workalikes of StreamDevice and asyn; seq-like examples include monochromators, motor anti-bumping and motor multiplexing. A queue_iocs-based counterpart of areaDetector is also introduced, which overcomes areaDetector's limitations in performance and architecture. Interesting byproducts involved include a simple but expressive architecture for GUIs, workalikes of procServ/procServControl, and extensions to ADGenICam that allow it to support a much wider range of detectors.

Abstract publication:

I agree that the abstract will be published on the web site

Posters / 125

PyStxm: STXM data acquisition using BlueSky at the Canadian Light Source

Author: Russ Berg[1]

Co-authors: Charan Kuppili [2]; Chithra Karunakaran [2]; Jay Dynes [2]; Jian Wang [2]; Yingshen Lu [2]

[1] *Canadian Light Source*

[2] *Canadian Lightsource*

Typically in the past, data collection software developed by individuals with non-software-focused roles often placed low priority on the User Interface (UI), considering it time-consuming and prone to software bugs. However, for software developers primarily tasked with delivering reliable software, the aim is to create a robust platform enabling users of all levels to efficiently collect data, automate repetitive and error-prone tasks, and offer user-friendly features to enhance efficiency at the beamline. A comprehensive user interface is fundamental to achieving these objectives.

Not wanting to reinvent the wheel the user interface incorporates ideas inspired from successful well known existing commercial and open source software, Adobe Photoshop and the open source 3D animation software project, Blender. 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[1] files that conform to the NXstxm[2] NEXUS application definition. pyStxm connects to the underlying motors and other devices via EPICS[3] applications.

The application is written in Python using Qt as the application framework. Device driver support is through EPICS drivers and IOC applications. The CLS spectro microscopy beamline (10ID1) contains 2 STXM end stations in series that get their names from the vacuum environments they operate in, Ambient STXM (ASTXM) for vacuum pressure regions 10-5 and the UHV or Cryo STXM (CSTXM) for vacuum down to 10-8. The CSTXM commissioned in 2019 uses version 2.5 of pyStxm which is dependent on EPICS SSCAN records to perform the function of the scanning engine, the latest version (3) uses the BlueSky [4] data collection framework for the same function. Version 3 of the software was ported to BlueSky as part of an upgrade project for the ASTXM. The bulk of the upgrade for the software was carried out on a testbench of duplicate hardware and using a laser as the source so as to not interfere with existing beamline operations which for the most part achieved the desired

goal. Due to numerous hardware failures, commissioning of the software and end station pushed from late 2023 into 2024 which unfortunately contained a lengthy 6-month outage to replace the LINAC which was originally installed in the 1960's. However, the majority of commissioning work was completed which included the addition of a ptychographic capability for the end station. The largest barrier to completion of commissioning at time of writing is a significant vibration issue preventing image resolutions below 150 nm.

References:
[1] Mark Könnecke, E. et al., (2015). The NeXus data format. Journal of Applied Crystallography. Volume 48, Part 1 (301-305)
[2] Benjamin Watts and Jörg Raabe., (2016), A NeXus/HDF5 based file format for STXM, AIP Conference Proceedings 1696, 020042, https://doi.org/10.1063/1.4937536 Published Online
[3] Dalesio, E. et al.,(1991). EPICS architecture (Conference: International conference on accelerator and large experimental physics control systems, Tsukuba (Japan), 11-15 Nov 1991)
[4] Daniel Allan, E. et al., (2019). Bluesky's Ahead: A Multi-Facility Collaboration for an a la Carte Software Project for Data Acquisition and Management. Synchrotron Radiation News, Volume 32, Issue 3.

Abstract publication:

I agree that the abstract will be published on the web site

Posters / 85

State of the azimuthal integration at the ESRF and pathway to pyFAI2

Author: Jerome Kieffer[1]

Co-authors: Edgar GUTIERREZ FERNANDEZ [1]; Loic Huder [1]; Thomas Vincent [1]

[1] *ESRF*

This contribution summarizes the development of pyFAI over the past year:
* Detectors flipping: orientation can now be specified in PONI
* Geometry is more open for exchanging experimental setup with other software
* Support for grazing incidence geometry
* Azimuthal error models in addition to Poissonian
* Offers unweighted azimuthal integration, like *legacy* version but with correct uncertainties
* Diffraction mapping tool, with exploration of the results

Abstract publication:

I agree that the abstract will be published on the web site

Posters / 71

Karabo and Tango interoperability

Authors: Gabriele Giovanetti[1]; Michael Smith[1]

Co-authors: Alessandro Silenzi [1]; Ana Garcia-Tabares Valdivieso [1]; Andrea Parenti [1]; Ayaz Samadli ; Dennis Goeries [2]; Florian Sohn [1]; Ivars Karpics [1]; Wajid Ehsan [1]

[1] *European XFEL*

[2] *European XFEL GmbH*

The Karabo control system is used facility-wide at European XFEL (EuXFEL) to steer experiments and collect scientific data. As a user-centered facility, EuXFEL deals with ever-changing requirements and faces often the need to integrate new instrumentation, or even to cope with user-provided hardware on relatively short notice.
Tango is a well-established control system, and many hardware devices that are new in the EuXFEL context, have a Tango server available.
For this reasona generic Karabo interface to Tango server has been implemented, the Karabo Tango Mirror. It takes full advantage of the PyTango API and can be configured to connect to a Tango Server and expose any Tango attribute to the Karabo world. A more specific Karabo TangoMotor device has been implemented, to steer motor controllers which come with a Tango server by means of a Karabo motor interface. The Karabo webProxy, wich expose Karabo device properties by mean of a REST API, can be used to expose Karabo devices to any other control system, including Tango.

Abstract publication:

I agree that the abstract will be published on the web site

Posters / 30

Plotting Refactoring for SasView Neutron Scattering Software

Author: Julius Karliczek[1]

Co-authors: Jeffery Krzywon [2]; Lucas Wilkins [3]; Miguel Angel Gonzalez [4]; Paul D. Butler [2]; Piotr Rozyczko [5]

[1] *ILL*

[2] *NIST*

[3] *ISIS*

[4] *Institut Laue-Langevin*

[5] *ESS*

SasView is undergoing a plotting GUI refactoring project. Emphasis lies on a new version of plotting functionality, which is supposed to improve the user experience when using SasView and creating fits and plots with the program. A goal is, to move the triggers of already existing functions to a place, where users can see them more easily. Also, the function of displaying data in a dedicated fitting window with tabs for different datasets is the current scope of the project.

Abstract publication:

I agree that the abstract will be published on the web site

Posters / 128

DrILL builder interface for ILL users in Mantid

Authors: Eric Pellegrini[1]; Remi Perenon[1]

[1] *Institut Laue-Langevin*

These last years, ILL made a committed effort to implement data reduction for a large share of its instruments into the Mantid framework.
By doing so, ILL provides its users with access to the common Graphical User Interface (GUI) of Mantid (Mantid Workbench), bringing the effort on data reduction to a wide audience.

Taking advantage of their proximity with users, Mantid ILL developers could also enhance the GUI by providing new visualization tools.
This required to discuss intensively with scientists leading experiments to gather a maximum of feedback. On the other hand, implementation was designed to be compliant with the constraints coming from the Mantid project.

We will focus on one of these tools, namely the "DrILL Builder".
DrILL Builder intends to help users performing high-throughput data processing by providing easy-and-quick data visualization and automating the sorting/labeling of data files.
We will discuss users specifications and implementation status of this tool.

Abstract publication:

I agree that the abstract will be published on the web site

Posters / 34

Deployment of C++ software, also with Python embedding and extension

Authors: Ammar Nejati[1]; Mikhail Svechnikov[1]; Joachim Wuttke[2]

[1] *FZ Jülich*

[2] *Forschungszentrum Jülich*

We discuss the manifold difficulties in deploying a cross-platform, cross-language software. We consider a software with a C++ core that has an embedded Python interpreter and is exposed to Python with bindings automatically generated by Swig. We explain how such a software can be deployed to Windows, Linux, and macOS, in form of Python wheels or binary installers. Our solutions are based on proven experience with the physics software BornAgain.

Abstract publication:

I agree that the abstract will be published on the web site

Posters / 131

Managing Experiment Configurations in IBEX

Authors: Christopher Moreton-Smith[None]; Daniel Maclaren[None]; David Keymer[None]; Evan Smith[None]; Freddie Akeroyd[1]; Isaac Hill[None]; Jack Harper[None]; Kathryn Baker[None]; Lilith Cole[None]; Lowri Jenkins[None]; Sudeepta Chakraborty[None]; Zsolt Kebel[None]

[1] *STFC ISIS Neutron and Muon Source*

IBEX is the EPICS based experiment control program used on beamlines at the ISIS Neutron and Muon source. A key feature of IBEX is experiment configurations and how these can be split into components that can be re-used across many other configurations. In this context a configuration represents the collection of devices and their settings as well as other options such as data logging

that are required for a given experiment setup, a component is a smaller set of this information and many components can be included in a single configuration. We will describe our configuration architecture and how configurations are created and managed by the instrument scientists within IBEX to allow streamlined setup for experiments.

Abstract publication:

I agree that the abstract will be published on the web site

Posters / 42

Shaping a modern approach to open data from Diamond Light Source

Author: Terence Tan[1]

Co-authors: Philippe Rocca-Serra [2]; Steve Collins [3]; Susanna-Assunta Sansone [2]

[1] *University of Oxford & Diamond Light Source*

[2] *University of Oxford*

[3] *Diamond Light Source*

A collaboration between Diamond Light Source and the University of Oxford funds a PhD project aimed at understanding the technical, social and policy implications of adopting the FAIR (Findable, Accessible, Interoperable, Reusable) Principles, and evaluating the effects of its implementation on synchrotron data. The work focuses on the early stage of the science life cycle, when the scientists submit their experiment proposals to the facility, to ensure FAIRness is embedded in the process at design stage.

To start, we will use Machine Learning techniques to extract metadata from experiment proposals and link the metadata to ontologies; the semantic annotation is essential to make the information machine-actionable, a fundamental goal for FAIR. For the next step, we will investigate how this could trigger a cascading effect and improve the FAIRness of other stages in the science life cycle. Lastly, we will evaluate the impact of the proposed process improvements, looking at cost-benefit for the organisation.

The project is designed to leverage on the activities of existing communities and projects, and is also guided by a Stakeholder Group with representatives from Diamond, other facilities, and European infrastructure projects.

Abstract publication:

I agree that the abstract will be published on the web site

Posters / 56

Hybrid cloud-based instrument control system for remote experiments at MLF, J-PARC

Author: Kentaro Moriyama[1]

Co-author: Hiroyuki Hasemi [2]

[1] *Comprehensive Research Organization for Science and Society (CROSS) Neutron Science and Technology Center*

[2] *Materials and Life Science Division, J-PARC Center, Japan Atomic Energy Agency (JAEA)*

Providing a remote experimental environment for facility users and utilizing advanced computational resources of cloud environments for instrument control are challenging issues for scientific user facilities such as J-PARC MLF. Therefore, we have developed a hybrid cloud-based instrument control system by modifying IROHA2, which is the standard instrument control software framework for neutron scattering experiments in J-PARC MLF. IROHA2 comprises four core server components, namely, device control, instrument management, integrate control, and sequence management, each of which serves as a web interface.
In the newly developed system, the server components for device control and instrument management are deployed in the local computing environment of instruments, while the front-end server components are deployed in the Amazon Web Services (AWS), one of the leading cloud environments, thereby creating a hybrid cloud-based distributed system. Moreover, a loosely coupled distributed system is realized by introducing event-driven asynchronous communication between computing environments via a message broker service in AWS. This prevents communication and system/service failures from spreading to the entire distributed system, improving fault tolerance and availability of the distributed system. Furthermore, Pub/Sub messaging via message brokers is highly compatible with microservice architecture, enabling flexible and easy integration of our system with various services and computing resources in cloud environments (e.g., for automatic measurement and feedback control).
Using the developed system, users can access the front-end of the cloud through a web browser to remotely control instruments, including data collection and control of sample environment equipment.
In this presentation, we discuss the features of our hybrid cloud-based system and further present future development plans.

Abstract publication:

I agree that the abstract will be published on the web site

Posters / **13**

An unified attitude tuning architecture for HEPS beamlines

Authors: Pengcheng Li[1]; Xiaoxue Bi[1]; Yi Zhang[1]; Yu Liu[1]

[1] *Institute of High Energy Physics*

In order to provide higher quality X-ray beams, synchrotron radiation light sources are becoming increasingly complex, including dozens of highly delicate optical devices and sensitive experimental setups. This creates a common problem for all users, namely the need to quickly and accurately adjust the attitude and optimize the beamline to achieve the optimal photon beam and experimental scenarios; at present, this kind of attitude tuning is often done manually, wasting time and manpower. This presentation begins with the introduction of a generic beamline attitude tuning framework for advanced light sources, which supports flexible input/output ports, diverse data processing and evaluation functions, as well as rich choices of numerical optimization libraries. Next, we will describe the integration of our framework in Mamba, the universal experiment control and data acquisition software system for HEPS in China; in this way, user-friendly GUIs can be provided, and ML/AI techiques may be further employed. Finally, we will show real applications of our framework on the attitude tuning of a focusing capillary, an X-ray emission spectrometer and a Raman spectrometer; with similar applications in mind, we expect our attitude tuning framework to be capable of fulfilling a majority of beamline attitude tuning requirements in a unified, efficient and cost-effective manner.

Abstract publication:

I agree that the abstract will be published on the web site

Posters / 28

Motion planning for triple-axis spectrometers

Author: Tobias WEBER[1]

[1] *ILL*

As part of our efforts in the field of autonomous instrument control, we present a motion planning algorithm for triple-axis spectrometers and its open-source implementation. Due to angular constraints and walls in the instrument space, not every (Q, E) coordinate point is accessible to the spectrometer. A careful mapping of the available positions is usually required before each experiment to avoid any collisions. The present algorithm is able to automatically find the best safe path for the instrument, keeping it at the furthest possible distance from obstacles. It does so by calculating the Voronoi bisectors of the instrument's angular configuration space. Of these it creates a map of possible paths and finds the shortest path along the bisectors.

- Paper DOI: 10.1016/j.softx.2023.101455
- Code: https://github.com/ILLGrenoble/taspaths

Abstract publication:

I agree that the abstract will be published on the web site

Posters / 50

A scientific data analysis software framework for exabyte scale data challenges from HEPS

Author: Yu Hu[1]

[1] *IHEP, CAS*

Recent advances in X-ray beamline technologies, including the advent of very high brilliance beamlines at next generation synchrotron sources and advanced detector instrumentation, have led to an exponential increase in the speed of data collection. As a consequence, there is an increasing need for a data analysis platform that can refine and optimize data collection strategies in real time and effectively analyse data in large volumes after the data collection. The increased volume and rate of data push the demand for computing resources to the edge of current workstation capabilities. Advanced data management and analysis methods are required to keep up with the anticipated data rates and volumes.
To address the data challenges at High Energy Photon Source in China, we proposed a software framework and system for the full life cycle of the advanced light source experiment. In this talk, we will focus on the data analysis software framework software framework in this system. We will introduce the design of this framework and the scientific software developed based on the framework. The future plan will also be presented.

Abstract publication:

I agree that the abstract will be published on the web site

Posters / 32

MXCuBE all in one - user interface, control, collaboration

Authors: Antonia Beteva[1]; Marcus Oskarsson[1]; Daniele De Sanctis[1]

Co-author: Jens Meyer [1]

[1] *ESRF*

he MXCuBE (Macromolecular Xtallography Customised Beamline Environment) software, originally developed at ESRF, provides a common user interface for the Crystallography experiments. The software offers an abstraction to control systems and hardware so that the application can run within most beamline environments. Today, it has evolved into a very successful collaboration between 12 Synchrotrons on 4 continents, celebrating its 15th anniversary. The project's aim is to render the same user experience, no matter where the experiment is taking part, and be a platform for sharing domain-specific knowledge, best practices as well as data acquisition protocols.

MXCuBE is used at the core of all the ESRF MX beamlines, enabling remote, high-throughput attended and unattended data collections.

The project is divided into two parts - mxcubecore and UI, both open source, hosted on GitHub and distributed under the terms of the GNU Lesser General Public License. The software benefits from a well-established technology stack. Within the MXCuBE developer community, there is a well-defined process for versioning, deployment and contributions.

mxcubecore is a library that provides software abstraction to different control systems and beamline instrumentation. It is written in Python. There are a number of already supported control systems, protocols and sequencers: Bliss, Tango, Sardana, Epics, Ophyd, TINE, SPEC, Exporter. It is also relatively easy to add new ones. The hardware can be controlled via something called channels and commands (e.g. Tango, Epics) or with direct access to a control system (e.g. bliss).

The heart of mxcubecore is a set of abstract classes. The role of these classes is to provide an API that can be used by any client. It defines methods, attributes and signals. Most of the methods are defined only to give a blueprint and should, when appropriate, be overloaded by the inheriting classes. This makes mxcubecore easy to step in for newcomers and cheap in time for development.

There are two distinct MXCuBE user interfaces, based on different technologies - mxcubeweb and mxcubeqt. The underlying beamline control layer is the same for both and is implemented using the mxcubecore library.

mxcubeweb is a web application, running in any recent browser. The application is built using standard web technologies, based on Flask, React and web sockets and does not require any third-party plug-ins to be installed in order to function.

mxcubeqt is a Qt based application. It is written in Python, using the PyQt bindings for Qt. mxcubeqt was the first implementation of the MXCuBE application. Being a desktop application, mxcubeqt needs a separate solution to run remote experiments.

Abstract publication:

I agree that the abstract will be published on the web site

Posters / **29**

A new version of the TAS software Takin

Author: Tobias WEBER[1]

[1] *ILL*

We present a new version of the triple-axis software suite *Takin* whose most recent addition includes a magnon calculator based on the same formalism as SpinW. The new module is written in fast C++20 and features a graphical user interface for quickly manipulating magnetic structures and couplings, but can also be used as a library and scripted using *Python*. The module is furthermore usable as a stand-alone program as well as a plug-in into *Takin's* resolution convolution simulator and fitter where it allows for the resolution-corrected refinement of magnetic coupling constants.

- Paper DOI: 10.1016/j.softx.2023.101471
- Code: https://github.com/ILLGrenoble/takin

Abstract publication:

I agree that the abstract will be published on the web site

Posters / **41**

METABOLATOR: Establishing a Citable Web Application for Automated Metabolic Load Analysis

Authors: David Pape[1]; Mani Lokamani[1]; Ayush Seal[1]; Oliver Knodel[1]; Jeffrey Kelling[1]; Karim Fahmy[1]; Guido Juckeland[1]

[1] *Helmholtz-Zentrum Dresden - Rossendorf (HZDR)*

METABOLATOR is a web application for automated analysis of microcalorimetric metabolic data using Monod's equation. The software was developed in collaboration between the Institute of Resource Ecology and the Department of Information Services and Computing at Helmholtz-Zentrum Dresden - Rossendorf (HZDR), and is now offered as a web service for the community. In addition to publishing the software under an open source license, we made the service, which is hosted on HZDR infrastructure, citable by registering its metadata with DataCite and minting a dedicated Digital Object Identifier (DOI). In this talk, we will present the results of our collaboration from the point of view of a Research Software Engineer (RSE). We will introduce the METABOLATOR software, and discuss its development from initial trials into an installable package and web service. Moreover, we will debate the importance of persistent identifiers (PIDs) for reproducible, citable, and overall FAIR data analysis workflows.

