DAPHNE4NFDI Annual Meeting 2025

Report of Contributions

Contribution ID: **60** Type: **Use case flash talk (Mon)**

UC1: X-ray imaging

LMU and Uni Göttingen

Session Classification: Use-cases (flash talks)

Contribution ID: **70** Type: **Use case flash talk (Mon)**

UC11: High energy x-ray diffraction

DESY, CAU Kiel and HZDR

Session Classification: Use-cases (flash talks)

Contribution ID: **75** Type: **Talk**

Welcoming speech HZB

Wednesday 26 March 2025 08:45 (15 minutes)

Presenter: RECH, Bernd (HZB)

Session Classification: Facility Updates

Contribution ID: **78** Type: **Facility update (Wed)**

DESY

Wednesday 26 March 2025 09:00 (20 minutes)

Presenter: MURPHY, Bridget

Session Classification: Facility Updates

Contribution ID: **88** Type: **Poster**

Pydidas - A modular framework for diffraction data analysis: The road towards FAIR software

Helmholtz-Zentrum Hereon operates multiple X-ray diffraction (XRD) experiments for external users and while the experiments are very similar, their analysis is not. The variety in data analysis workflows is challenging for creating FAIR analysis workflows because a lot of the analysis is traditionally done with small scripts and not necessarily easily reproducible.
Pydidas [1, 2] 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. 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.
Plugins are straightforward in their 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://pydidas.hereon.de
[2] https://github.com/hereon-GEMS/pydidas

Primary author: STORM, Malte (Helmholtz-Zentrum Hereon)

Presenter: STORM, Malte (Helmholtz-Zentrum Hereon)

Session Classification: Poster

Contribution ID: **91** Type: **Facility update (Wed)**

HEREON

Session Classification: Facility Updates

Contribution ID: **92** Type: **Use case flash talk (Mon)**

Flash talk instructions

Monday 24 March 2025 14:35 (5 minutes)

Presenter: Dr TRAGESER, Christian (JCNS)

Session Classification: Use-cases (flash talks)

Contribution ID: **93** Type: **Poster**

Performance Benchmarking: one of the System Design's pillars

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.

Primary author: KHOKHRIAKOV, Igor (DESY)

Presenter: KHOKHRIAKOV, Igor (DESY)

Session Classification: Poster

Contribution ID: **94** Type: **Talk**

snip –user centric collaborative digital lab book

Tuesday 25 March 2025 09:15 (15 minutes)

We present current updates on the development of snip –the user friendly collaborative lab book to document your current thoughts about the experiment online and live.
Recent improvements include:

- Hyperlinks to other systems,
- Hyperlinks to specific snips,
- Single Sign On / OpenID Connect,
- Progressive Web App for mobile devices,
- Python library.

Primary authors: Dr OSTERHOFF, Markus (Röntgenphysik Göttingen); KÖSTER, Sarah; MOHR, Sebastian

Presenter: Dr OSTERHOFF, Markus (Röntgenphysik Göttingen)

Session Classification: Metadata, Catalogues and Software (TA1, TA2 and TA3)

Contribution ID: **95** Type: **Poster**

Evaluating ELNs for experiment documentation and how to make an informed decision

Electronic laboratory notebooks (ELNs) are essential for fair and open research, contributing to both the reproducibility and reusability of data and ensuring that research is easily accessible and retrievable. Here we dive into the importance of ELNs for large-scale photon and neutron research infrastructures and outline key specifications for successful ELN implementation. These may be features such as real-time collaboration, data integration and secure access control. Its necessary to introduce a framework for evaluating existing ELNs based on these specifications, including a formula for calculating a figure of merit to aid in the selection process. A survey of existing ELNs revealed that while many solutions are technically mature, only a smaller percentage fully meet the specific requirements of this research domain.

Primary author: JORDT, Philipp (Kiel University)

Co-authors: Dr HAKIM, Bishoy (FAU); MURPHY, Bridget; DALL'ANTONIA, Fabio (European XFEL GmbH); MAURER, Florian; WEBER, Frank (Karlsruhe Institute of Technology); GRUNWALDT, Jan-Dierk; PITHAN, Linus (DESY, FS-EC); AMELUNG, Lisa (DAPHNE4NFDI | DESY); Dr OSTERHOFF, Markus (Röntgenphysik Göttingen); DOLCET, Paolo (Karlsruhe Institute of Technology); Mr BINIYAMINOV, Vitaly (KIT / ITCP); LOHSTROH, Wiebke; TYMOSHENKO, Yuliia (KIT, IQMT)

Presenter: MURPHY, Bridget

Session Classification: Poster

Contribution ID: **96** Type: **Poster**

Spectroscopy Use Case: Advancing Towards FAIR Data

Inelastic scattering is a fundamental technique for probing lattice dynamics and magnetic excitations, serving as a cornerstone in materials science and condensed matter research. While it has historically been linked to neutron scattering, recent innovations in modern synchrotron facilities have facilitated the acquisition of complementary x-ray scattering data, which are vital for contemporary scientific investigations. This encompasses experiments conducted under high-pressure conditions on small single crystals and within materials that demonstrate substantial neutron absorption.
In use case 5, we aim to establish suitable metadata vocabularies specifically for Inelastic Neutron Scattering (INS) and Inelastic X-ray Scattering (IXS) techniques, alongside fostering the use of electronic laboratory notebooks (ELNs). This poster demonstrates our efforts in developing metadata schemas for a triple-axis spectrometer and provides an update on the status of the electronic lab notebook being developed at MLZ.

