O 63: Poster: Data Management

Time: Wednesday 18:00–20:00 Location: P2/EG

O 63.1 Wed 18:00 P2/EG

SciCat - a meta data catalog and research data management system — •Linus Pithan[1], Massimiliano Novelli[2], Dylan McReynolds[3], Laura Shemilt[4], Carlo Minotti[5], Anastasiia Pylypenko[1], Alexander Gerlach[1], Alexander Hinderhofer[1], Stephan Egli[5], Tobias Richter[2], and Frank Schreiber[1] — [1]Uni Tübingen, DAPHNE (DE) — [2]European Spallation Source (DK/SE) — [3]Lawrence Berkeley National Laboratory (USA) (US) — [4]The Rosalind Franklin Institute (UK) — [5]Paul Scherrer Institut (CH)

SciCat is a modern and flexible data catalogue that facilitates research data management on multiple scales and thereby suits the needs of large scale research facilities, individual research groups (e.g. at universities) as well as scientific communities. SciCat [1] was initially built to serve Photon- and Neutron sources (in the beginning developed at PSI and ESS, later suppored by ExPaNDS & PaNOSC and further institutions). Through the engagement of DAPHNE4NFDI (DAta from PHoton and Neutron Experiments) in the SciCat project there are new use cases emerging especially - but not exclusively - for small scale installations. The use of an homogenized interface (API) to create and search datasets SciCat can e.g. also be used as infrastructure backbone for machine learning (ML) projects in communities by serving datasets for validation, testing and training of ML models. On this poster we aim to give an overview on specific use-cases of SciCat within DAPHNE, which may serve as blueprints to use SciCat also in other scientific communities.

[1] http://scicatproject.github.io, http://github.com/SciCatProject

O 63.2 Wed 18:00 P2/EG

A local solution for automated data acquisition and storage in catalysis — •Abdulrhman Moshantaf[1], Mike Wesemann[1], Patrick Oppermann[1], Heinz Junkes[1], Robert Schlögl[1,2], and Annette Trunschke[1] — [1]Fritz-Haber-Institut der Max-Planck-Gesellschaft, Berlin (Germany) — [2]Max Planck Institute for Chemical Energy Conversion, Mühlheim 45470 (Germany)

In order to solve the current challenge in catalysis research in the development of new, scalable catalysts for hydrogen-based future technologies, a better integration of theory and experiment is required. The necessary data exchange demands extensive digitalization in catalysis. Experimental data must be generated reproducibly and with sufficient diversity, and must be available in machine-readable form. Artificial intelligence can then contribute to the discovery of correlations. We have developed a concept for a local data infrastructure and implemented it in a catalysis laboratory. In research projects, handbooks are written (preferably in machine-readable form) detailing how experimental data are obtained, including the definition of benchmark catalysts. To implement the concept of handbooks, automated systems for data acquisition and storage have been designed. Such a system consists of (i) EPICS for communication with devices and data acquisition, (ii) a database (archive), (iii) an archiving appliance for storing time series, (iv) Phoebus for creating graphical user interfaces, (v) Python/Bluesky/Jupyter notebooks for creating automations and evaluations, and (vi) S3 storage for long-term storage.