Abstract publication:

I agree that the abstract will be published on the web site

Posters / **48**

Back in the future with the Recovery Portal: a tool to restore control system components at European XFEL

Authors: Alessandro Silenzi[1]; Andrea Parenti[2]; Dennis Goeries[1]; Gero Flucke[1]; Ivars Karpics[2]; Michael Smith[2]; Raul Costa[1]; Steffen Hauf[2]; Valerii Bondar[2]; Wajid Ehsan[2]

[1] *European XFEL GmbH*

[2] *European XFEL*

Stable and continuous operation of large-scale distributed control systems are based on well- established configuration management and data logging. During service interruptions, like power cuts, hardware failures, network outages, planned maintenance and software deployment, parts of control system components may lose crucial configuration and restoring them to working condition may take time, resulting in missing service hours. To avoid fixing system components one by one an overall recovery procedure and tools might be used.In this contribution a software solution for restoring control system components to an arbitray point in the past at European XFEL is presented.

Abstract publication:

I agree that the abstract will be published on the web site

Posters / **55**

Raydata: An application to facilitate analysis at X-ray light source

Authors: Andi Barbour[1]; Evan Carlin[2]; Hiran Wijesinghe[1]; Lutz Wiegart[1]; Max Rakitin[1]; Nathan Cook[2]; Paul Moeller[2]; Raven O'Rourke[2]

[1] *Brookhaven National Laboratory*

[2] *RadiaSoft LLC*

The execution and analysis of light source experiments requires the use of sophisticated simulation, controls and data management tools. Existing workflows require significant specialization to accommodate specific beamline operations and data pre-processing steps necessary for more intensive analysis. We present a prototype analysis platform with real-time analysis at the beamline. A part of the Sirepo scientific gateway, Raydata provides a graphical user interface for controlling analysis of beamline scans and leverages a flexible run engine to execute user configurable Python-based analyses with customizable queuing and resource management. The system currently supports CSX and CHX beamlines at the National Synchrotron Light Source II. We discuss the system's architecture and user interface.

Abstract publication:

I agree that the abstract will be published on the web site

Posters / **77**

Motor device interfaces and a multi-axis motor framework at European XFEL

Author: David Hickin[1]

Co-authors: Alessandro Silenzi [1]; Ana Garcia-Tabares Valdivieso [2]; Andrea Parenti [1]; Anna Klimovskaia [1]; Ayaz Samadli [1]; Bernard Baranasic [1]; Dennis Göries [1]; Florian Sohn [1]; Gabriele Giovanetti [1]; Gero Flucke [1]; Ivars Karpics [1]; Leandro Zanellato [1]; Michael Smith [1]; Nerea Jardon [1]; Patrick Geßler [1]; Raul Costa [1]; Riccardo Fabbri [1]; Sergey Esenov [1]; Steffen Hauf [1]; Sylvia Huynh [1]; Tobias Freyermuth [1]; Valerii Bondar [1]; Wajid Ehsan [1]

[1] *European XFEL GmbH*

[2] *European XFEL*

Karabo is a device-based distributed control system used to implement the control and data acquisition systems of the 3 tunnels and 7 instruments of European XFEL.

Motion systems, both PLC and non-PLC-based, are controlled through single and multi-axis Karabo devices with standardised interfaces and behaviour.

Combined motion of multiple motors is provided through multi-axis devices, which are presented as a number of virtual axes. A device framework using Karabo's middlelayer API facilitates the creation of multi-axis devices with minimal code, ensuring the expected interface and behaviour, and includes motion coordinated at the PLC level, hardware and software limits, automatic operator panel generation and a unit testing framework.

Abstract publication:

I agree that the abstract will be published on the web site

Posters / 133

Remote Collaboration via Distributed HDF File Access

Authors: Justin Wozniak[1]; Raymond Osborn[1]

[1] *Argonne National Laboratory*

Cross-institutional data sharing is still a challenging problem for the large datasets collected at the Advanced Photon Source (APS). Sector 6 at the APS routinely collects single-crystal x-ray diffraction data at a rate of several terabyes per day, which is streamed for automated data reduction in local file stores. Such large data volumes make it challenging to collaborate on data analysis with remote collaborators, without the inefficiencies of transferring full data sets. There is a need to be able to collaborate with remote users performing visualization, lightweight analysis, and small data modification on these datasets, tasks for which the full dataset is not needed. Thus, we desire to create a system in which it is possible to enable remote data slicing and selection. The HDF interface allows for the possibility of lightweight data access, at least at the interface level. We are prototyping this approach using the HDF Highly Scalable Data Service (HSDS), a distributed service that can be hosted on an institutional cluster or commercial cloud and accessed by a remote client. This proposed solution integrates with existing analysis routines and visualization tools such as NeXpy. In this presentation, we will provide more detail on the use cases for remote collaboration, outline the architecture in more detail, and provide preliminary measurements on the costs of the approach, both from a performance and financial perspective.

Abstract publication:

I agree that the abstract will be published on the web site

Posters / 79

GeCo: The Elettra 2.0 Beamline Interlock System

Author: Valentina Chenda[1]

Co-authors: Abrami Alessandro [1]; Adriano Contillo [1]; Luca Cristaldi [1]; Luca Rumiz [1]; Luca Sancin [1]; Matteo Lucian [1]; Michele TURCINOVICH [1]; Milan Prica [1]; Peter Sossi [1]; Roberto Borghes [1]; Roberto Pugliese [1]

[1] *Elettra Sincrotrone Trieste*

The Elettra Synchrotron, located in Italy near Trieste, has been operating for users since 1994 being the first third generation light source for soft X-rays in Europe. To stay competitive for world-class photon science, a massive upgrade of the storage ring has been planned in 2025. The goal is to build an ultra low emittance light source with ultra-high brilliance in the same building as the present storage ring. The Elettra 2.0 project includes the construction of new beamlines and an extensive upgrade of most of the existing beamlines to fully exploit the high brightness and high degree of coherence offered by the new source. In this scenario the original beamline interlock and personnel safety systems are going to be upgraded using state of the art technologies. Siemens PLCs are used for low level control, while higher level applications are developed using the Tango framework. Particular attention was paid to the design of the hosting crates in order to have a modular, compact and easy to assemble solution. Also the software architecture has been designed for having a standard reusable schema. PLC data structures and code are automatically created by means of Python scripts, minimizing development time and human errors. This work presents and describes the architecture of the next Elettra 2.0 Beamline Interlock System named GeCo (Gestione e Controllo in italian).

Abstract publication:

I agree that the abstract will be published on the web site

Posters / **80**

FaXToR data processing

Author: Gabriel Jover-Manas[1]

[1] *ALBA synchrotron*

On-the-fly 3D data reconstruction is a challenging need in synchrotron micro-tomography facilities. This presentation will have two parts. In the first part we will show the data workflow and infrastructure implemented at the FaXToR beamline of the ALBA Spanish synchrotron to follow dynamics inside the samples. In the second, we will present a new approach to implement tomography processing tools in the Orange framework. This user-friendly yet powerful set of tools allow users to process and visualize tomographic datasets in our servers and is ready to be expanded with new libraries and tools.
The beamline is expected to present a high data throughput making use of state-of-the art CMOS fast detectors. Therefore, particular care is required in order to cope with the computing requirements. The IT infrastructure will support 3D data reconstruction, distributed processing and PB data storage.

Abstract publication:

I agree that the abstract will be published on the web site

Posters / **91**

The Practice of CI/CD in Advancing the Ecosystem Development of photon Source Software

Authors: Jianli Liu[1]; Hao-Kai Sun[1]; Lei Wang[2]; Shiyuan Fu[2]; Yu Hu[3]

[1] *Computing Center, Institute of High Energy Physics, Chinese Academy of Sciences*

[2] *IHEP*

[3] *IHEP, CAS*

Continuous Integration/Continuous Delivery (CI/CD) can facilitate the development and integration process of advanced photon source software and algorithms. A significant amount of repetitive

tasks, such as compiling, testing, deploying, and releasing, may impede the progress of algorithm and software development. Developers often need to expend considerable effort maintaining servers and development environments to ensure the smooth operation of the overall development process. In contrast to professional software developers and maintainers, scientists are more concerned with improving algorithm implementation and integrating scientific software, rather than focusing on the underlying software environment deployment.
This report discusses how to advance the ecosystem of advanced light source software and the latest developments using CI/CD systems, including implemented instances of software and algorithm integration, technical challenges in the ecosystem development of advanced light source software, and future work plans.

Abstract publication:

I agree that the abstract will be published on the web site

Posters / 157

SNAPRed: A tool for data reduction and instrument calibration for the SNAP high-pressure diffractometer.

Author: Malcolm Guthrie[1]

Co-authors: Andrei Savici [1]; Michael Walsh [1]; Pete Peterson [1]

[1] *ORNL*

The SNAP diffractometer at SNS is highly re-configurable, featuring movable detectors, choppers, and optional neutron optics, introducing challenges in automated data reduction. The complexities of this data orchestration necessitated the development of the ongoing SNAPRed software. SNAPRed addresses this by defining unique Instrument States based on the instrumental configuration. A streamlined workflow for rapid, standardized calibration, accompanied by metrics for calibration quality, is a key feature. Pressure cells ("containers") impact background and attenuation with wavelength-dependent effects. SNAPRed integrates existing approaches to manage these challenges. User experience (UX) is prioritized, balancing accessibility for novices with advanced controls for experts. Performance is optimized to support real-time Rietveld analysis by the end user, supported by a "lite"mode for faster data processing without loss of diffraction resolution. The software standardizes the reduction workflow, capturing all parameters, metadata, and processes for complete provenance of the reduced data.

Abstract publication:

I agree that the abstract will be published on the web site

Posters / 10

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

Author: Stuart Fisher[None]

Co-authors: Jens Meyer [1]; Marcus Oskarsson [1]; Valentin Valls [1]

[1] *ESRF*

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 can launch data processing and interact with the ESRF workflow system ewoks using ewoksweb and sidecar (a processing bridge). Daiquiri also provides real-time monitoring and short term acquisition review via mimosa, a separate dedicated "off-line"web interface.

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 three beamlines in 2024.

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

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

Abstract publication:

I agree that the abstract will be published on the web site

Posters / 9

Unsupervised clustering for extracting fine structural information in ARPES

Author: Lingzhu Bian[None]

ARPES(ARPES : angle-resolved photoemission spectroscopy) is a powerful tool for observing electronic structures in solid-state materials, widely used in characterizing quantum materials. Spatially resolved ARPES (Nano-ARPES) allows measurements on relatively inhomogeneous surfaces, due to its sub-micrometer beam size. However, challenges remain due to the complexity of the surface, particularly in positioning the area of interest which only relies on the scanning photoemission microscopy (SPEM). However, the SPEM data is quite heavy, which needs to be well analyzed in order to distinguish the desired regions. This is important but time consuming.

In recent years, unsupervised clustering method has shown strong capabilities in automatically categorizing the ARPES spatial mapping dataset[1,2]. However, it is only for real space currently and usually has limited ability in distinguishing subtle differences caused by complex and variable scenarios, such as different layers and substrates. Here, we propose a novel method called High-order Unsupervised Clustering Approach (HUCA). Using the K-means clustering results/metrics for real space in different energy-momentum windows as the input of the second round K-means clustering for momentum space, the energy-momentum windows that exhibit subtle inhomogeneity in real space will be highlighted. It recognizes different types of electronic structures both in real space and momentum space in spatially resolved ARPES dataset. Moreover, some subtle band differences, such as band shift or splitting, can still be pointed out by HUCA. Our results demonstrate that the clustering accuracy and identification limit can be significantly improved by HUCA. This method can be used to quickly capture the areas of interest, and is especially suitable for unknown samples, extremely small areas and areas with dispersed distribution.

Furthermore, by combining HUCA with the ARPES data acquisition system, it will achieve online fine clustering and band structure extraction, opening an era of intelligent ARPES experimental data collection and promoting the widespread application of machine learning in high-dimensional data clustering processing.

1 H.~Iwasawa et,al, "Unsupervised clustering for identifying spatial inhomogeneity on local electronic structures",npj Quantum Mater. 7, 24 (2022)
2 C.~N et al,"K-means-driven Gaussian Process data collection for angle-resolved photoemission spectroscopy", Mach. Learn.: Sci. Technol. 1 045015 (2020)

Abstract publication:

I agree that the abstract will be published on the web site

Posters / 37

Enhancing Operational Efficiency at SPring-8: Automated Operation Mode Scheduling and Proposal task management for Measurement Proxy

Author: Takahiro Matsumoto[1]

Co-author: Nobuhiro Mizuno [1]

[1] *Japan Synchrotron Radiation Research Institute*

At large synchrotron radiation facility SPring-8, accommodating diverse experimental user demands requires various operation modes and advanced task management for measurement proxies. Despite these complexities, the need for efficient operation through automation has become increasingly essential. This presentation discusses initiatives to enhance operational efficiency at SPring-8, which operates under multiple bunch modes including three several-bunch modes and five hybrid bunch modes. 1 Traditionally constrained by user demands and operational limitations, manual scheduling and task management were performed semi-annually. To address these challenges, we introduced an automated system based on mathematical optimization and tested since 2023. This system dynamically generates operational schedules based on user-input constraints, significantly reducing administrative burdens and improving flexibility, cost-efficiency, and fairness. Additionally, we developed a web-based management system for the Structural Biology Beamlines that automates tasks such as guidance dispatch, sample verification, and data handling. This has notably reduced staff workload and improved operational continuity. These innovations have enhanced service quality and user satisfaction, while also improved the operational management at this complex research facility.

Reference
1 http://www.spring8.or.jp/en/users/operation_status/schedule/

Abstract publication:

I agree that the abstract will be published on the web site

Posters / 53

Pushing the speed limit of hardware triggered scans using the PandABox

Author: Oliver Silvester[1]

Co-authors: David Perl [1]; Dominic Oram [1]

[1] *Diamond Light Source*

In Macromolecular Cystallography (MX) experiments at Diamond Light Source (DLS), crystals are located by scanning samples around a rectangular grid, under synchrotron light, and looking for diffraction. The speeds of these scans have been limited by the position capture unit which coordinates trigger signals - the Zebra. Using a Zebra, these scans have historically had a step-like motion, where the goniometer stops at each trigger point before accelerating to the next one. While the detector and goniometer could theoretically support grid scan trigger rates of 500hz, trigger points became too inaccurate to use beyond a 250hz scan. The PandABox, with its larger FPGA and faster data transfer, can handle a constant speed scan motion, where the goniometer continuously travels across each row while triggers are sent. With this new motion, the grid scans can be performed at maximum goniometer speeds and minimum detector exposure times –doubling the frame rate of grid scans to 500hz, thus increasing the overall throughput of the beamline. This talk will discuss how the PandA has been exploited to design these new grid scans and its integration with the data acquisition software

Abstract publication:

I agree that the abstract will be published on the web site

Posters / 143

Online Multimodal Data Analysis At The Hard X-ray Micro/Nano Probe P06.

Author: Jan Garrevoet[1]

[1] *DESY*

Abstract. Many X-ray based techniques allow for simultaneous acquisition, e.g. XRF, XRD, SAXS, WAXS, XBIC, XBIV, and
many more. Although previously hardly every utilised, it is nowadays becoming the standard. Now that many technical challenges
are solved the next challenge lays in the analysis of the acquired data.
INTRODUCTION
Adding the possibility of multimodal data acquisition is a must for any beamline. The current challenges are not only
the amount of data and computational resources needed, but also the know-how. This because many users are not
specialists any more in the x-ray field but in another research field. To make the best of the data, data analysis often
is/was outsourced to beamline personnel.
ARCHITECTURE
The main architecture is based on micro services and distributed on as many compute nodes as required for the task at
hand. All acquired data is accompanied by metadata that allows for fully automated data analysis for every available
measurement modality. The default data pipeline constitutes out of 3 components, a master, worker, and data sorter.
A task distributor and a sink complete the default standard blocks.
AIM
The goal of the data pipelines is to provide processed data asap to the visiting scientist and control system. The delay
between data acquisition and analysed data is dependant on several components. Software stack of the used detector,
streaming or file based, and the methodology and the required computational power. The analysed data is provided to
the user in 3 forms:

-
-
-

Queryable database;

HDF5 file;
Visualised in an online data viewer.
ACKNOWLEDGMENTS
We acknowledge DESY (Hamburg, Germany), a member of the Helmholtz Association HGF, for the provision of
experimental facilities. Parts of this research were carried out at the PETRA III P06 Beamline.

Abstract publication:

I agree that the abstract will be published on the web site

Posters / **141**

SHIVER - A graphical user interface for visualization of single crystal inelastic neutron experiments

Author: Andrei Savici[1]

Co-authors: Maria Patrou [2]; Ross Whitfield [2]; Chen Zhang [2]

[1] *Oak Ridge National Laboratory*
[2] *ORNL*

Historically, data workflows for single crystal inelastic neutron experiments were using a pre-histogramming step, storing data for each detector as a histogram in energy transfer. Therefore, each bin of these histograms is then a measurement of the dynamic structure factor (or double differential cross section). A useful representation of this data includes transformation into momentum transfer, and a final histogramming step in the desired momentum and energy transfer grid. The main deficiencies of such existing approach are (i) the large amount of memory resources needed to store the histogram bins with zero neutron counts, and (ii) the difficulty of error propagation in re-histogramming, which resulted in the absence of weighting of contributions with different statistical significance.

We have developed a different approach, where the histogramming to energy-momentum space is based on neutron detection events and accounts for their statistical significance. The algorithm is implemented in the Mantid software.

The script based interface was made available a few years ago. This presentation will focus on showing the newly implemented graphical interface, and will emphasize lessons learn for improvement of the software code and the user experience.

This research used resources at the Spallation Neutron Source, a DOE Office of Science User Facility operated by the Oak Ridge National Laboratory.

Abstract publication:

I agree that the abstract will be published on the web site

Posters / **114**

Deployment strategy of Beamline and Experiment Control components across development and production environments

Author: Ivan Usov[1]

Co-authors: Borys Sharapov ; Klaus Wakonig ; Leonardo Sala ; Simon Ebner

[1] *PSI*

The Beamline and Experiment Control (BEC) is a new solution for beamline operations that targets the Swiss Light Source upgrade (SLS2.0) at Paul Scherrer Institute.

We present a deployment strategy for BEC components and dependencies, leveraging on-premise GitLab pipelines, runners, and Ansible roles/playbooks. GitLab pipelines orchestrate automated workflows, integrating version control with continuous integration/continuous deployment (CI/CD) practices. GitLab runners serve as execution agents, executing pipeline jobs on virtual machines within the development and production environments, and enabling scalable and efficient deployment processes. Finally, Ansible playbooks and roles streamline configuration management and deployment tasks.

By combing GitLab's CI/CD capabilities with Ansible automation, we expect BEC to achieve a scalable deployment mechanism across all beamlines, facilitating adaptation to evolving requirements and ensuring optimal user configuration interface within the SLS 2.0 ecosystem.

