

MS01-P01 | DATA ANALYSIS INFRASTRUCTURE FOR SERIAL CRYSTALLOGRAPHY

EXPERIMENTS AT THE EUXFEL

Dall'Antonia, Fabio; Beg, Marijan; Bergemann, Martin; Bielecki, Johan; Bondar, Valerii; Carinan, Cammille; Costa Junior, Raul; Danilevski, Cyril; Ehsan, Wajid; Esenov, Sergey; Fabbri, Riccardo; Flucke, Gero; Fullà-Marsà, Daniel; Giovanetti, Gabriele; Göries, Dennis; Hickin, David; Jarosiewicz, Tobiasz; Kamil, Ebad; Kirienko, Yury; Kirkwood, Henry; Klimovskaia, Anna; Kluyver, Thomas; Mamchyk, Denys; Michelat, Thomas; Mohacsi, Istvan; Parenti, Andrea; Rosca, Robert; Rück, Denivy; Santos, Hugo; Schaffer, Robert; Silenzi, Alessandro; Spirzewski, Michal; Trojanowski, Sebastian; Youngman, Christopher; Zhu, Jun; Mancuso, Adrian P.; Fangohr, Hans; Brockhauser, Sandor: European XFEL GmbH, Schenefeld, GER

Data analysis applications for serial crystallography and single-particle imaging experiments at the European X-ray Free Electron Laser (EuXFEL) instrument SPB/SFX [1] range from near real-time data processing for experimental optimisation to computationally intensive offline analysis on stored data long after the experiment. Online analysis tasks for providing fast feedback during experiments are integrated into the EuXFEL-distributed control software Karabo [2]. Alternatively, user groups can connect their own processing tools to the Karabo-Bridge, a client software that provides a network interface to the data streams. Currently a C++ and a Python API are available.

For offline analysis, we provide the Karabo-Data Python package [3] to help users working with the experimental data stored at EuXFEL. It provides routine analysis capabilities and data file conversion to be compatible with the most common analysis tools applied within the serial crystallography community.

Another field of active development is the implementation and integration of semi-automated high-performance computing crystallography pipelines to support user groups with computationally intensive offline data analysis tasks. We also aim at providing a full set of standard data analysis “sandbox” tools with example datasets and Jupyter notebooks to combine documentation and tutorials for less experienced users.

We present an overview of the available tools and interfaces with a selection of use-cases for online and offline data processing.

[1] A. P. Mancuso *et al.* (2019), *J. Synchrotron Rad.* 26:3, 1-17.

[2] B. Heisen *et al.* (2013), *Proc. ICALEPCS2013*, 1465-8.

[3] H. Fangohr *et al.* (2018), *Proc. ICALEPCS2017*, 245-52.