Abstract publication:

I agree that the abstract will be published on the web site

Posters / **97**

Control of Sample Environment via Secop With Bluesky, While Performing Measurement Procedures on the Beamline

Author: Peter Wegmann[1]

Co-authors: Dirk Wallacher [1]; Klaus Kiefer [1]; Marcel Bajdel [1]; Nico Grimm [1]; Simone Vadilonga [1]; William Smith [1]

[1] *Helmholtz-Zentrum Berlin*

The easy integration of sample environment hardware into Bluesky represents a crucial step for achieving the goals for automation of in-situ and operando experiments at photon sources required by the ROCK-IT project. The use of Ophyd as a common hardware abstraction layer facilitates the integration of sample environment hardware into an existing Bluesky environment. Based on the metadata and structural data provided by the Sample Environment Communication Protocol (SECoP), the SECoP-Ophyd Integration, developed by us, is capable of automatically generating and configuring Ophyd devices upon connection to a Sample Environment Control (SEC) node. A successful integration test for a catalysis experiment at BESSY II demonstrated the control of a SECoP enabled sample environment and performing measurement techniques on the beamline within the same Bluesky environment. This poster presents an introduction to the SECoP-Ophyd integration, as well as an overview of the control of the sample environment at the experiment.

Abstract publication:

I agree that the abstract will be published on the web site

Posters / **78**

A Tango control system for the MicroTomo2@STAR imaging station

Author: Michele Belletti[1]

Co-authors: Diego Dreossi [1]; Iztok Gregori [1]; Katia CASARIN [1]; Marco Peloi [1]; Massimo DEL BIANCO [1]; Roberto Borghes [1]; Salvatorangelo SBARRA [1]; Valentina Chenda [1]

[1] *Elettra Sincrotrone Trieste*

MicroTomo2 is an X-ray imaging station installed at the STAR accelerator facility of the University of Calabria. The STAR accelerator, currently under construction, will produce photons with energies up to 350 keV generated by the laser light-electron collisions using the Thomson back-scattering phenomenon. The MicroTomo2 experimental station will provide full-field X-ray radiography and tomography for 3D imaging and analysis of a large variety of samples and materials. The experimental setup is composed of a six axes sample manipulator and two distinct X-ray imaging detectors that will allow to cover a wide range of scientific cases spanning from high resolution investigations to large scale sample tomography. The focus of this work is the MicroTomo2 control system, developed using the TANGO framework, that is in charge of managing all the elements of the station from the photon transport to the experiment management and data visualization.

Abstract publication:

I agree that the abstract will be published on the web site

Posters / **127**

Documentation is communication. Tips to make documentation more successful.

Author: Olga Merkulova[1]

[1] *DESY*

Breaking news, Instagram Reels, TikTok, and YouTube Shorts,…Living in the era of information intoxication means the average person's attention span was ~8 seconds in 2022. No wonder it's hard to conquer readers'attention with documentation. That's why it's important to make documentation exciting and interesting to read still keeping it informative.

Documentation is communication. And it is an important part of a community builder. It is also a visit card of your project. Via documentation you bring a message to your clients or community. It is a way of communication. And without communication there is no sustainability, consistency, etc. How to communicate to succeed with your community?

Documentation within a group or a facility may also be a single source of truth.
Documenting processes and protocols may be even more important than documenting project itself, especially if it is some kind of a library.

Another big issue in creating documentation is that your team probably doesn't know how to do it. If you make the process understandable and a standard one, there will be no problem.

I will tell what types of documentation exist, share some tips and tools to simplify the process of writing documentation.

Abstract publication:

I agree that the abstract will be published on the web site

Posters / **8**

An efficient ptychography reconstruction strategy through fine-tuning of large pre-trained deep learning model

Authors: Chenglong Zhang[1]; Chun Li[1]; Liang Zhou[1]; Shuo Wang[1]; Xinyu Pan[2]; Peng Liu[1]; Wenhui Wang[1]; Yi Zhang[1]; Yuhui Dong[1]; Zhongzheng Zhou[1]

[1] *IHEP*

[2] *The Institute of High Energy Physics of the Chinese Academy of Sciences*

With pre-trained large models and their associated fine-tuning paradigms being constantly applied in deep learning, the performance of large models achieves a dramatic boost, mostly owing to the improvements on both data quantity and quality. Next-generation synchrotron light sources offer ultra-bright and highly coherent X-rays, which are becoming one of the largest data sources for scientific experiments. As one of the most data-intensive scanning-based imaging methodologies, ptychography produces an immense amount of data, making the adoption of large deep learning models possible. Here, we introduce and refine the architecture of a neural network model to improve the reconstruction performance, through fine-tuning large pre-trained model using a variety of datasets. The pre-trained model exhibits remarkable generalization capability, while the fine-tuning strategy enhances the reconstruction quality. We anticipate this work will contribute to the advancement of deep learning methods in ptychography, as well as in broader coherent diffraction imaging methodologies in future.

Abstract publication:

I agree that the abstract will be published on the web site

Posters / **74**

Web-Based control system for the QUATI beamline at Sirius

Author: Igor Ferreira Torquato[1]

Co-authors: Santiago Figueroa [1]; Alexey Espíndola [1]; Eduardo Pereira [1]; Amélie Rochet [1]

[1] *Brazilian Synchrotron Light Laboratory (LNLS), Brazilian Center for Research in Energy and Materials (CNPEM)*

High-level control systems and Graphical User Interfaces (GUIs) play essential roles in enabling users to interact with complex systems, particularly in beamline environments where precise control and real-time monitoring are crucial. Beamlines must provide tools that facilitate this interaction. On the Quati beamline[1], which is the XAS beamline of Sirius, the experiment control system is designed to offer a simple yet robust web-based interface to scripts and scan engines. In contrast to specialized frameworks for beamlines, this system prioritizes versatility, providing users with intuitive control and configuration capabilities of available scripts. The system architecture is composed of a centralized control server designed to serve multiple clients. This server runs a robust message queue system (RabbitMQ[2]), that ensures decoupled applications and scalability. This architecture enables the use of asynchronous communication with the client, client-server decoupling, and non-simultaneous availability, offering greater flexibility and fault tolerance. The system's flexibility extends to its ability to run multiple scripts across different servers, facilitated by its design. The web server was developed using Flask[3] and incorporates SocketIO for real-time updates, offering users a responsive and dynamic interface. The system is under final development and will be implemented for the beam commissioning of Quati, in 2024.

References:

1. Figueroa, Santiago JA, et al. "QUATI beamline: QUick x-ray Absorption spectroscopy for TIme and space-resolved experiments at the Brazilian Synchrotron Light Laboratory." Radiation Physics and Chemistry 212 (2023): 111198.

2. RabbitMQ Documentation | RabbitMQ. (n.d.). https://www.rabbitmq.com/docs

3. Flask Documentation. https://flask.palletsprojects.com/en/3.0.x/

Abstract publication:

I agree that the abstract will be published on the web site

Posters / 6

A WEB APPLICATION FOR BIOSAXS HIGH-THROUGHPUT DATA COLLECTION AND EXPERIMENTAL CONTROL

Author: Jean Baptiste FLorial[1]

Co-authors: Petra PERNOT [2]; Mark Tully [2]; Marcus Oskarsson [2]; Andrew MC CARTHY [1]

[1] *EMBL*

[2] *ESRF*

The ESRF extremely brilliant source (EBS), Europe's first high energy 4th generation synchrotron, started user operation in August 2020. To benefit from the exceptional high quality X-rays produced by the ESRF-EBS all the experimental control and data analysis pipelines have been significantly upgraded on all the EMBL-ESRF Joint Structural Biology Group (JSBG) beamlines at the ESRF. On the biological X-ray scattering beamline, BM29, a new web-based user interface, BioSAXS Customized Beamline Environment (BSXCuBE), was developed to simplify and automate bioSAXS data collection. The interface was designed to be user-friendly, for expert and non-expert users alike, to easily define and run bioSAXS experiments, as well as to visualize both raw and processed data. In 2022 BSXCuBE was used by over 277 users during more than 107 individual experiments on BM29. It has been tested on a wide range of biological samples and is constantly being optimized, extended and enhanced to suit the specific needs of bioSAXS users. The user feedback has been positive and found to be instrumental in simplifying bioSAXS data collection.

Abstract publication:

I agree that the abstract will be published on the web site

Posters / 135

User Interfaces for SIRIUS beamlines

Author: Ana Clara de Souza Oliveira[1]

Co-authors: Bruno Vasco de Paula Carlos [1]; Matheus Luís Bernardi [1]; Rafael Lyra [1]; Thiago Donato Ferreira [1]

[1] *Brazilian Synchrotron Light Laboratory (LNLS) - The Brazilian Center for Research in Energy and Materials (CNPEM)*

SIRIUS is a 4th generation synchrotron light source facility that was designed, built and is operated by the Brazilian Synchrotron Light Laboratory (LNLS/CNPEM). Currently, SIRIUS has 6 fully operational beamlines and other 8 beamlines in technical commissioning, scientific commissioning or installation phases. Most SIRIUS beamlines currently have their experiment control solutions based on Python scripts, Jupyter notebooks and desktop graphical interfaces. The desktop graphical user interfaces are mainly implemented in Python and based on the Qt framework, using the

PyDM library on top of EPICS. These solutions brought flexibility and the possibility of implementing graphical interfaces with zero or low-code approaches, making development possible by both support groups and beamline staff. Most of the interfaces that were made available as standardized solutions by support groups have as their main objective the monitoring and control of variables, with a focus on equipment control details. Inspired by positive feedback from beamline staff and users of web solutions such as MxCuBE, web implementations were explored by the team with the aim of offering standardized solutions for experiment control interfaces and the possibility of integration with the new bluesky-based experiment orchestration layer. This presentation will describe the user interface options that have been explored for beamline control and experiment orchestration at SIRIUS.

Abstract publication:

I agree that the abstract will be published on the web site

Posters / **36**

Concept for an exchangeable metadata structure for electronic labbooks based on Mediawiki

Author: Thomas Gruber[None]

Co-authors: David Pape [1]; Martin Voigt [1]; Oliver Knodel [1]

[1] *Helmholtz-Zentrum Dresden-Rossendorf*

An essential challenge by creating FAIR datasets is the often underestimated I, which stands for interoperability. Especially for a dataset that is meant to be exported from its ecosystem, it is important to store the metadata and data in the appropriate exchangeable format based on standards.
One possible source for metadata is an electronic lab notebook that stores it in a structured manner. In many cases the internal structure does not match any established metadata scheme and a mapping is required for a meaningful export. This talk presents a concept of what is necessary to make a generic export from an electronic Labbook based on semantic Mediawiki for ingestion into SciCat or interoperable Nexus files.
First an existing scheme (e.g. Nexus application class) needs to be imported and the class dependencies is stored. Reference to the origin of created classes and properties are essential. Within Mediawiki, templates are used to assign a set of properties to a page. Together with a configuration which describes through which properties the templates are interconnected a generic export is possible. Now by selecting a certain measurement, the export script can extract all essential metadata to create e.g. a SciCat export with project, instrument, sample and dataset information or create a nexus file and knows which groups needs to be created and under which path a property is stored.
Since the information is bound to the property, the application of a mapping can be optimized in an iterative manner. This makes it a flexible procedure that is perfectly usable for existing documentation where metadata schemes are applied at a later stage or need to be updated. In addition, the reference to the original metadata scheme is known in the whole pipeline and could be included in the export.

Abstract publication:

I agree that the abstract will be published on the web site

Posters / **158**

Integrated control of a chip scanner for time-resolved crystallography at the NSLS-II FMX beamline

Author: Robert Schaffer[1]

Co-authors: Kevin Rollet [1]; Martin Fuchs [1]; Venkateswaran Shekar [1]

[1] *Brookhaven National Laboratory*

The FMX (Frontier Microfocusing Macromolecular Crystallography) beamline at the NSLS-II light source has been developing a new experimental station for fixed target time-resolved serial crystallography on biological systems. We present here the controls-system for a chip scanner to enable the rapid collection of large numbers of room temperature crystallographic measurements on biological samples. In addition to static measurements, samples can be excited in a pump-probe scheme by the injection of compounds suspended in liquid through a microdrop dispensing system, at timed intervals preceding the measurement. Enabling this has required the implementation of a full stack integrated solution, involving direct programming of the powerPMAC motion controller, control of motion, triggering and detectors through EPICS, data collection through Ophyd/Bluesky, and the implementation of an optional GUI for control of the experiment. Here, I will outline the components involved in this process, as well as the successes and pitfalls which we have encountered during the implementation and testing of the scanner, and show the results from the first experiments at the beamline.

Abstract publication:

I agree that the abstract will be published on the web site

Posters / **75**

EasyTexture: a new software for data reduction at POWTEX

Author: Oleksandr Koshchii[1]

Co-authors: Andrew Sazonov [2]; Jens Walter [3]; Marina Ganeva [1]

[1] *Forschungszentrum Juelich*

[2] *European Spallation Source*

[3] *Georg-August Universität Göttingen*

POWTEX is a high-intensity TOF diffractometer at the FRM-II research reactor in Garching bei München, Germany. The instrument will serve the needs of the solid-state chemistry, geoscience, and materials science communities through neutron scattering measurements on POWder and TEXture samples.

The important part of the data processing workflow at POWTEX is *data reduction*, i.e. correction of the collected data by experimental artifacts caused by the instrument itself or its environment. The focus of our work is the development of data reduction workflow for texture samples, as post-processing needs for analysis of powder samples at POWTEX are substantially different.

In our poster, we present a new software called EasyTexture, which will help to reduce detector data at POWTEX and prepare texture intensity resolved spectra for analysis within the MAUD package 1. The graphical user interface of EasyTextrure is designed using the EasyScience framework 2, which provides tools for creating intuitive and comprehensive interaction experience for users. Our software is developed as a free and open-source project, as we keep in mind the idea of FAIR and sustainable data ecosphere that enhances open science.

1 Lutterotti, Luca, et al., Z. Kristallogr. Suppl. 26 (2007), 125

2 https://easyscience.software/

Abstract publication:

I agree that the abstract will be published on the web site

Posters / **66**

PiXiu: software for calculating inelastic neutron scattering spectra in four dimensions with high precision

Authors: Ming Tang[1]; Ni Yang[1]; Xiao-Xiao Cai[1]; Zi-Yi Pan[1]

[1] *Institute of High Energy Physics, Chinese Academy of Sciences*

PiXiu is a program bridging the gap between first-principles density functional theory (DFT) and inelastic neutron scattering (INS) dynamical structure factor calculations. In addition to performing powder-averaging for powder samples, PiXiu is capable of calculating the dynamical structure factor in four dimensions for single crystals. Under the hood, PiXiu combines the Quantum Espresso and Phonopy to generate the force constants and phonon properties. A new adaptive sampling policy is implemented to sample Q-point in the 3D reciprocal space. It is applied to calculate the contributions of coherent scattering for one and two quantum order events, resulting in elevated spectrum resolution. Temperature effects on INS spectra are also accounted for the Debye-Waller factors and Bose-Einstein populations. Additionally, PiXiu can optionally obtain the crystal structure according to the material ID of the Materials Project database. PiXiu can be used standalone to generate the neutron dynamical structure factor for various crystals. It can also be used with Prompt, a dual-mode Monte Carlo engine for reproducing thermal scattering experiments, to simulate the whole INS experiment, hence establishing a direct connection between computational and experimental data.

Abstract publication:

I agree that the abstract will be published on the web site

Posters / **68**

ROCK-IT: Remote, operando controlled, knowledge-driven, IT-based catalysis research at large-scale facilities

Authors: Alexander Schökel[1]; André Rothkirch[1]; Britta Höpfner[2]; Christina Widmann[3]; Devin Burke[1]; Hector Perez Ponce[2]; Heike Görzig[2]; Linus Pithan[1]; Mehdi Kazemi[1]; Olaf Schwarzkopf[2]; Rolf Krahl[2]; Sandra Hamann[4]; Shrouk Ehab[1]; Sonal R. Patel[2]; Thomas Gruber[4]; William Smith[2]; Zeynep Isik Dursun[1]

[1] *DESY*

[2] *HZB*

[3] *KIT*

[4] *HZDR*

Insights into “catalysts at work” are of high interest to academic and industrial users, prompting the ROCK-IT project partners DESY, HZB, HZDR, and KIT to enhance capabilities for in situ and operando experiments. ROCK-IT aims to meet the need for a holistic workflow through common remote access protocols, FAIR-data management standards, automation, robotics, experiment and beamline control software including AI aspects, and real-time evaluation. The results of the project will allow for more effective catalyst development and an increased attractiveness for non-expert users as well as users from the industry and will be easily transferable to other scientific areas. Wherever possible, Free and Open-Source Software (FOSS) will be used, and compatibility and integrability with the Helmholtz Federated IT Services will be ensured.

Abstract publication:

I agree that the abstract will be published on the web site

Posters / 152

Automation developments at SOLEIL

Author: Yves-Marie ABIVEN[1]

Co-authors: Arafat Noureddine [1]; Laura Muñoz [1]

[1] *synchrotron SOLEIL*

SOLEIL II project is set to commence in 2025. This project aims to develop an ambitious Diffraction Limited Storage Ring (DLSR) that will enhance brilliance, coherence, and flux. Additionally, the upgrade will include the improvement in the experimental techniques on the beamlines. Automation has been prioritized to adress evolving requirements and simplify user experiences at the beamlines and accelerators operations. SOLEIL aims to develop systems for flexible control of instruments, ranging from manual to fully automated mode of opéeration. This approach involves optimizing experimental procedures, increasing beamline efficiency, and improving sample and data throughput.

As part of its strategic focus , SOLEIL is embracing robotics, with new applications based on 6-axis robot arms. The main purpose is on automating repetitive and time-consuming tasks, such as constant switching between measurements and sample replacements. Recent applications include automatic detector positioning for the NANOSCOPIUM beamline, liquid sample injection for the SWING beamline, and mechanical and magnetic adjustments for the insertion device modules. Through these advancements, SOLEIL is driving towards enhanced automation and operational efficiencies in its state-of-the-art research facilities.

Abstract publication:

I agree that the abstract will be published on the web site

Posters / 111

BEC's Scanning Approach: Devices Get Ready!

Author: Christian Appel[1]

Co-authors: Andreas Menzel [1]; Jan Wyzula [2]; Klaus Wakonig [3]; Matias Guijarro [4]; Mirko Holler ; Sven Augustin [1]; Xiaoqiang Wang

[1] *Paul Scherrer Institut*

[2] *Paul Scherrer Institute*

[3] *PSI*

[4] *Paul Scherrer Institute (PSI)*

The Swiss Light Source at the Paul Scherrer Institut is undergoing an upgrade to a 4th generation synchrotron, presenting an opportunity to enhance its current software stack. To consolidate efforts, a software package for Beamline and Experiment Control (BEC) has been developed, primarily written in Python and leveraging established software tools. For the underlying hardware abstraction layer, BEC utilizes the Ophyd library from NSLS-II.

BEC's server comprises multiple services interconnected by a Redis message broker, also serving as an in-memory database. This architecture offers flexibility in service configuration, enabling adaptability to various requests. In preparation for a scan, the scan server service publishes a message containing all relevant information for the upcoming acquisition. All devices have direct access to this information via the shared memory of Redis and are thus capable of preparing themselves for the upcoming measurement. This concept is envisioned to be applicable to the majority of all implemented scans.

This contribution will elaborate on this idea in more detail, present initial integration tests that implemented the concept in September 2023 with the last light of SLS, and discuss the advantages this brings for beamlines and future operations.

Abstract publication:

I agree that the abstract will be published on the web site

Posters / 122

OpenEM: Open research data infrastructure for electron microscopy

Author: Spencer Bliven[1]

Co-authors: Attila Nacsa [2]; Carlo Minotti [3]; Despina Adamopoulou ; Philipp Wissmann [4]; Sofya Laskina [5]; Yves Tittes [6]

[1] *Paul Scherrer Institute*

[2] *University of Geneva*

[3] *PSI*

[4] *ETH Zurich*

[5] *EPFL*

[6] *University of Basel*

The Open Electron Microscopy Data Network (OpenEM) is a Swiss-wide collaboration to improve data management at electron microscopy (EM) facilities and make the dissemination of EM data open and FAIR (findable, accessible, interoperable, and reusable). The collaboration is based around a central SciCat instance hosted at the Paul Scherrer Institute (discovery.psi.ch), which stores metadata from all facilities and provides data access and management to microscope users, facility managers, and the public.

Electron microscopy has become an essential tool in structural biology and other fields, with modern microscopes producing above 10 TB per day. Managing these datasets is challenging for researchers and facilities producing the data. Making the data publicly accessible is also difficult due to limited integration of existing databases with the EM infrastructure, as well as a lack of repositories for all use cases.

The OpenEM project aims to improve this by providing a central repository for all EM data and metadata from EM facilities in Switzerland. Metadata schemas are under development as part of the Open Science Community for EM, in cooperation with other community-driven ontologies such as PDBx/mmCIF, the EM Glossary, NXem, and the cryoEM ontology. Data is stored on tape at the Swiss National Supercomputing Centre (CSCS) or the ETHZ Long-Term Storage, from which it can be retrieved by collaborators and the public.

Interoperability with other data repositories is a priority, particularly the EMDB and EMPIAR databases for macromolecular data. The OpenEM infrastructure aims to streamline the full data lifecycle for EM data, from data collection through publication and reuse.

Abstract publication:

I agree that the abstract will be published on the web site

Posters / **116**

Multimodal data acquisition system for sub-second time resolution using motor trajectory control in Sardana

Authors: Vanessa Silva[1]; Mirjam Lindberg[1]

Co-authors: Daphne van Dijken [1]; Mikel Eguiraun [2]; Marcelo Alcocer [1]; Michele Cascella [2]; Justus Just [1]; Konstantin Klementiev [1]

[1] *MAX IV Laboratory*

[2] *MAX IV*

The Balder beamline is placed at the 3 GeV storage ring at MAX IV Laboraty (a 4th generation synchrotron) and is dedicated to X-ray absorption and emission spectroscopy in the energy range of 2.4–40 keV 1. The beamline can deliver a very high photon flux of 1013 ph/s and is suitable for experiments under in situ / operando conditions. This kind of experiment requires support for fast continuous scan with good data quality and support for sequential multi-techniques data acquisition (DAQ) for a better analysis and study of the sample dynamics 2.
This work describes the multimodal DAQ system implemented at the Balder beamline, combining the complementary techniques X-ray absorption spectroscopy (XAS) and X-ray diffraction (XRD) in a single experiment. To achieve stable and quick energy scans, multiple energy edges and XRD acquisition positions are loaded all at once to the motor controller ACS SPiiPlusEC. The motion is performed in trajectory which allows sequential energy spectrums and XRD acquisitions to be performed sequentially in sub-second time resolution. The experiment synchronization is performed with PandABox 3, a FPGA trigger unit, which can generate TTL signals based on the motors current position described in a sequencer table. For the multimodal system, each technique (XAS and XRD) related detectors receive a different pulse train based on the current energy motor position. Furthermore, the experiment is orchestrated and performed with Sardana [4]. A dedicated Sardana scan macro is used to load XAS-XRD scan conditions as well to prepare the trajectory in the motor controller. User-friendliness is achieved using dedicated Taurus GUI on top of Sardana, where each technique can be sequenced and configured by a visual drag-and-drop system.

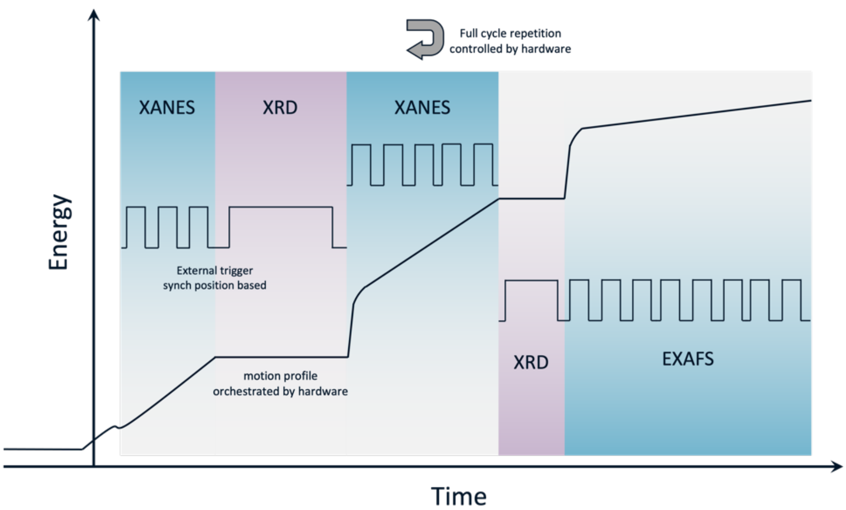

Figure 1: Multimodal XAS-XRD experiment diagram correlating sequential XAS energy scan with XRD acquisition positions together with generated external triggers

1 K. Klementiev, K. Norén, S. Carlson, K. G. V. S. Clauss, and I. Persson, "The balder beamline at the max iv laboratory,"Journal of Physics: Conference Series 712, 012023 (2016)

2 1. K. Klementiev, H. Tarawneh, and P. F. Tavares, "Wiggler radiation at a low-emittance storage ring and its usage for X-ray absorption spec- troscopy,"Journal of Synchrotron Radiation 29, 462–469 (2022)
3 S. Zhang et al., "PandABox: A Multipurpose Platform for Multi-technique Scanning and Feedback Applications", in Proc. ICALEPCS'17, Barcelona, Spain, Oct. 2017, pp. 143–150. doi:10.18429/JACoW-ICALEPCS2017-TUAPL05
[4] T. M. Coutinho et al., "Sardana: The Software for Building SCADAS in Scientific Environments" , in Proc. ICALEPCS'11, Grenoble, France, Oct. 2011, paper WEAAUST01, pp. 607–609.

Abstract publication:

I agree that the abstract will be published on the web site

Posters / 102

Construction of XASDB

Author: ⊠⊠⊠[1]

Co-authors: ⊠⊠⊠ [1]; ⊠⊠⊠ [1]

[1] *IHEP*

X-ray Absorption Spectroscopy (XAS) is a pivotal technique for material characterization. In the field of XAS, data assessment typically involves comparing the data with reference spectra from previous measurements, placing high demands on the quality of the spectra and measurement data. In order to advance data-driven scientific research, it is urgent to establish a reusable infrastructure for XAS data.
Over the past decade, we have collected a series of standard and well-characterized XAS data on the core beamline B8 at the BSRF. Additionally, in the future, the High Energy Photon Source (HEPS) build in China is estimated to generate a large volume of XAS data. Those massive amount of data provides a solid foundation for the construction and ongoing development of the database.
Foucsing on these problems, we have designed and implemented a database for XAS data. This database aims to integrate multiple resources of XAS data, while supplementing information on the characteristics, sources, and analysis of spectral information. Then, We have established specifications and standards for relevant data formats and provided comprehensive displays of both raw and processed data. Furthermore, we have provided a spectral matching tool capable of identifying the most similar spectra in the database to a given input data.
This database will provide effective foundational support for the opening and sharing of XAS data, enabling users to conveniently access, share, and reuse the required scientific data, thereby enhancing the value of the data and accelerating the productivity of scientific output.
In the future, this database will be available for free access and the researchers can conveniently utilize it for further analysis, such as matching with their own experiments data and developing machine learning models.

Abstract publication:

I agree that the abstract will be published on the web site

Posters / 110

BEC Widgets: A Modular GUI Framework for Beamline Experiment Control

Author: Jan Wyzula[1]

Co-authors: Christian Appel ; Ivan Usov [2]; Matias Guijarro [3]; Sven Augustin [1]; Klaus Wakonig [2]

[1] *Paul Scherrer Institute*

[2] *PSI*

[3] *Paul Scherrer Institute (PSI)*

BEC Widgets is an innovative Qt-based GUI framework specially designed to provide graphical user interfaces for the Beamline Experiment Control (BEC), tailored for users at the Swiss Light Source at Paul Scherrer Institute. BEC Widgets ensures seamless integration and a plug-and-play experience that significantly improves workflow efficiency and interactivity for beamline scientists.

With shared memory access through a REDIS server, BEC Widgets facilitates the development of modular, independent GUI components covering all aspects of beamline operations, including data visualization, device/scan control, and scripting. Users can drive experiments through GUI desktop applications, modular widgets with a CLI interface using the RPC, and event scripting. This modular and flexible design significantly reduces setup times, increases data accuracy, and enhances experiment flexibility, addressing specific challenges in beamline control.

Utilizing the QtPy abstraction layer, BEC Widgets ensures compatibility with major Qt Python bindings, PyQt and PySide. This compatibility extends the framework's adaptability and future-proofs it against updates in the Qt ecosystem. BEC Widgets leverages PyQtGraph for high-performance data visualization, providing dynamic and interactive graphical representations of experimental data. This capability allows users to manage complex data sets efficiently, enhancing the analytical and operational aspects of beamline experiments.

Abstract publication:

I agree that the abstract will be published on the web site

Posters / **96**

Controls for dynamic tomography at the TOMCAT beamlines

Author: Istvan Mohacsi[1]

Co-authors: Anne Bonnin [1]; Christian Matthias Schlepütz [1]; Federica Marone [1]; Goran Lovric [1]; Marco Stampanoni [1]; Tine Celcer [1]

[1] *Paul Scherrer Institute (CH)*

Computed tomographic microscopy is one of the milestone techniques of X-ray science. The I- and S-TOMCAT beamlines of the Swiss Light Source (SLS) specialize in in-vivo, in-situ and operando high-speed tomography with over 10 kHz sustained frame rates. The improved flux and brilliance of the new TOMCAT beamlines present new frontiers for dynamic applications. Hence, they will serve as a showcase to outline the next generation controls infrastructure of the ongoing SLS 2.0 upgrade. The beamlines'control systems will be constructed as a combination of an EtherCAT-based motion backbone and custom systems for high-speed motion, synchronization, and triggering. The talk will focus on their implementation, configuration and interfacing with the data path and other sub-systems via EPICS and a new high-level beamline experimental control system.

Abstract publication:

I agree that the abstract will be published on the web site

Posters / **95**

High-Speed 2D detector DAQ at the ESRF

Authors: Alejandro Homs Puron[1]; Laurent Claustre[1]; Samuel Debionne[1]

Co-authors: Andrew Goetz [1]; Jens Meyer [1]

[1] *ESRF*

The ESRF EBS upgrade has meant a tremendous increase in X-ray photon flux in the experimental beamlines (BLs), requiring faster and more advanced DAQ techniques. Faster and larger 2D detectors are being developed and need to be integrated into the BL control system. The BLISS control software, designed to push the BL instrumentation to its limits, fulfills the opportunities offered by this upgrade. As a result, the control of 2D detector DAQ must meet the challenge of an increasing frame rate, data throughput, online data reduction (ODR) tasks, as well as the simultaneous access from Online Data Analysis (ODA) software. This work presents the latest optimizations in LIMA (a.k.a. LIMA1) in order to cover the 1-10 kHz frame rate range, efficiently deal with the increasing demand of ODR tasks and the coexistence with advanced ODA applications such as Ewoks. Furthermore, we cover the recent developments in the LIMA2 project components: the core library, the detector plug-ins and the processing pipelines. New core features include the state machine and an improved control of the processing pipeline Tango clients. In addition to the PSI Jungfrau detector, in production since 2023, while the Dectris Eiger2 and the ESRF Smarpix detectors are starting to be used on the BLs, new detectors like the Rigaku XSPA-1M and the Fraunhofer IIS LAD (Large Area Detector) for BM18 are being integrated. New processing pipelines for XPCS and tomography are under development as well. Finally, the LIMA2 integration into BLISS is also presented.

Abstract publication:

I agree that the abstract will be published on the web site

Posters / 93

Handling different analysis workflows in a modular framework

Author: Malte Storm[1]

[1] *Helmholtz-Zentrum Hereon*

Helmholtz-Zentrum Hereon operates multiple X-ray diffraction (XRD) experiments for external users and while the experiments are very similar, their analysis is not. Pydidas 1 is a software package developed for the batch analysis of X-ray diffraction data. It is published as open source and intended to be widely reusable. Integration is based on the ESRF's pyFAI package.

Because the wide range of scientific questions tackled with the technique of XRD, a limited number of generic tools will not be sufficient to allow all possible analysis workflows. Easy extensibility of the core analysis routines is a key requirement. A framework for creating plugin-based workflows was developed and integrated in the pydidas software package to accommodate different analytical workflows in one software tool. We present the architecture of the pydidas workflows and plugins along with the tools for creating workflows and editing plugins.

Plugins are fairly simple in design to allow users/collaborators to extend the standard pydidas plugin library with tailor-made solutions for their analysis requirements. Access to plugins is handled through a registry which automatically finds plugins in specified locations to allow for easy integration of custom plugins. Pydidas also includes (graphical) tools for creating and modifying workflows and for configuring plugins, as well as for running the resulting workflows.

While pydidas was develop with the analysis of X-ray diffraction data in mind and the existing generic analysis plugins reflect this field, the architecture itself is very versatile and can easily be re-used for different research techniques.

1 https://github.com/hereon-GEMS/pydidas

Abstract publication:

I agree that the abstract will be published on the web site

Posters / **92**

Fast Data Analyser for X-ray spectroscopy beamlines

Author: Przemyslaw Karczmarczyk[1]

Co-author: Carlo Marini [1]

[1] *ALBA Synchrotron*

FDA (Fast Data Analyser) is an application developed at ALBA Synchrotron to analyse the data produced with X-ray Absorption Spectroscopy (XAS), as well as with X-Ray Diffraction (XRD). It provides a quick and convenient way of loading, processing, and analysing the data with different methods, such as XANES (X-ray absorption near edge structure) normalization, EXAFS (Extended X-ray absorption fine structure) extraction and Fourier Transforming, LCA (Linear combination analysis), PCA (Principal component analysis), or MCR (Multivariate Curve Resolution). Raw data from transmission, fluorescence, and electron yield experiments can be loaded and preprocessed in different ways, including: deglitching, self-absorption correction, or background subtraction. With this versatile tool utilizing multiple numerical algorithms, the user can perform quick calculations with selected technique, extract analysis results together with plots, and export them externally in various formats.

FDA is a user-friendly application that comes with a full-featured forms-based environment for interacting with and processing the data implemented using PyQt toolkit. It is currently capable of loading hundreds of spectra at once and provides access to a number of features of XAS & XRD techniques, making it a robust & flexible tool for any type of analysis. Already in use at several spectroscopy-oriented beamlines at ALBA Synchrotron, it provides a simple yet powerful way of quickly analysing the data produced in the experiments.

Abstract publication:

I agree that the abstract will be published on the web site

Posters / **31**

CAMEO: Orchestrate, communicate with any app

Author: Yannick Le Goc[None]

CAMEO[1] is a general-purpose middleware providing services to manage distributed heterogeneous applications. It has been developed in the Instrument Control Service at the ILL to solve many common issues e.g. how to have a *C++* control server, many *Java* GUI clients displaying plots calculated by a *Python* library, or how to integrate a proprietary software running on a dedicated *Windows* OS, or how to restart safely the entire control system of an instrument from a single entry point without synchronization issues?
CAMEO provides services in two main topics. First, it provides a way to orchestrate the remote applications which can be started, stopped, monitored, synchronized. Second, high-level communication patterns are provided i.e. Request/Respond, Publish/Subscribe. A third simple Function pattern specific to CAMEO is also provided. The CAMEO interface is simple and hides the complexity of the network programming layer making a powerful easy-to-use tool for non-experts.

Moreover CAMEO is very versatile as it was designed to adapt to many environments. For instance, a black box program can be controlled by CAMEO and interact with other CAMEO applications. It is also possible to orchestrate the applications with CAMEO and make them communicate with other protocols like *HTTP*, *AMQP*, etc.
All that makes CAMEO very interesting for a control system environment where heterogeneous applications running in different languages on different platforms have to be integrated.
CAMEO is now a mature middleware and in this article we present an overview of the functionalities including the proxy router to secure communications between two networks and with a focus on the ILL environment.

1 https://code.ill.fr/cameo/cameo

Abstract publication:

I agree that the abstract will be published on the web site

Posters / **39**

The Design of HDF5 Data Formats for HEPS

Authors: Fazhi Qi[1]; Hao Hu[2]; Haofan Wang[1]; Qi Luo[1]

[1] *Institute of High Energy Physics, CAS*

[2] *Institute of High Energy Physics, Chinese Academy of Sciences*

The High Energy Photon Source (HEPS) encompasses a variety of experimental types, including diffraction, scattering, imaging, and spectroscopy. The data generated from these experiments are highly dimensional, uncertain, and computationally complex. Considering the users' needs for interoperable data analysis and high-performance I/O processing, it is necessary to organize and manage the data and metadata in a rational manner. We have adopted HDF5 and NeXus as the unified data format standards for HEPS. Sequentially interfacing with the first phase of 14 beamlines, we have organized the data formats and summarized common requirements, eventually developing three generic data formats. These serve as foundational templates for the subsequent phases of 90 beamline stations.

Initially, the report will introduce the characteristics of experimental data types at HEPS and the current status of experimental data formats. From the perspectives of experimental users and beamline station managers, we will detail the organization and specification requirements of various experimental data HDF5 formats designed according to NeXus. Furthermore, we will demonstrate the data format organization schemes with examples from the B2 Hard X-Ray Nanoprobe Multimodal Imaging Beamline and the B7 Hard X-Ray Imaging Beamline, both of which are high-throughput stations. Finally, we will summarize some challenges and future plans concerning HEPS data formats.

Abstract publication:

I agree that the abstract will be published on the web site

Posters / **44**

The Transition from CentOS to Rocky Linux at SESAME

Authors: Ibrahim FOUDEH[1]; Mostafa Zoubi[2]; Salman Matalgah[2]

Co-authors: Abdalla Ahmad [2]; Amro Aljadaa [2]; Anas Abbadi [2]; Anas Mohammad [2]; Malik Almohammad [2]; Rami Khrais [2]

[1] *Sesame*

[2] *SESAME (Synchrotron-light for Experimental Science and Applications in the Middle East)*

The emergence of Rocky Linux, a community-driven enterprise operating system, represents a significant shift in SESAME's Scientific computing landscape. This paper explores the transition from CentOS to Rocky Linux by the computing group, examining motivations, advancements, and implications institution-wide.
In 2008, the Computing and Control group at SESAME initially adopted Scientific Linux as the primary Linux distribution due to its maintenance by CERN and widespread use in research communities. However, technical challenges and the cessation of support for Scientific Linux led to a switch to CentOS in 2013. Unfortunately, CentOS's announced of reaching End of Life (EOL) by end of 2021, means no further updates or support, posing significant challenges for ongoing operations and development at SESAME.
Motivated by the need for stability and long-term support amid changes in CentOS, the decision to migrate to Rocky Linux ensures continuity in computing operations. Technical aspects of Rocky Linux development, such as system configurations and integration, are examined, highlighting collaborative efforts within SESAME to enhance the operating system.
Adoption of Rocky Linux fosters open collaboration and innovation, empowering research teams to tackle complex tasks. Challenges, including compatibility issues, are addressed through proactive planning and testing.
In summary, SESAME's switch to Rocky Linux shows their commitment to teamwork and using the best tools available. This change demonstrates SESAME's ability to adapt to new challenges, driving progress in science and innovation.

Abstract publication:

I agree that the abstract will be published on the web site

Posters / 64

The DESY use case or: From demonstrator beamline instances towards usefulness and large scale deployment.

Author: Regina Hinzmann[None]

Co-authors: Igor Khokhriakov [1]; Anton Barty ; Peter van der Reest ; Linus Pithan [2]

[1] *DESY, FS-SC*

[2] *DESY*

DESY joined DESY the SciCat collaboration later than other big European laboratories. In Jan 2022 it was decided to adopt SciCat to DESY needs. Since then, IT has set up an OpenStack and Kubernetes infrastructure maintaining about 14 SciCat instances of which a handful are actively used. This infrastructure allows us to tackle various issues independently. We would touch upon those topics which we think are most important to overcome first to reach our overall goal of deploying SciCat at all beamlines (23+1 at PETRAIII and 4 at FLASH).

At DESY, it was difficult to agree on what metadata actually is, what SciCat shall be used for. Initially, we went from ingesting what is technically possible to developing an uniformed validation layer before ingestion.

The DESY proposal management system (DOOR) has its own way of identifying persons and organisations of granting access to the beamline for a certain beamtime. Access to data and access to computing resources are handled by yet another system (Gamma Portal) and need to be ported over to SciCat to allow for browsing and downloading data.

Within a national programme (DAPHNE), DESY has the opportunity to provide an open data catalogue for numerous national institutions. How this is technically solved is to be seen. Currently, various technologies are available and being sounded out.

Definitely, DESY aims to provide DOIs for DESY data. This workflow is currently investigated and to be implemented.

Abstract publication:

I agree that the abstract will be published on the web site

Posters / **109**

NXRefine: An Automated Workflow for Single Crystal X-ray Diffuse Scattering

Authors: Raymond Osborn[1]; Matthew Krogstad[1]; Stephan Rosenkranz[1]; Guy Jennings[1]; Justin Wozniak[1]

[1] *Argonne National Laboratory*

Recent advances in synchrotron x-ray instrumentation have enabled the rapid acquisition of x-ray diffraction data from single crystals, allowing large contiguous volumes of scattering in reciprocal space to be collected in a matter of minutes, with data rates of several terabytes per day. NXRefine implements a complete workflow for both data acquisition and reduction of single crystal x-ray scattering to produce three-dimensional reciprocal space maps 1. Advanced workflows already exist for the generation of Bragg peak intensities, but the goal of NXRefine is to generate a three-dimensional mesh of scattering intensity that includes both Bragg peaks and the diffuse scattering that arises from deviations from the average structure. After the initial refinement of the sample orientation, the workflow is automated in order to reduce data in real time so that it is available for inspection before a set of measurements, e.g., as a function of temperature, are complete. Furthermore, the results can be transformed into 3D-ΔPDF maps or analyzed with machine learning tools that extract the temperature dependence of peak intensities in thousands of Brillouin zones. The workflow is written as a set of Python modules that can either be run from the command line, launched from a GUI that is implemented as a plugin to NeXpy 2, or by submitting jobs to a batch queue, using an integrated workflow manager.

This work was supported by the U.S. Department of Energy, Office of Science, Basic Energy Sciences, Materials Sciences and Engineering Division.

1. https://nexpy.github.io/nxrefine/
2. https://nexpy.github.io/nexpy/

Abstract publication:

I agree that the abstract will be published on the web site

Invited Speakers / **166**

20 years of the COD –disseminating crystallographic data

Author: Saulius Grazulis[1]

[1] *VU Institute of Biotechnology, Life Science Center*

For more than 20 year, the Crystallography Open Database (COD) collects published crystal structure data and makes it available on the Web under the CC0 license in an organised, machine readable and searchable for. Currently, the collection of the COD has over 500 thous. records and is used for crystal analysis, material identification, DFT calculations, machine learning, teaching and much more. To be usable for such applications, the COD data must satisfy certain quality criteria. All data that are deposited to the COD undergo tree levels of checks –syntax checks, semantic validation against the IUCr dictionaries and COD specific crystallographic checks based in the IUCr journal publication rules.

```
Over the years, software tools were developed for these tasks that are routinely used in the C
```

```
The use of Open Source software, the support of the community, European and Lithuanian funders
```


Beamline Control Systems / **46**

Hyperion: Ultra-High Throughput Automated Macromolecular Crystallography Data Collection Using the Bluesky Framework

Authors: David Perl[1]; Dominic Oram[1]; Oliver Silvester[1]

Co-authors: Graeme Winter [1]; Paul Hathaway [1]

[1] *Diamond Light Source*

At Diamond Light Source, several Macromolecular Crystallography (MX) beamlines focus on, or include, completely automated data collection. This is used primarily for high throughput collection, which has historically meant several hundred samples per day. Diamond is building its next generation, service-based, data acquisition platform Athena using NSLS-II's Bluesky experiment orchestration library. Using this platform, the MX data acquisition team at Diamond is developing an application named Hyperion to increase the possible rate of automated MX data collection, both for immediate use and in preparation to take advantage of the upgraded Diamond-II synchrotron, due in several years. The automated data collection routines are currently built on legacy experiment orchestration software which includes a lot of redundancy originally implemented for safety when human users are controlling the beamline, but which is inefficient for automated data collection processes.
Using an agile development approach, we have systematically replaced sections of the original data collection routines with calls to the new Hyperion service. This has allowed continuous uptime for delivery of automated data collection to users, without interruption by lengthy testing periods, while immediately seeing the incremental throughput gains as the new sections are implemented.
Hyperion currently runs on beamline I03 and, although it has not yet been thoroughly optimised, is responsible for collecting around 1000 samples per day. Although this is already double the throughput of the legacy implementation there is a requirement to increase this to 5000 samples a day, which Hyperion is expected to achieve by the end of the year. Currently available routines include triggering robot exchanges of samples, sample orientation and centring, and finally collection of rotational data, which represent everything needed for standard data collection. Detailed timing metrics are collected about all of the above to facilitate further speed-up. To come are fluorescence scanning, collection strategy determination, and rotation data collection at multiple positions on a single sample pin –which will enable us to reach the high end of the target throughput –as well as expansion to other MX beamlines.

Abstract publication:

I agree that the abstract will be published on the web site

Beamline Control Systems / **134**

Bluesky at SIRIUS light source

Author: Ana Clara de Souza Oliveira[1]

Co-authors: Allan da Silva Pinto [1]; Paulo Baraldi Mausbach [1]; Rafael Lyra [1]; Thiago Donato Ferreira [1]

[1] *Brazilian Synchrotron Light Laboratory (LNLS) - The Brazilian Center for Research in Energy and Materials (CNPEM)*

SIRIUS is a 4th generation synchrotron light source facility that was designed, built and is operated by the Brazilian Synchrotron Light Laboratory (LNLS/CNPEM). Currently, SIRIUS has 6 fully operational beamlines and other 8 beamlines in technical commissioning, scientific commissioning or installation phases.

The distributed control system is based on EPICS and the software solutions for orchestrating experiments adopted by the beamlines are mostly based on scripts and graphical user interfaces implemented in Python. Baseline beamline software is implemented and presents satisfactory results, enabling operation for external users. However, there is plenty of room for improvements in terms of robustness, dead times in scan scripts, code standardization and experimental data standardization.

In some of the beamlines that are in operation, the bluesky project packages were explored in customized solutions, which demonstrated that this framework could provide solutions to the known problems. The bluesky project has been explored as a standard software platform for orchestration on new SIRIUS beamlines, in addition to providing solutions that could replace legacy software on beamlines already in operation in the future.

This presentation will describe the strategies adopted in the implementation of bluesky on SIRIUS beamlines, which aim to facilitate the standardization and reuse of solutions between beamlines, enable collaboration between support groups and beamline staff to implement orchestration, in addition to enabling control of experiments in single queues of tasks in each beamline, accessed both via Python scripts and notebooks, and via graphical user interfaces.

Abstract publication:

I agree that the abstract will be published on the web site

Beamline Control Systems / 35

Karabo goes AMQP: Replacement of the Core Communication Broker

Authors: Alessandro Silenzi[1]; Anna Klimovskaia[1]; Dennis Goeries[1]; Gero Flucke[1]; Gianpietro Previtali[1]; Janusz Malka[1]; Janusz Szuba[1]; Noushadali Anakkappalla[1]; Raul Costa[1]; Sergey Esenov[1]; Steffen Hauf[1]

[1] *European XFEL GmbH*

Karabo is the control and data processing framework operating the instruments and photon beamlines at the European XFEL. Its event driven nature is enabled by a central message broker that distributes control information to subscribed software processes.
Originally, Karabo was developed using the Java Messaging System (JMS) broker, and the OpenMQc library to interface it from C++ and Python. This library is not maintained since years and a process stalling broker communication unnoticed for a few hours can put the entire broker communication at risk. Accordingly, it was decided to investigate new broker technologies.
This contribution will show the careful steps taken to move Karabo communication to other brokers, explain why finally the RabbitMQ broker with the AMQP protocol was chosen and discuss the experiences after the Karabo production environments at the European XFEL have completely moved to RabbitMQ.

Abstract publication:

I agree that the abstract will be published on the web site

Beamline Control Systems / 99

Integrating blissdata within PSI Beamline Experiment Control (BEC) system

Author: Matias Guijarro[1]

Co-authors: Christian Appel ; Jan Wyzula [2]; Klaus Wakonig [3]

[1] *Paul Scherrer Institute (PSI)*

[2] *Paul Scherrer Institute*

[3] *PSI*

BEC (Beamline Experiments Control) is a new Python-based experiment control software currently developed within the Paul-Scherrer Institute (PSI). It will be available to Swiss Light Source (SLS) users starting from January 2025 after the SLS 2.0 upgrade program. BEC provides services dealing with every aspect of a modern beamline control software.

Blissdata is a Python library developed at the European Synchrotron Radiation Facility (ESRF) as part of the Beamline Instrumentation Support Software (BLISS) project. It is especially designed to store acquisition data from beamline experiments, and provides memory management and efficient data streaming as well as an h5py-like API to fetch data from memory or from saved files transparently for clients like data processing or data visualization.

Embarking Blissdata within BEC is an example to promote synergies between development teams towards better interoperability and collaboration within the synchrotron community, for the benefit of synchrotron users.

This contribution makes a retrospective of Blissdata integration within BEC, with a focus on technical aspects.

Abstract publication:

I agree that the abstract will be published on the web site

Beamline Control Systems / 38

Implementing Bluesky at APS: A Strategy for Enhanced Data Acquisition and Collaboration

Author: Fanny Rodolakis[1]

Co-authors: Eric Codrea [1]; Joseph Sullivan [1]; Peter Jemian [1]

[1] *Argonne National Laboratory*

The Advanced Photon Source (APS) at Argonne National Laboratory is at the forefront of facilitating groundbreaking scientific research by providing state-of-the-art X-ray capabilities. Recognizing the critical role of software in maximizing the scientific output of user facilities, APS has embarked on a strategic deployment of Bluesky, a comprehensive software framework designed for data acquisition, management, and analysis in synchrotron research. This initiative aims not only to enhance the productivity of facility users through more intuitive and efficient data handling but also to foster

collaboration within and across user groups by standardizing data acquisition processes.
Bluesky's architecture, developed with the flexibility to adapt to a wide range of experimental setups, aligns with NOBUGS' vision of better software for better science. It offers an open-source solution that encourages contributions from the scientific community, fostering collaborative development and sharing of best practices. At APS, the deployment strategy includes comprehensive training for staff and users, development of user-friendly interfaces, and integration with existing data management systems to ensure a seamless transition and adoption.
This presentation will outline the strategic approach taken by APS for the Bluesky deployment, highlighting the challenges faced, solutions implemented, and the anticipated benefits in terms of operational efficiency, scientific discovery, and community engagement. We will also discuss the importance of collaborative initiatives and share insights into how Bluesky can serve as a model for other facilities looking to upgrade their data acquisition and analysis software ecosystem. By sharing our journey, we aim to contribute to the NOBUGS community's ongoing dialogue on improving scientific software and to inspire collaborative efforts that enhance the capabilities and impact of user facilities worldwide.

Abstract publication:

I agree that the abstract will be published on the web site

Beamline Control Systems / 130

Bluesky UI at Advanced Light Source

Author: Seij De Leon[1]

Co-authors: Alexander Hexemer [1]; Antoine Islegen-Wojdyla [1]; Dylan McReynolds [1]; Per-Anders Glans [1]

[1] *Lawrence Berkeley National Laboratory*

The learning curve for beamline control systems is often challenging due to the use of command line controls, or various custom made GUIs. Access to beamlines is limited and time constrained, so learning command line controls or scripts takes up valuable time that could be used for the experiment.

This talk focuses on the creation of a web browser application that acts as a beamline controls system. This app eliminates the learning curve for beamline controls by providing an intuitive browser based interface that utilizes a modern open source controls system (Bluesky 1) on the backend. It also creates a common interface that can be used at all EPICS-based 2 beamlines to unify the user experience.

By studying workflows of beamline scientists at ALS, we established a set of minimum user requirements as a starting point to decide the essential UI components and functionality. Additionally, existing projects that provide similar browser based UI for control systems were reviewed including ESRF Daiquiri 3, Diamond Light Source UI Library [4], and the ANSTO Beamline GUI [5]. These existing applications and UI libraries helped guide decisions on suitable frameworks and architecture.

The Web-browser app was initially developed using simulated motors and area detectors to mock connections with physical hardware devices to enable offline development and testing. We then deployed the solution with physical devices at the ALS beamline 5.3.1.
The current web app works with both simulated devices and physical devices. Live updates to motor positions and area detector frames can be displayed on the browser, and the motor positions can be set with the app.

The initial results prove that a browser based controls system is feasible at ALS. Future development will continue to expand on this core functionality to incorporate the Bluesky Run Engine for experiment planning and the Queue Server for instrument coordination.

References:

1 Allan, D. et al. (2019) 'Bluesky's Ahead: A Multi-Facility Collaboration for an a la Carte Software Project for Data Acquisition and Management', Synchrotron Radiation News, 32(3), pp. 19–22. doi: 10.1080/08940886.2019.1608121.

2 EPICS, https://epics.anl.gov/

3 Fisher S, Oscarsson M, De Nolf W, Cotte M, Meyer J. Daiquiri: a web-based user interface framework for beamline control and data acquisition. J Synchrotron Radiat. 2021 Nov 1;28(Pt 6):1996-2002. doi: 10.1107/S1600577521009851. Epub 2021 Oct 29. PMID: 34738955; PMCID: PMC8570207.

[4] Diamond Light Source, https://diamondlightsource.github.io/web-ui-components/

[5] ANSTO, https://github.com/bluesky/bluesky-web/files/14053727/ANSTO.Australian.Synchrotron.beamline.GUIs.p

Abstract publication:

I agree that the abstract will be published on the web site

Beamline Control Systems / **72**

Ophyd-Async (formerly Ophyd v2) update

Author: Abigail Emery[None]

Bluesky facilities have been using ophyd as a comprhensive hardware abstraction layer for step scanning use cases but legacy constraints have made it difficult to move towards flyscanning. We present an update on the status of ophyd-async (formerly ophyd v2), which has now seen action in user experiments at both Diamond and NSLS-II to perform hardware triggered measurements.

Ophyd-async can be deployed alongside/instead of ophyd and focuses on correctness, testability and easy integration of generic flyscanning solutions. We summarize the performance improvements, additions to the device API, feedback from user-facing operations and information about the joint agile development programme running between Diamond and NSLS-II.

Abstract publication:

I agree that the abstract will be published on the web site

Beamline Control Systems / **129**

Enhancing Sardana Continuous Scans: Insights from a Collaborative Workshop

Authors: Johan Forsberg[1]; Jordi Aguilar Larruy[2]; José Centeno Gabadinho[2]; Michal Piekarski[2]; Oriol Vallcorba[3]; Roberto Homs Puron[2]; Steven Wohl[2]; Vanessa Silva[4]; Zbigniew Reszela[2]

[1] *MAX IV*

[2] *ALBA Synchrotron*

[3] *ALBA Synchrotron (CELLS)*

[4] *MAX IV Laboratory*

Over a decade ago, Sardana 1 integrated generic continuous scans 2, initially meeting only basic requirements. More complex scenarios, were implemented by migrating generic logic into plugins,

often relying on hooks or ad-hoc solutions. In the past year, SOLARIS Synchrotron and the Sardana Community co-hosted a collaborative continuous scans workshop 3 with participation from similar frameworks like Bliss and Flyscan, as well as various institutions. This event fostered an open exchange of experiences and approaches. Subsequently, a midterm roadmap for Sardana enhancements emerged from the workshop discussions. This paper outlines the initial steps taken along this roadmap, focusing on:

- Implementing the acquisition of experimental channels at different rates
- Introducing the concept of shutters
- Establishing primary and secondary relationships of synchronizers

These enhancements are being deployed across different beamlines at ALBA and MAXIV Synchrotrons.

1 T. M. Coutinho et al., "Sardana: The Software for Building
SCADAS in Scientific Environments", in Proc. 13th Int.
Conf. on Accelerator and Large Experimental Control Systems (ICALEPCS'11), Grenoble, France, Oct. 2011, paper WEAAUST01, pp. 607-609
2 Z. Reszela at al., "Iterative Development of the Generic Continuous Scans in Sardana", in Proc. 15th Int. Conf. on Accelerator and Large Experimental Control Systems (ICALEPCS'15), Melbourne, Australia, Oct. 2015, pp. 524-528, doi:10.18429/JACoWICALEPCS2015-TUB3O02
3 https://indico.solaris.edu.pl/event/5

Abstract publication:

I agree that the abstract will be published on the web site

Experiments automation / 33

MXCuBE Goes Serial

Authors: Antonia Beteva[1]; Jerome Kieffer[None]; Marcus Oskarsson[1]

Co-authors: Daniele De Sanctis [1]; Jens Meyer [1]; Julien Orlans ; Samuel Debionne [1]; Samuel Rose [1]; Shibom Basu [2]

[1] *ESRF*

[2] *EMBL*

MXCuBE goes serial

M.Oskarsson, A.Beteva, D. de Sanctis, S.Basu, J. Orlans, S.Rose, S. Debionne, A. Homs, J. Kieffer, J.Meyer

The ESRF Extremely Brilliant Source upgrade programme included the construction of the new ID29 beamline, the first in the world beamline fully dedicated to room-temperature experiments and time-resolved macromolecular serial crystallography. The beamline presents a new layout in terms of design, and the experimental setup was designed to fully benefit from the characteristics of the EBS machine. The primary objective is to collect diffraction data from micrometer sized crystalline samples at room temperature with a microsecond time resolution. The beamline presents a flexible sample environment that can accommodate fixed target, viscous injectors, microfluidics or tape drives. The experimental setup is completed with a Jungfrau 4M detector that has been integrated in the ESRF data acquisition pipeline and can be operated at high data acquisition rates.

MXCuBE, the state-of-the-art experiment control software used at all ESRF MX beamlines, for remote, high-throughput attended and unattended data collections, was expanded with new Serial Macromolecular Crystallography features for ID29. Several new scanning functions have been introduced to perform microsecond data acquisitions on Silicon chips, foils, jets, and tape drive experiments. Additionally, the implementation of newly developed or prototyped chips can be easily

added. A custom dialog has been developed within MXCuBE for the chips so that the user can easily select which parts of the chip to scan. The PSI Jungfrau 4M detector is controlled through Bliss and the latest version of the LImA2 library, which integrates an automatic hit-finding algorithm and optional online frame rejection and compression.

Abstract publication:

I agree that the abstract will be published on the web site

Data Analysis / **17**

HTTomo, a new high-throughput GPU tool for large tomographic data processing at DLS

Authors: Daniil Kazantsev[1]; Jessica Verschoyle[1]; Naman Gera[1]; Yousef Moazzam[1]

[1] *Diamond Light Source*

HTTomo stands for High Throughput Tomography pipeline for processing and reconstruction of parallel-beam tomography data. The HTTomo project was initiated in 2022 at Diamond Light source in anticipation of major data increase with the Diamond-II upgrade. With the support of modern developments in the field of High Performance Computing and multi-GPU processing, the main goal is to mitigate I/O bottlenecks and through GPU acceleration enable higher throughput for big data.

HTTomo is a user interface (UI) written in Python for parallel data processing using MPI protocols. It orchestrates I/O data operations and enables processing on a CPU and/or a GPU using computing cluster or a personal computer. HTTomo utilises other data analysis libraries, such as TomoPy for CPU and HTTomolibgpu for GPU, as backends for data processing. The methods from the libraries are exposed through YAML templates to enable fast task programming and pipeline building.

The main concept of HTTomo is to split the data optimally in accordance to the given computational resources and methods requirements. It is a GPU-memory aware system that exploits the GPU device in a way, that the data stays, for as long as possible, on the device. We use CuPy's API to help with the data transfers and faster computations.

In the presentation, the main elements of HTTomo's framework will be demonstrated, as well as, the benchmarks with the current tomographic software in production at DLS, Savu.

Abstract publication:

I agree that the abstract will be published on the web site

Data Analysis / **67**

The Challenge and Change of Data Analysis at China Spallation Neutron Source II

Author: Junrong Zhang[1]

[1] *China Spallation Neutron Source*

Since the user program of China Spallation Neutron Source (CSNS) was open in 2018, eight neutron beamlines have been operational, with the number of users reaching approximately 6000. The data

portal of CSNS provides the services in data access, data reduction, data analysis and simulations for over 130,000 experimental runs. It is continuously evolving to meet the requirements of users for data analysis and software. In the March of 2024, the upgrade project of CSNS II was launched, which will result in the construction of nine neutron instruments, in addition to muon and proton stations. This will present new challenges for the data analysis and management too.

The perspectives and design of data analysis for CSNS II will be presented. The presentation will address a number of key issues, including the real-time processing and reduction of event data streams, the complete simulation of neutron scattering experiments using Monte Carlo/Molecular Dynamics/Density Functional Theory, the web-based access to data visualization and remote control, the reconstruction of neutron imaging data based on High Performance Computing and Artificial Intelligence, and the possible applications of artificial intelligence in data analysis.

Abstract publication:

I agree that the abstract will be published on the web site

Data Analysis / 118

Frame Tripping at the EMBL P14 Beamline: from Raw Data to Tomographic Reconstruction in 30s

Authors: Marina Nikolova[1]; Gleb Bourenkov[1]; Jonas Albers[1]

Co-authors: Angelika Svetlove [1]; Andrea d'Amato ; Matthew Lawson [1]; Ksenia Denisova [1]; Stefan Fiedler [1]; Thomas R. Schneider [1]; Elizabeth Duke [1]

[1] *EMBL*

Since the initial proof-of-principle X-Ray imaging experiments at the EMBL beamline P14 at Petra III 1, a high throughput tomography (HiTT) setup has been established 2, and is now offered as a regular user service. The setup features a high performance data acquisition and processing software system.
Here we present a recent highlight: the achievement of significant gains in the 3D reconstruction speed. We are now automatically producing a fully reconstructed data set ~30s after the completion of the acquisition. This includes the flat-field correction and phase retrieval of 4 x 1810 frames of size 4 Megapixels followed by reconstruction using TomoPy Gridrec 3. The implementation of the latter has been modified in order to reduce its memory usage and to limit any application of padding i.e. extraneous looping over frames to only once. Furthermore Gridrec has been incorporated in our phase-retrieval software: *TOMO CTF*, thereby bypassing all of the TomoPy python wrapper routines and directly calling its C routines instead.
MPI 4 has been employed for cluster inter-node communication thus removing the previously necessary step of reading the CTF-corrected data from disk before proceeding with the reconstruction.
A variable number of cluster nodes can be configured for the tomographic processing. By default we are using 4 during regular beamtime data acquisition, a total of 352 cores. The data is written directly onto a 1.2 PB parallel file cluster storage, BeeGFS.
Through a further refactoring of the software more advanced acquisition strategies can also be automatically processed fast. At present the reconstruction of a 360 degree rotation with an extended field of view is being added to *TOMO CTF*.
A newly integrated PCO camera contributes to a more stable and error-free operation.
These recent developments have been instrumental in providing streamlined and user-friendly computational services during HiTT experiments at P14.
We thank the BMBF for funding the beamline compute resources ('Verbundforschung', 05K16GU1 and 05K19GU1).

1 Polikarpov, M., Bourenkov, G., Snigireva, I., Snigirev, A., Zimmermann, S., Csanko, K., Brockhauser, S. & Schneider, T. R. (2019). Acta Cryst. D75, 947–958.
2 Albers J, Nikolova M, Svetlove A, Darif N, Lawson MJ, Schneider TR, Schwab Y, Bourenkov G and

Duke E. High Throughput Tomography (HiTT) on EMBL beamline P14 on PETRA III. Journal of Synchrotron Radiation (2024). DOI: 10.1107/S160057752300944X
3 Gürsoy D, De Carlo F, Xiao X, and Jacobsen C. Tomopy: a framework for the analysis of synchrotron tomographic data. Journal of Synchrotron Radiation, 21(5):1188–1193, 2014.
4 https://www.mpich.org/

Abstract publication:

I agree that the abstract will be published on the web site

Data Analysis / 89

High-performance coherent X-ray imaging data analysis using PyNX

Author: Vincent Favre Nicolin[1]

Co-authors: Jerome Kieffer ; Clément Atlan [2]; Yuriy Chushkin [2]; Olga Stamati [3]; Pierre Paleo [2]; Steven Leake

[1] *ESRF-The European Synchrotron*

[2] *ESRF*

[3] *Univ. Grenoble Alpes*

4th generation synchrotron sources provide two orders of magnitude more coherent photons, and thus the ability to collect coherent X-ray imaging datasets faster and/or with a higher resolution. Consequently, the increased volume of data requires dedicated tools to fully take advantage of the improved coherent flux.

PyNX[1,2,3] is developed at ESRF - it has been written from the ground up to provide state-of-the-art performance using optimised GPU programming, both in terms of speed and memory requirements (to handle larger datasets). It is used on multiple beamlines notably at ESRF, Soleil, Petra-III, Sirius, TPS, with scripts for data analysis easily expandable to new instruments (only the data input functions need to be updated). Input/output using the standard CXI format is also supported.

We will present the applications to various experimental techniques: Coherent Diffraction Imaging (CDI) and Ptychography (far field and near field) for two and three-dimensional imaging, also in the Bragg geometry to provide strain information in nano-crystals, and finally holo-tomography (currently in development).

References:
1 J. Appl. Cryst. 53 (2020), 1404, arXiv:2008.11511 main PyNX documentation: http://ftp.esrf.fr/pub/scisoft/PyNX/doc/
3 git repository: https://gitlab.esrf.fr/favre/PyNX

Abstract publication:

I agree that the abstract will be published on the web site

Data Analysis / 150

McStas: 25 years of simulating neutron instrumentation and experiments

Author: Peter Willendrup[1]

Co-authors: Emmanuel Farhi [2]; Kim Lefmann [3]; Mads Bertelsen [4]

[1] *European Spallation Source and DTU Physics*

[2] *Synchrotron SOLEIL*

[3] *University of Copenhagen*

[4] *European Spallation Source*

The McStas[1-3] neutron Monte Carlo ray-tracing simulation project was started at RISØ in 1997 and has thus now served the neutron scattering community for more than 25 years.
The presentation will give a brief overview of highlights from the 25 year history of McStas and further update the NOBUGS community on recent developments and future plans for both McStas and its X-ray counterpart McXtrace4.

1 K. Lefmann and K. Nielsen, Neutron News 10, 20, (1999).
2 P. Willendrup, and K. Lefmann, Journal of Neutron Research, 22, p.1 (2020)
3 P. Willendrup, and K. Lefmann, Journal of Neutron Research 23, p.7 (2021)
4 E.B. Knudsen et al. Journal of Applied Crystallography, 46 p.679 (2013)

Abstract publication:

I agree that the abstract will be published on the web site

Data Reduction / 51

Digital Twin Design and Implementation at the Institut Laue-Langevin

Author: Shervin Nourbakhsh[1]

[1] *Institut Laue-Langevin*

Data from virtual experiments are becoming a valuable asset for research infrastructures: to develop and optimize current and future instruments; to train in the usage of the instrument control system; to study quantifying and reducing instrumental effects on acquired data. Furthermore large sets of simulated data are also a necessary ingredient for the development of surrogate models (supervised learning) for faster and more accurate simulation, reduction and analysis of the data.
So far, the production and usage of data from virtual experiments have been mostly reserved to simulation experts. With this work, data from virtual experiments are made available to the general users. The presented framework wraps in a digital twin of the facility instrument the knowledge of the physical description of the instrument, the simulation software and the high performing computing setup. The twin presented in this article has been developed at the ILL in the framework of the PANOSC European project in close collaboration with other research facilities (ESS and EuXFel) for some of its essential components. An overview of the core simulation software (McStas), its Python API (McStasScript), the public instrument description repository and the instrument control system (NOMAD) are given. The choices on the the communication patterns (based on ZQM) and interaction (via CAMEO) between the different components are also detailed. Example twins are also presented. The adoption of FAIR principles for data formats and policies in combination with open-source software make it a sustainable project both for development and maintenance in the mid and long-term.

Abstract publication:

I agree that the abstract will be published on the web site

Data Reduction / 58

Dranspose: A constrained map-reduce live analysis pipeline for fast experimental feedback

Author: Felix Engelmann[1]

Co-author: Paul Bell [1]

[1] *Max IV, Lund University*

During beamtimes, critical decisions on how to proceed with an experiment must be made constantly. As a result, it is important to provide feedback with the best possible data analysis, mostly in the form of visualizations, with the lowest possible latency. For low data rates, writing and monitoring a file works well. However, processing tens of gigabytes per second is difficult with a filesystem, and streaming solutions are therefore preferable. Apart from the high throughput, it is important that the users consuming the feedback are able to quickly tweak the analysis to adapt to changing conditions. While there exist generic stream processing frameworks like Apache Spark, or application-specific OnDA (JAC 46, 2016), our goal is to combine all major relevant features from those tools.

We propose Dranspose (http://dranspo.se), a data analysis pipeline for high-throughput data acquisitions. It is a horizontally scalable, distributed system deployable to a Kubernetes cluster, relying on Redis for coordination and ZeroMQ for data streaming. As a programming paradigm, we propose a novel constrained map-reduce. In a classical map-reduce, every event is processed by an independent worker which forwards their result to a reducer. Often, experimental data has temporal dependencies such as the evolution over time, which in standard map-reduce is only analyzable in the reducer. Our constrained map-reduce enables stateful workers combined with load balancing for fast temporal difference calculations by sending consecutive frames to the same worker (see Figure). Additionally, it enables event formation from multiple sources to get a complete data view for a specific trigger. This is especially useful to normalize one detector by another. The reducer writes the analyzed data to an HSDS service, which provides standardized, simultaneous access for visualization tools like silx, h5pyd, or h5web. Dranspose takes care of the orchestration and distribution of data but allows users to easily adapt and update the map and reduce Python functions. For ease of development, we provide record and replay actions to develop the analysis on a local machine before the experiment.

Several applications for Dranspose include merging rotation stage encoder positions to a CMOS camera at 3GB/s for live tomographic reconstruction, azimuthal integration with normalization to I_0, live XRF concentration mapping, or crystallography spot finding. With gained confidence from the scientific users, Dranspose is useful for data reduction, and the feedback may be directly consumed by the control system to decide on the progression of a scan in a closed loop.

Abstract publication:

I agree that the abstract will be published on the web site

Data Reduction / 98

Fast GPU-friendly indexing (cell reorientation) algorithm for X-ray crystallography.

Author: Hans Christian Stadler Kleeb[1]

Co-author: Markus Janousch [2]

[1] *Paul Scherrer Institute*

[2] *Paul Scherrer Institut*

Authors: Piero Gasparotto, Luis Barba, Hans-Christian Stadler,
Greta Assmann, Henrique Mendonça, Alun W. Ashton, Markus Janousch,
Filip Leonarski and Benjamín Béjar

TORO (TOrch-powered Robust Optimization) 1 is a new algorithm for indexing diffraction patterns, applicable when the unit cell geometry is known. Originally based on the PyTorch framework a dedicated version in CUDA has been developed. The algorithm is capable of indexing frames in less than a millisecond. It therefore allows real-time data processing. Not only being fast its quality is also on par with the XGandalf indexing algorithm. A user-friendly interface allows the configuration to the particular needs of an experiment and the inclusion into the CrystFEL data processing pipeline.

1. Gasparotto P, Barba L, Stadler H-C, Assmann G, Mendonça H, Ashton A, et al. TORO Indexer: A PyTorch-Based Indexing Algorithm for kHz Serial Crystallography. ChemRxiv. 2023; doi:10.26434/chemrxiv-2023-wnm9n (submitted to Journal of Applied Crystallography, in proof).

Abstract publication:

I agree that the abstract will be published on the web site

Data Reduction / 59

Integrated Sample Management, Data Acquisition and Processing for MX at the Canadian Light Source

Author: Michel Fodje[1]

[1] *Canadian Light Source, Inc*

The Canadian Macromolecular Crystallography Facility (CMCF) consists of two beamlines (CMCF-ID and CMCF-BM). The beamlines are operated through a modern computer software system for on-site and remote collection. It consists of a user-friendly graphical user interface for experiment-focused data collection (MxDC), a laboratory information management system for remote planning, experiment monitoring, beamline intelligence and data access (MxLIVE), and an automated pipeline for data reduction and strategy determination (AutoProcess).
Together, these software systems provide an integrated environment facilitating remote access to the facility, which has been an invaluable resource for the Canadian crystallographic community since 2006. More than 95% of access to the facility is now by remote access.

Abstract publication:

I agree that the abstract will be published on the web site

Data Reduction / 140

ASAPO: A high-speed streaming framework to support an automated data-processing pipeline.

Authors: Mikhail Karnevskiy[1]; Tim Schoof[1]

Co-author: Martin Gasthuber [1]

[1] *DESY*

Modern scientific experiments often generate large amounts of data, posing challenges for real-time processing and analysis. ASAPO, a high-performance streaming framework developed at DESY, addresses these challenges by providing a robust solution for online and offline data processing. Leveraging TCP/IP and RDMA over Ethernet and Infiniband, ASAPO facilitates high-bandwidth communication between detectors, storage systems, and analysis processes.
ASAPO offers user-friendly interfaces for C/C++ and Python on all major platforms, streamlining the
development of data processing pipelines. A high-level Python library reduces boilerplate code and enables the creation of complex analysis workflows with ease. Key features include automatic retransfer, trivial parallelization on a per-image basis, support for multi-module detectors, and web-based
monitoring capabilities.
Several experimental facilities at Petra III already benefit from ASAPO, employing it in various dat-aprocessing
pipelines. Examples include azimuthal integration of X-ray scattering data, peak finding, and indexing of diffraction patterns. These applications demonstrate ASAPO's versatility and effectiveness in accelerating scientific discovery.

Abstract publication:

I agree that the abstract will be published on the web site

Data Reduction / 86

The High-throughput Data I/O framework for HEPS

Authors: Shiyuan Fu[1]; Yu Hu[2]; Rui Liu[1]; Hao-Kai Sun[3]; Jian Liu[4]; Lei Wang[1]; Shuang Wang[1]

[1] *IHEP*

[2] *IHEP, CAS*

[3] *Computing Center, Institute of High Energy Physics, Chinese Academy of Sciences*

[4] ☒☒☒☒☒☒☒☒☒☒☒☒

The High Energy Photon Source (HEPS) is estimated to produce a substantial volume of raw data, presenting significant computational challenges in scientific research. To address this problem, we have developed a high-throughput data I/O framework specifically tailored for HEPS, aimed at mitigating the I/O bottlenecks.Firstly, within this framework, we have devised a unified I/O interface for computational tasks, which serves to shield the difference in underlying data sources and formats. Subsequently, an asynchronous prefetch method has been integrated into the framework to expedite data read and write speeds. This includes the dynamic adjustment of prefetch data volume based on computational tasks and memory space, thereby optimizing the utilization of computational node memory space. Lastly, in order to overcome the issue of slow data access resulting from the process of writing data to disk and subsequent reading, the framework has been extended to encompass a streaming data module. This module dynamically parses data streams from the DAQ and stores them in a distributed cache pool, thereby further accelerating the data retrieval process through the utilization of data streams.

Abstract publication:

I agree that the abstract will be published on the web site

Data Reduction / 106

DAMNIT

Author: Thomas Michelat[1]

[1] *European XFEL*

The Data And Metadata iNspection Interactive Thing (DAMNIT) is a tool developed by the Data Analysis group at the European XFEL (EuXFEL) to help scientists and users effortlessly create overviews of their experiments.

Traditionally, at EuXFEL, many user groups and beamline scientists use spreadsheets and electronic logbooks to track experimental settings, metadata, and analysis results. This manual process is time-consuming, error-prone, and lacks reproducibility. DAMNIT aims to replace these spreadsheets with an automated solution.

Integrated into the EuXFEL data system, DAMNIT collects metadata automatically and manages data analysis workflows during experiments, giving users a real-time overview of their progress.

DAMNIT offers a simple frontend application with a spreadsheet-like interface showing all tracked quantities for each data measurement. It allows users to set up complex data analysis workflows on large computing clusters using the Slurm scheduling system. This streamlines data analysis during experiments, with results saved for later use through an accessible API and displayed in the overview table. DAMNIT also makes it easy to update computed and collected data using a simple Python API and re-processing features.

Abstract publication:

I agree that the abstract will be published on the web site

Remote user interfaces / 54

SVG based control system synoptic solution on the web application Taranta

Authors: Yimeng Li[1]; Vincent Hardion[2]; Mikel Eguiraun[1]; Johan Forsberg[1]

Co-authors: Dominika Trojanowska [3]; Matteo Canzari [4]; Hélder Ribeiro [5]; Ajaykumar Dubey [6]; Athos Georgiou [7]; Valentina Alberti [8]

[1] *MAX IV*

[2] *MAX IV Laboratory*

[3] *S2Innovation*

[4] *INAF-OAAB*

[5] *Atlar Innovation*

[6] *Persistent Systems*

[7] *CGI*

[8] *INAF-OATs*

Presently at the MAX IV laboratory, Scalable Vector Graphics (SVG) based synoptic are widely used for monitoring accelerator and beamline statuses due to their intuitive, interactive and easy controllable feature. However, synoptic views have historically been confined to desktop applications, rendering remote access less user-friendly. In addition, supporting and managing updates of synoptic views entails a challenging process to allow scientists to modify the synoptic themselves. Can we simplify access to synoptic and reduce the need for software developers to engage in the development process? Facilitating user autonomy in managing control systems holds high importance at

MAX IV. The existing integrated user interface, Taranta, allows scientists to effectively monitor and control their experiment. This paper introduces the SVG based web synoptic solution on Taranta and the applications in the MAX IV laboratory. Synoptic on Taranta enables remote access to control systems and allows for basic component updates through simple configurations. Moreover, it offers users a convenient overview of all connected Tango devices. This solution promises a significant enhancement in user experience and a notable reduction in the development cycle.
KEYWORD: SVG, Synoptic, Taranta, control system

Abstract publication:

I agree that the abstract will be published on the web site

Invited Speakers / 165

Linking Scientific Instruments and Computation: Patterns, Technologies, Experiences

Author: Ian Foster[1]

[1] *ANL*

Powerful detectors at modern experimental facilities routinely collect data at multiple GB/s. Online analysis methods are needed to enable the collection of only interesting subsets of such massive data streams, such as by explicitly discarding some data elements or by directing instruments to relevant areas of experimental space. Such online analyses require methods for configuring and running high-performance distributed computing pipelines—what we call flows—linking instruments, data center computers (e.g., for analysis, simulation, AI model training), edge computing (for analysis), data stores, metadata catalogs, and high-speed networks. In this talk, I review common patterns associated with such flows, describe methods for instantiating those patterns, and present experiences with the application of these methods to the processing of data from a range of experimental facilities, each of which engages HPC resources for data inversion, machine learning model training, or other purposes. I also discuss implications of these new methods for operators and users of scientific facilities.

FAIR data management / 139

Extending FAIR data management with processed data integration in the ESRF data portal

Authors: Alejandro De Maria Antolinos[None]; Andrew Goetz[1]; Guillaume Gaisne[1]; Mael Gaonach[None]; Marjolaine Bodin[1]

[1] *ESRF*

ince its launch in August 2020, the ESRF's Extremely Brilliant Source (ESRF-EBS) has resulted in a notable transformation in data management, driven by the heightened volume and complexity of data. The ESRF data policy 1 has been aligned in 2024 by the inclusion of processed data.
The ongoing development of the ESRF data portal 2 based on the ICAT metadata catalogue 3, initiated in 2017, has been upgraded to facilitate the integration and presentation of processed data. A standardized framework has been established across all beamlines to incorporate processed data into the portal, and will provide users with the possibility to reprocess data as needed. Based on a micro front-end architecture, the new data portal 4 offers customized displays tailored to specific techniques, exemplified by the MX data. We will illustrate the seamless integration of specialized

portals for diverse domains, such as the Human Organ Atlas 5 or the Paleontology Portal [6]. In addition an overview will be given of the features implemented for sample tracking, e-logbook, minting dois, and tools for data managers.

References
1 https://doi.org/10.15151/ESRF-DC-1534175008
2 https://data.esrf.fr
3 https://repo.icatproject.org/site/icat/server/5.0.0/
4 https://data2.esrf.fr
5 https://human-organ-atlas.esrf.eu/
[6] https://paleo.esrf.fr

Abstract publication:

I agree that the abstract will be published on the web site

FAIR data management / 126

Tiled in the Context of Data and Metadata Services

Author: Dan Allan[1]

Co-authors: Dylan McReynolds [2]; Eugene Matviychuk [3]; Juan Marulanda [3]; Padraic Shafer [3]; Stuart Campbell ; Tom Caswell ; Wiebke Koepp [2]

[1] *Brookhaven National Laboratory*

[2] *Lawrence Berkeley National Lab*

[3] *BNL*

Providing users with remote and random access to structured data is emerging as an important challenge for user facilities in the next decade. Our peers in industry and in other scientific areas are building such services. Tiled is a solution tuned to the requirements of user facilities, applying web standards and widely-adopted technologies in numerical computing to address search, random structured remote access, transcoding between storage and access formats, authentication, and authorization. Tiled is an open source project, primarily developed by contributors from NSLS-II and ALS, deployed at facilities including APS and BESSY-II, and increasingly being explored for adoption at others.

This talk takes a broad look at the ecosystem of data and metadata services. It aims to place the goals of Tiled, and other projects like it, in context. This speedy review will touch on traditional file transport services, metadata catalogs, and on other structured data access services. It will explore their various goals and features; how these support practical data visualization and analysis during and after the experiment, archival and discoverability; and the goals of FAIR data.

Abstract publication:

I agree that the abstract will be published on the web site

FAIR data management / 47

The Data Catalogue service and use case at SESAME

Authors: Malik Almohammad[1]; Mostafa Zoubi[1]; Salman Matalgah[1]

Co-authors: Alejandro De Maria Antolinos [2]; Andy Götz [2]; Gianluc Iori [1]; Ibrahim FOUDEH [1]; Kevin Philipps [3]; Marjolaine Bodin [2]; Patrick Austin [3]; Rolf Krahl [4]

[1] *SESAME (Synchrotron-light for Experimental Science and Applications in the Middle East)*

[2] *ESRF*

[3] *Science and Technology Facilities Council (STFC, UKRI)*

[4] *Helmholtz-Zentrum Dresden - Rossendorf (HZDR)*

In today's research facility's landscape, experimental data management, metadata catalogue and access are playing a vital role on enabling the full research lifecycle, allowing the users communities and scientific institutions to collaborate, transfer and share data on a well-defined collaborative platform.
Following the community best practices on delivering and exporting data, SESAME is implementing the Data Portal with key services to authenticate, access, organize and sharing experimental data to researchers. Collaboratively a decision was taken by the Scientific Computing group at SESAME to consider the metadata catalogue service namely ICAT, this solution has a good reputation based on its wide use in many research labs as an open-source solution, further this step will foster and drive the SESAME data policies adaptation and assure compliancy with the FAIR data principles.
This paper explains the ingestion journey and workflow of SESAME Experimental Data (SED) to the ICAT metadata catalogue service, this process always starts on an early stage at the call for proposal (CFP) at SESAME user portal (SUP), data generated per beamtime are in compliance with the Data acquisition system standards and stored on the short-term storage (STS) and moved and archived on Long-term Storage (LTS) at a later stage.
Data portal service is a fully integrated user experience with SESAME research infrastructure subsystems, it ensures smooth delivery of data management, browsing and access and with some helpful plugins such as H5web view and E-logbook. The current development is executed with the support of The European Synchrotron Radiation Facility (ESRF) and the EU project BEATS, and current challenges and future enhancement will be addressed as well as lessons to learn.

Abstract publication:

I agree that the abstract will be published on the web site

FAIR data management / 156

Rolling out a sample lifecycle and experimental data catalog at SOLEIL

Author: Gwenaëlle Abeillé[1]

Co-authors: Frédéric Potier [1]; Patrick Madela [1]

[1] *Synchrotron SOLEIL*

SOLEIL Information System has a 20-year legacy characterized by non-uniform and siloed IT solutions that have been continuously evolving in response to changing business requirements, thereby increasing its complexity. A redesign of our information system architecture was deemed necessary to address this challenge, requiring a new, homogeneous, and flexible approach.

Currently, we are in the process of implementing new use cases to support our core mission, such as hosting diversified samples from external users on SOLEIL experimental stations. We have initiated this process by introducing a brand-new application dedicated to the lifecycle of samples, alongside an existing experimental data catalog called SciCat 1. Given that these two applications rely on numerous data exchanges with existing and future applications, we have integrated their development and integration efforts within a comprehensive architectural framework,

which includes a technical platform with Apache Kafka and WSO2 solutions.

This presentation will outline our architectural approach aimed at addressing the future evolution of the SOLEIL information system.

1 https://scicatproject.github.io/

Abstract publication:

I agree that the abstract will be published on the web site

FAIR data management / 16

DOMAS: a data management software framework for advanced light sources

Author: Hao Hu[1]

[1] *Institute of High Energy Physics, Chinese Academy of Sciences*

In recent years, China's advanced light sources have entered a period of rapid construction and development. As modern X-ray detectors and data acquisition technologies advance, these facilities are expected to generate massive volumes of data annually, presenting significant challenges in data management and utilization. These challenges encompass data storage, metadata handling, data transfer, and user data access. In response, the Data Organization Management Access Software (DOMAS) has been designed as a framework to address these issues. DOMAS encapsulates four fundamental modules of data management software, including metadata catalogue, metadata acquisition, data transfer, and data service. For light source facilities, building a data management system only requires parameter configuration and minimal code development within DOMAS. This paper firstly discusses the development of advanced light sources in China and the associated demands and challenges in data management, prompting a reconsideration of data management software framework design. It then outlines the architecture of the framework, detailing its components and functions. Lastly, it highlights the application progress and effectiveness of DOMAS when deployed for High Energy Photon Source (HEPS) and Beijing Synchrotron Radiation Facility (BSRF).

Abstract publication:

I agree that the abstract will be published on the web site

FAIR data management / 115

Latest news from the SciCat ecosystem: ElasticSearch integration and Job sub-system

Authors: Carlo Minotti[1]; Despina Adamopoulou[None]; Dylan McReynolds[None]; Igor Khokhriakov[None]; Massimiliano Novelli[2]; Regina Hinzmann[None]; Spencer Bliver[1]

[1] *PSI*

[2] *European Spallation Source*

SciCat is an open-source data catalog providing data management, annotation, and publishing features for scientific facilities (https://scicatproject.github.io/). It enables tracking of data provenance, annotation with metadata, and publication of datasets with a unique DOI. SciCat is built on a flexible microservice architecture, allowing easy configuration for diverse use cases. The adoption of

SciCat by multiple research facilities has helped its community to expand development, improve functionality and usability.

Development of SciCat is currently focused on version 4, which migrated the codebase to a more modern technology stack based on Typescript and the node.js framework Nest.js. The latest release improves search functionality through ElasticSearch integration, and the next one will improve interoperability with other services through a highly configurable job sub-system.

ElasticSearch integration is part of the effort of increasing data FAIRness, and comes as a consequence of the increased number of use cases that the community has access to and their review. During the review process, it became evident that the search capabilities needed to be improved and expanded with a better free text search option, similar to a search engine. This led to the choice of ElasticSearch, which is the industry standard.

The growing number of adopters and their varying landscape of IT infrastructure has exposed a need for flexible integration with additional services, such as archiving systems or custom APIs. The choice has been to design a configurable job sub-system which is dedicated to manage and dispatch configurable commands with third party services through a variety of protocols and systems, including calls to REST APIs, posting messages to RabbitMQ queues or Kafka topics. This allows each facility to easily configure SciCat to interact with services that are or will be deployed and best suit their needs.

We will present how SciCat integrates with ElasticSearch, provide an overview of the technical challenges, and a few examples. We will then provide an overview of the configurable jobs sub-system, present the idea behind it and the use cases that drove their design (such as the OpenEM project from the Swiss electron microscopy facilities). We will highlight the implementation details and the current status of this effort. We will conclude with a brief overview of future development.

Abstract publication:

I agree that the abstract will be published on the web site

FAIR data management / 15

Navigating the Data Patchwork: Strategies for Integrating Metadata Catalogs, Data Publications, and Archives

Author: Oliver Knodel[1]

Co-authors: David Pape [1]; Guido Juckeland [1]; Martin Voigt [1]; Stefan Müller [1]; Thomas Gruber [1]

[1] *Helmholtz-Zentrum Dresden-Rossendorf*

In the ever-expanding landscape of data management, navigating the diverse array of metadata catalogs such as SciCat, data publications on Invenio derivatives, and internal archives presents a formidable challenge. However, with the right strategies, this mosaic of data can be effectively combined and represented to unlock its full potential. In this talk, we delve into the intricacies of data fusion, exploring innovative approaches to harmonize metadata catalogs, data publications, and archives seamlessly.

We will discuss the importance of interoperability and standardization in facilitating the integration process, enabling disparate data sources to coalesce into a cohesive ecosystem. Through conceptional examples and case studies, we will provide insights into the practical application of strategies.

Abstract publication:

I agree that the abstract will be published on the web site

FAIR data management / 108

ALBA: Towards FAIR data management principles

Authors: Emilio Centeno[1]; Fernan Saiz[2]; Fulvio Becheri[1]; Gemma Rosas[1]; Marc Armenter[1]; Nicolas Soler[1]; Oriol Vallcorba[1]; Rodrigo Cabezas[1]; Zbigniew Reszela[2]

[1] *ALBA Synchrotron (CELLS)*

[2] *ALBA Synchrotron*

ALBA Synchrotron 1 is actively implementing FAIR data management principles 2 across all operational beamlines. Data is cataloged in ICAT, 3 preferably using the NeXus data format 4, alongside metadata sourced from various information systems.

To ensure all metadata is accessible for data interpretation and reuse, gathering beamline and experimental conditions during data collection is essential. Most ALBA beamlines utilize Sardana 5 as the data acquisition framework, built atop a TANGO Control System [6].

To standardize the process as much as possible without disrupting scientists' habits and needs, an additional layer of processing is considered between the control system and ICAT ingestion. This involves publishing necessary information to a Redis [7] database, enabling consumers to retrieve it for relevant operations. The published information and operations vary depending on the beamline and technique and two general scenarios are considered:

1. For beamlines generating NeXus files, most metadata is encapsulated within these files, and only the file location details are published to Redis. A NeXus recorder, configurable for different Application Definitions, has been developed for beamlines that use Sardana.
2. In cases where NeXus files are not generated (e.g., some proprietary software), the folder containing collected files is published to Redis, along with additional processing information if needed. The intermediate layer is responsible for NeXus file generation tailored to each technique before ICAT publication.

Publication to Redis can be achieved directly from Sardana macros or via a Tango Device Server. Using Redis as the message broker aligns with Sardana's roadmap, which considers publishing scans data to Redis to decouple acquisition from storage. ICAT pubication will also benefit of this feature when available.

1 Alba Synchrotron, http://www.cells.es

2 Wilkinson, M., *et al.* (2016). The FAIR Guiding Principles for scientific data management and stewardship. *Sci Data* **3**, 160018. https://doi.org/10.1038/sdata.2016.18

3 Collaboration, T. I. C. A. T. (2014). The ICAT Project. The ICAT Collaboration. https://doi.org/10.5286/SOFTWARE/ICAT

4 Konnecke, M., *et al.* (2015). The NeXus data format. *J. Appl. Cryst.* **48**, 301-305.

5 Sardana, https://sardana-controls.org/

[6] Tango Control System, http://www.tango-controls.org

[7] Redis, https://redis.io/

Abstract publication:

I agree that the abstract will be published on the web site

FAIR data management / 147

The data pipeline for the European Spallation Source ERIC

Authors: Fredrik Bolmsten[1]; Jesper Rude Selknaes[1]; Thomas Holm Rod[1]; Torben Roland Nielsen[1]

[1] *European Spallation Source*

ESS was born with open and reusable data in mind. Based on lessons learned from other research infrastructures, the data pipeline for experiments at the European Spallation Source ERIC (ESS) was outlined from the very beginning and designed to allow for FAIR data and real time data processing and analysis. In this presentation we will present the integrated data pipeline at the ESS, its current status, and how it supports the generation of FAIR data. We will walk through the data pipeline step by step, beginning with the user office software, over experiment control, data processing and analysis and concluding with the data management step, where ESS plays an active role in the development of the Photon and Neutron Open Science Cloud (PaNOSC) in the realm of the European Open Science Cloud.

Abstract publication:

I agree that the abstract will be published on the web site

FAIR data management / 124

Ontological definition of experimental techniques for FAIR data

Authors: Andrew Goetz[1]; Ioannis Koumoutsos[None]; Wout De Nolf[1]

[1] *ESRF*

Ensuring the accessibility of FAIR data within the scientific community is crucial, especially given the vast volume of generated data and its potential for future reprocessing. While findability is often perceived as the simplest aspect of FAIR, achieving it necessitates robust ontology harmonization within an open science framework. Without it, scientists may struggle to efficiently locate datasets.

To address this challenge, it is imperative to employ common ontology concepts to enrich the metadata associated with conducted experiments. A pivotal initial step involves integrating the Photon and Neutron Experimental Techniques (PaNET) ontology as it offers invaluable identification of the techniques utilized in each dataset. To align the techniques employed at the ESRF with their semantically equivalent terms in PaNET, the development of an ESRF-Experimental Techniques ontology is deemed the optimal approach for initiating semantic negotiations with PaNET. The overarching objective is to uphold semantic integrity as defined by ESRF scientists, while fostering an ongoing dialogue with PaNET and between the domain experts, enabling them to resolve ambiguities that might impede mutual understanding, comparison, and dialogue.

In this presentation, we will outline the methodology employed to construct this bridging ontology, which endeavors to define and differentiate techniques based on their distinct characteristics. A similar approach could be adopted at other large-scale facilities.

Abstract publication:

I agree that the abstract will be published on the web site

FAIR data management / 88

DRACO-HELIPORT integration for metadata enhanced data-acquisition

Author: Mani Lokamani[1]

Co-authors: David Pape [2]; Oliver Knodel [1]; Stefan Bock [1]; Ulrich Schramm [1]; Jeffrey Kelling [2]; Guido Juckeland [2]

[1] *Helmholtz-Zentrum Dresden-Rossendorf*

[2] *Helmholtz-Zentrum Dresden - Rossendorf (HZDR)*

The Dresden laser acceleration source (DRACO) is a state-of-the-art high-power ultra-short pulse laser system[1,2], that uses an Amplitude Technologies Pulsar architecture to form main and diagnostics beams at different focal lengths and target density conditions. The setup can deliver from 6J to 45J of pulse energy at a typical pulse duration of 30fs and a typical frequency of 1Hz. During the diagnostic phase, the beam characteristics are recorded in the form of images and several instrument parameters, that shape the beam to desired characteristics.

In this talk, we present our approach of implementing FAIR principles to DRACO operations and monitoring using our in-house guidance system HELIPORT3, with the goal of making them reusable irrespective of the downstream experiment. We employ FAIR workflows4 to post-process data collected by DRACO's built-in data acquisition system and enrich it with metadata for subsequent utilization in machine-learning and optimization algorithms for accurate control of the beam characteristics. The integration of DRACO and HELIPORT demonstrates the first step towards establishing a digital twin for the laser source facility at HZDR.

References

1 First results with the novel Petawatt laser acceleration facility in Dresden, U. Schramm et al, J. Phys. Conf. Ser. 874 012028 (2017)
2 High dynamic, high resolution and wide range single shot temporal pulse contrast measurement, T. Oksenhendler et. al., Opt. Express 25, 12588-12600 (2017)
3 HELIPORT: A Portable Platform for FAIR {Workflow | Metadata | Scientific Project Lifecycle} Management and Everything, O. Knodel et. al., P-RECS (2021)
4 FAIR Computational workflows, C. Goble et. al., Data Intelligence (2020) 2, 108 (2020)

Abstract publication:

I agree that the abstract will be published on the web site

Metadata / **52**

SciLog –An Electronic Logbook for User Experiments: Live at PSI

Authors: Carlo Minotti[1]; Klaus Wakonig[1]

[1] *PSI*

We present the latest developments, the future directions and the current adoption of SciLog at PSI. SciLog is an electronic logbook developed and operated at PSI, which stores unstructured information captured by humans and systems during an experiment. It aims to replace paper notebooks often used during experiments and enhance scientists' experience by enabling them, e.g., to jot unstructured data and track to-dos. It supports collaborative editing and can, but is not limited to, display information as time-ordered messages in an interactive "chat-like"view.

It is based on a REST API layer which uses the Loopback4 framework, and stores entries in a Mongo database to ensure flexibility. The User Interface relies on Angular. It supports facility-specific authorisation systems, thanks to its OIDC integration, and can automatically import new experiment proposals for a better user experience.

In addition to presenting the software, we will focus on its maintenance, use and infrastructure at PSI, mentioning its evolution and future plans.

Abstract publication:

I agree that the abstract will be published on the web site

Remote user interfaces / 136

Web Applications to Support X-ray Spectroscopy Experiments

Author: Jacob Filik[1]

Co-authors: Christopher Reynolds [1]; Giannantonio Cibin [1]; Joshua Elliot [1]; Sofia Diaz-Moreno [1]; Thomas Penfold [2]; Victor Rogalev [1]

[1] *Diamond Light Source*

[2] *Newcastle University*

Using web applications in a Software As A Service approach is increasingly becoming an important route for science facilities to provide tools to their users.
Diamond has developed two such applications; the XAS data repository and web-CONEXS.
The XAS data repository is a database of XAS data collected on standard or well characterized compounds, facilitating the storage and retrieval of this data, both through the user interface and a REST API.
Web-conex is a simulation as a service portal, developed as part of the EPSRC funded COllaborative NEtwork for X-ray Spectroscopy (CONEXS) and offers a simplified interface for launching simulations on the STFC's IRIS HPC infrastructure.
This presentation will give a brief overview of these two applications and then focus common issues, such as web technologies, authentication and authorization, and deployment and how they are addressed at Diamond.

Abstract publication:

I agree that the abstract will be published on the web site

Remote user interfaces / 113

Evolution management of CERN's Accelerator Controls web applications with Micro Frontends

Author: Anti Asko[1]

Co-authors: Chris Roderick [1]; Stephane Deghaye [1]

[1] *CERN*

After implementing a new Micro Frontend architecture two years ago, CERN's accelerator Controls group has acquired substantial expertise across various parts of the field. With an emphasis on agile development practices, Micro Frontends have facilitated shorter migration cycles, faster deliveries, and fostered a transformation in the development culture of web applications for accelerator Controls at CERN. This paper will introduce the new generic architecture as well as challenges faced and how they were tackled. Additionally, this paper will present the new workflows adopted and highlight collaboration and contribution schemes established with external teams. Micro Frontends

introduces a new era in web application development, presenting both exciting opportunities and challenges.

Abstract publication:

I agree that the abstract will be published on the web site

Remote user interfaces / **40**

RICA : A system to enhance remote user experience using secure real-time video-streaming over the web.

Author: Cristina Cocho Martinez[None]

Historically, experiments were carried out by scientists physically present on the instrument where their sample was being measured. Recent socio-economic changes have increased the need for tools and mechanisms providing remote secured access to control the instruments and to access the generated scientific data.

Amongst those tools, a real time visualization of the instrument, including all its moving parts, plays an important role because it allows scientists to have a better vision of their experiment and to anticipate any potential problematic situation.

At the Institute Laue Langevin we have developed RICA (Remote Instrument Camera Access) whose aim is to provide a real-time secured remote video access to the instruments. RICA provides a mechanism allowing only authenticated and authorized users to access video streams remotely via their browsers. Access-control linked to the experiment schedules of the ILL and the Nomad instrument control software ensures that only team members of an active experiment can access the video streams.
This article describes the technical challenges to convert in real-time a standard internet-enabled camera stream (Real Time Streaming Protocol) into one that can be sent via the web to a browser. A detailed description of the project architecture will be provided as well as a view of the containerization and continuous integration techniques used for the deployment of the application.

Abstract publication:

I agree that the abstract will be published on the web site

Remote user interfaces / **120**

Hey SlackBot, What's Up With my Beamline?

Authors: Abigail Alexander[1]; Dean Keeble[1]

[1] *Diamond Light Source*

The well-known application *Slack* is primarily used for instant messaging and sharing memes. However, according to the people who *make* it, Slack is "… a messaging app for business that connects people to the information that they need".

For instrument staff at a scientific facility that information might be the current state of hardware; statuses of various data acquisition services; or outputs from data processing. Furthermore, having access to this information on desktop and/or mobile clients is desirable for instrument staff who may be in the lab; at their desk; or off-site, away from the facility.

At Diamond, we have recently developed and deployed a Slack app which aims to provide some on-demand high-level information to beamline staff. This data is sourced from various services and attempts to connect beamline staff with the information they need (or in some cases, *want*).

Here we present some of the design architecture we have chosen to adopt, how we have implemented this functionality; and present plenty of examples of the kind of information that we have made accessible.

Abstract publication:

I agree that the abstract will be published on the web site

Workflow engines / **117**

Automated data processing with Ewoks for ESRF beamlines and users

Authors: Andrew Goetz[1]; Loic Huder[1]; Wout De Nolf[1]

[1] *ESRF*

The increasing complexity and speed of experiments at synchrotrons call for the need of efficient and well-established automated data processing pipelines. Workflows are the ideal approach for defining these pipelines given their ability to describe data processing recipes. This led to the emergence of many workflow systems, such as tomwer at tomography beamlines and pypushflow at MX beamlines at ESRF.

Ewoks aims at unifying these workflow systems via a “meta workflow” approach. By decoupling the actual recipe (the workflow) from the workflow system, we can ensure that the processing remains independent from the underlying technologies and can be replicated in other contexts or at later times.

In this presentation, I will show how we have used this “meta workflow” approach to provide ESRF beamlines, but also ESRF users, with consistent automated data processing, in line with FAIR data principles.

Abstract publication:

I agree that the abstract will be published on the web site

Workflow engines / **14**

Streamlining Scientific Discovery with Data Pipelines at the Advanced Photon Source

Authors: Hannah Parraga[None]; John Hammonds[None]; Steven Henke[None]; Siniša Veseli[None]; Ryan Chard[None]; Nicholas Schwarz[None]

Data are essential to the scientific discoveries enabled by experiments performed at the APS. When the facility resumes operation this year, it will generate an estimated 100PB of raw experimental data per year from its seventy-two 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. The amount of data generated at the APS will increase due to the newly upgraded storage ring and 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. Each beamline differs and can have multiple experiment techniques, types of detectors, data rates, data formats, operating systems, and processing workflows. Additionally, the users themselves will also vary. They come from different companies, universities, and research institutions, but all must be able to
access their data after leaving the lab. 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 by beamline staff. Beamline staff have different levels of technical experience. Some desire a hands-off approach to data management and some want the flexibility to program their own custom
tools. They use computers with a variety of operating systems. The APS must have a data management solution which works for each of these unique beamlines.

To address these challenges, workflows have been deployed at several beamlines which automate the data lifecycle from detector to data portal. These workflows integrate the Bluesky controls software, the
APS Data Management System, and Globus services to provide infrastructure which is agnostic to analysis technique, compute resource, and storage location. The reusability and flexibility of the configuration driven approach allows new analysis techniques to be supported with minimal development effort. Common analysis packages are supported which are used across multiple beamlines for X-ray techniques such as ptychography, crystallography, X-ray fluorescence microscopy, tomography, X-ray photon correlation spectroscopy, and far-field high-energy diffraction microscopy, and Bragg coherent diffractive imaging. Raw data, metadata, and analysis results are secured to only users from a given experiment. Data portals provide functionality to reprocess datasets using different parameters.

Although these data pipelines are meeting many of the challenges, further development is under way of additional features to provide users with an exceptional data management experience. Additional analysis packages are being implemented for coherent surface scattering imaging, crystallography, and combined ptychography plus X-ray fluorescence microscopy. Furthermore, additional beamlines are in the process of being onboarded to make use of the pipelines which have already been developed. Looking to the future, the existing data portal reprocessing features will be expanded to include data management tasks from any stage in the data lifecycle.

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

Abstract publication:

I agree that the abstract will be published on the web site

Workflow engines / 132

Tools for predicting and responding to anomalies in experiment data streams

Authors: Justin Wozniak[1]; Nicholas Schwarz[1]; Michael Prince[1]; Tong Shu[2]; Bogdan Nicolae[1]

[1] *Argonne National Laboratory*

[2] *University of North Texas*

Building resilient data streams for large-scale experiments is a critical problem in modern settings in which advanced computing techniques are more tightly integrated with data collection activities.

Resilience-aware application solutions must include 1) policy management, in which science-level goals are presented to the system; 2) data movement telemetry, which captures system responses; and 3) systems-level predictive and adaptive anomaly mitigation, which integrates telemetry streams and provides actionable predictions to the policy manager. In this effort, we consider a previously-developed workflow application in which Advanced Photon Source (APS) data is produced by the detector and is automatically picked up by the APS Data Management system which uploads it to APS central storage, then transfers the data to an HPC storage system at the Argonne Leadership Computing Facility, and triggers a job on supercomputer Polaris which reconstructs the scan. The goal is to apply models to simple anomaly prediction problems: for example, by applying machine learning (ML) methods to event streams to predict network degradation or reductions in available compute resources, and feed these predictions back to the policy component for application-level decisions. We can then link multiple such models (network, compute, applications, observation) into virtual infrastructure twins: i.e., predictive networks that approximate some aspect(s) of the real infrastructure and its applications. At the policy level, the system will then be able to generate alerts about imminent system congestion dynamics and fault-triggering conditions. As a tool, we want the system to be able to integrate interfaces that allow the use of a simple declarative notation to specify the stream(s) to be monitored and the model method(s) to be applied. This presentation will cover preliminary work in this effort, describe the raw data sources that can be integrated into telemetry streams for prediction, the selection and refinement of ML models for the task, and the quality of preliminary predictions from the system.

Abstract publication:

I agree that the abstract will be published on the web site

Workflow engines / **57**

Expanding the Galaxy platform to materials science workflows at STFC

Authors: Leandro Liborio[1]; Patrick Austin[2]; Subindev Devadasan[2]

Co-authors: Abraham Nieva de la Hidalga [3]; Alexander Belozerov [2]; Tom Underwood [2]

[1] *Scientific Computing Department, STFC, UKRI, UK*

[2] *Science and Technology Facilities Council (STFC, UKRI)*

[3] *Cardiff University*

With FAIR principles increasing in importance within many fields, the challenge of ensuring that these principles are fully embedded in research outputs applies not just to the (meta)data itself, but also to the methods used to process and analyse it. If metadata associated with the raw data is lost in the analysis process, the Findability (which relies on this metadata) may be compromised. If the parameters used in the analysis of the data, and the information about data provenance, are not recorded, then Reusability can be seriously compromised.

To address these challenges, at the Science and Technologies Facilities Council (STFC) in the UK, we have begun using the open source Galaxy web platform for workflow management. Galaxy offers a number of features that directly relate to ensuring that outputs retain all metadata needed for them to be reproduced: histories store the data and parameter inputs associated with all output data, software tools are strictly versioned and run in containers, and executions of workflows can be exported as Research Object Crates. Beyond this it also offers advantages, such as a supportive community of developers and users, a platform for hosting associated training materials and a system for job execution which supports multiple scalable options for compute resources whilst not requiring users to have any knowledge of the underlying submission system.

In particular, we have been developing Galaxy tool interfaces for the Larch library of tools for analysis of X-ray Absorption Spectroscopy (XAS) data, which enable the creation of workflows for performing near edge and extended fine structure analysis of data gathered at the Diamond Light Source in the UK. To direct development and provide a trial use case, in collaboration with the UK Catalysis Hub, we have attempted to reproduce results from papers published in this field. This example has helped us expose the challenges of reusing data when a system such as Galaxy is not employed to standardise analysis and capture all the required parameters. We are aiming to lower the barrier to entry for this kind of workflow to encourage more users to do it for their own data, where again the Galaxy platform can offer an advantage.

In this talk we will present the results of this effort, aiming to show the work required to bring Galaxy to a new domain and what the challenges and benefits of this are.

Abstract publication:

I agree that the abstract will be published on the web site

Research Software Engineering / 73

Python as a Statically Typed Language

Author: Callum Forrester[None]

Since the introduction of type annotations in python3.x and the arrival of static analysis packages like mypy and pyright, more and more Python code is statically typed. Type systems work by providing compile-time error checking and a fast feedback loop for particular types of error at the expense of increasing verbosity and reducing the overall number of possible programs a language can produce to a type-safe subset. For these reasons there has been some resistance to adopting typing in Python due to its historical position as a very simple, versatile language and easy in-route for novices and dabblers. However at Diamond we feel there is a good case for adopting it everywhere and treating Python as a statically typed language, especially for facility-scale software developed in close collaboration with users.

Abstract publication:

I agree that the abstract will be published on the web site

Research Software Engineering / 7

Performance Benchmarking of Node.js, Rust, Go, and Python in Web Applications

Author: Igor Khokhriakov[1]

[1] *DESY, FS-SC*

In the rapidly evolving landscape of web development, the performance of backend technologies is a critical factor influencing scalability, efficiency, and user experience. This research aims to present a comprehensive performance comparison of Node.js, Rust, Go, and Python —four prominent technologies widely adopted in web application development. Through a series of systematic tests, we evaluated each technology's capability to handle concurrent connections, process I/O operations, and manage CPU-intensive tasks within a web server context.

Our methodology involved setting up a standardized testing environment for each technology, focusing on key performance metrics such as requests per second (RPS), latency, and CPU/memory utilization under varying loads. The tests were designed to simulate real-world scenarios, including serving static content, executing database operations, and performing computational tasks.

Node.js, with its non-blocking I/O model, demonstrated excellent performance in handling I/O-bound operations, particularly in scenarios with high concurrency levels. However, its single-threaded nature posed limitations in CPU-bound tasks, despite the potential for scalability offered by its cluster module.

Rust, known for its speed and memory safety, excelled in CPU-intensive tests, showcasing its ability to leverage system resources efficiently. Its asynchronous runtime further enabled impressive handling of I/O-bound operations, positioning it as a robust choice for high-performance web applications.

Go's simplicity and built-in concurrency model, based on goroutines and channels, allowed it to perform remarkably well across all tests. It balanced CPU and I/O operations adeptly, making it a versatile option for a wide range of web applications.

Python, while not matching the performance of the other three technologies in raw throughput and latency, stood out for its developer productivity and vast ecosystem. Its asynchronous frameworks, such as asyncio, provided significant performance improvements for I/O-bound tasks, although it lagged in CPU-bound processing.

In conclusion, our research underscores the importance of choosing the right technology based on the specific needs of a web application. While Rust and Go offer superior performance for CPU-intensive and high-concurrency applications, Node.js remains a strong contender for I/O-bound scenarios. Python, with its ease of use and extensive libraries, is well-suited for projects where development speed is paramount. This benchmarking exercise serves as a guide for developers and architects in making informed decisions, balancing performance with other factors such as development efficiency, ecosystem maturity, and maintainability.

Abstract publication:

I agree that the abstract will be published on the web site

Research Software Engineering / 149

Preferred Practices Through a Project Template

Author: Peter Peterson[1]

Co-authors: Chen Zhang [1]; Jose Borreguero-Calvo [1]; Kevin Tactac [1]

[1] *Oak Ridge National Laboratory*

Software engineers, including those involved in scientific software, often mention that they follow best practices. While sounding like an excellent idea, this is often near impossible. Frequently, there is an opinion on what is the best practice. Some other things, like software licenses and naming conventions, are mostly left to the development team to decide. Our software team has developed preferred practices which are guided by best practices rather than constrained by them. These practices reflect the values and realities of the environments and users of our software. To share these decisions with the team, we created a project template. The project template promotes standardization and enables discussion of how various decisions are implemented. This paper will discuss what information belongs in a project template and the benefits of collecting preferred practices in this way.

Abstract publication:

I agree that the abstract will be published on the web site

Satellite Meeting / 24

LENS satellite meeting on analysis software

ILL50 110

Satellite Meeting / 163

Lima & fast 2D Detector DAQ

Satellite Meeting / 26

Mantid tutorials

MD-1-21

Satellite Meeting / 137

3rd Taranta Developers Face2Face

Satellite Meeting / 22

NIAC meeting

ESRF - CTRM 206
Dates: 27/09/24 to 30/09/24

Satellite Meeting / 23

Metadata catalogues

ESRF LOB 2-03

Satellite Meeting / 162

An introduction to developing in the Karabo SCADA Framework

BEL 1-01

Satellite Meeting / 161

Bluesky Community Meeting (Data Access)

LOB 1-45

Logistics & Welcome to NOBUGS / 3

Day 4

Logistics & Welcome to NOBUGS / 2

Day 3

Satellite Meeting / 25

An introduction to developing in the Karabo SCADA Framework

ESRF BEL-1-01

Posters / 12

"Intelligence Terminal"Multimodal Data Analysis System for Synchrotron Radiation Experiments

Author: Lina Zhao[1]

[1] *The institute of High Energy Physics CAS*

Synchrotron radiation (SR) light sources provide precise and deep insights that have been driving cutting-edge scientific research. Facing to SR scientific big data challenge, it is urgent to develop artificial intelligence (AI) analysis methods to enhance research efficiency including novel material discovery1. In this talk, I will focus on the construction of "Intelligence Terminal"multimodal data analysis system including AI analysis methods for image and diffraction SR data. First, regarding image data, we implement a novel localization quantitative analysis method based on deep learning to analyze X-ray nano-computed tomography (Nano-CT). We achieve localization three-dimensional quantitative Nano-CT imaging analysis of single-cell HfO2 nanoparticles and demonstrate the notable effect of the nanoparticles in tumor treatment2. Our approaches show the potential to explore the localization quantitative three-dimensional distribution information of specific molecules at the nanoscale level in Nano-CT. Second, regarding diffraction data, we develop two sets of data-driven and physics-knowledge-driven machine learning (ML) methods to analyze the X-ray diffraction and extract three-dimensional orientation information of nanofibers. The data-driven ML model achieves high accuracy and fast analysis of experimental data and is available to be applied in multi light

sources and beamlines3. The physics-knowledge-driven ML method enables high-precision, self-supervised, interpretable analysis and lays the foundation for systematic knowledge-driven online scientific big data analysis. Then, we develop the using interface named as intelligence photon source brain (IPSBrain, www.ipsbrain.com) based on the "Intelligence Terminal"system for users to utilize the novel AI driven data analysis methods. Overall, our work aims to analyze multimodal SR data accurately and quickly in real-time through AI algorithms, which support AI for SR-based Science strongly.

Reference:
1 Qingmeng Li, Rongchang Xing, Linshan Li, Haodong Yao, Liyuan Wu, Lina Zhao. *Synchrotron radiation data-driven artificial intelligence approaches in materials discovery. Artificial Intelligence Chemistry, 2(1): 2949-7477, (2024).*
2 *Zuoxin Xi, Haodong Yao, Tingfeng Zhang, Zongyi Su, Bing Wang*, Weiyue Feng, Qiumei Pu, Lina Zhao. *Quantitative Three-Dimensional Imaging Analysis of HfO2 NPs in Single Cells Via Deep Learning aided Nano-CT. ACS Nano, revision returned, (2024).*
3 *Minghui Sun, Zheng Dong, Liyuan Wu, Haodong Yao, Wenchao Niu, Deting Xu, Ping Chen, Himadri S Gupta, Yi Zhang*, Yuhui Dong, Chunying Chen, Lina Zhao*. Fast extraction of three-dimensional nanofiber orientation from WAXD patterns using machine learning, IUCrJ, 10, 3 (2023).

Abstract publication:

I agree that the abstract will be published on the web site