Primary author: TYMOSHENKO, Yuliia (KIT, IQMT)

Co-authors: SCHNEIDEWIND, Astrid (FZ Jülich); BAUDISCH, Josef (MLZ TUM); LOHSTROH, Wiebke; WEBER, Frank (Karlsruhe Institute of Technology)

Presenter: TYMOSHENKO, Yuliia (KIT, IQMT)

Session Classification: Poster

Contribution ID: **98** Type: **Poster**

White Paper On Metadata Catalogue Systems

DAPHNE4NFDI, part of Germany's National Research Data Infrastructure (NFDI), focuses on implementing FAIR principles for research data from Photon and Neutron (PaN) sources at large-scale facilities, universities, and research institutions. This includes the adoption of SampleDB as a flexible metadata solution, with comparisons to alternatives like ICAT and SciCat. Metadata catalogues are categorized into raw data repositories, institutional databases, and public-access systems, each with distinct roles. Key considerations include the integration of standards like NeXus, user-friendly searchability, and metadata enrichment. Case studies, including RefXAS, illustrate practical implementations and quality assessments. The paper also discusses Authentication and Authorization Infrastructure (AAI) strategies for secure, role-based data access, offering recommendations to enhance collaboration, data quality, and a unified, FAIR-compliant metadata ecosystem for the PaN research community.

The objective of this **white paper Catalogue** is to pinpoint a catalogue solution suitable for DAPHNE4NFDI partners. It aims to identify catalogue use-cases and the expected nature of searches along with outlining the metadata schema. The contents of this document provide a preliminary yet concise specification of (non-)requirements, a review of the current state and local solution landscape, incorporating lessons learned and a discussion leading to preliminary conclusions and decisions.

Primary author: Dr HAKIM, Bishoy (FAU)

Co-authors: GAUR, Abhijeet (ITCP, KIT); PEDERSEN, Björn (MLZ, TU München, Garching); FELDER, Christian (Forschungszentrum Jülich GmbH); DALL'ANTONIA, Fabio (European XFEL GmbH); Mr DALLMANN, Johannes (FAU); Mr HAYEN, Nicolas (Kiel University); KWEE-HINZMANN, Regina; KRAHL, Rolf (Helmholtz-Zentrum Berlin für Materialien und Energie); BUSCH, Sebastian (GEMS at MLZ, Helmholtz-Zentrum Hereon, Germany); Mr PARIPSA, Sebastian (University of Wuppertal); UNRUH, Tobias (FAU)

Presenter: Dr HAKIM, Bishoy (FAU)

Session Classification: Poster

Contribution ID: **99** Type: **Poster**

Continuous Integration and Continuos Deployment at DESY

Continuous Integration and **Continuous Deployment** is a modern **Software Engineering** best practice that enables efficient *large scale* software development and use. There are a variety of popular Ci/CD *tools* that help in adopting these practices. In this poster we focus on the kinds of software, their runtime environments,packaging and deployment tools and techniques used at DESY that can easily be leveraged by participating institutions under DAPHNE4NFDI.

Primary authors: KHOKHRIAKOV, Igor (DESY); TIRUMALAI NALLAM CHAKRAVARTY, Parthasarathy (DESY); KWEE-HINZMANN, Regina

Presenter: TIRUMALAI NALLAM CHAKRAVARTY, Parthasarathy (DESY)

Session Classification: Poster

Contribution ID: **100** Type: **Poster**

MIEZEPY: An Open-Source Tool for Efficient MIEZE Data Reduction

MIEZEPY is an open-source software package designed for the efficient reduction of data acquired in the MIEZE (Modulation of Intensity with Zero Effort) mode. MIEZE is a neutron resonant spin echo technique that enables the measurement of the intermediate scattering function, S(Q, τ), in depolarizing sample environments, such as under high magnetic fields. This technique is implemented at the RESEDA (Resonance Spin Echo for Diverse Applications) spectrometer at MLZ, which offers sub-μeV energy resolution and an exceptional dynamic range (~8 orders of magnitude). As part of its commitment to making data FAIR, MLZ is preparing to switch to the NeXus data format in the near future. Initial steps towards implementing the new data format include drafting a NeXus template for raw data recording in NICOS and enabling NeXus file readout in MIEZEPY for subsequent data reduction and analysis. Further progress in this transition is expected in the coming months.

Primary author: LYPOVA, Iryna (TUM/FRM II)

Co-authors: SCHOBER, Alexander; HERB, Christoph; BEDDRICH, Lukas; JOCHUM, Johanna; FRANZ, Christian; PFLEIDERER, Christian; LOHSTROH, Wiebke

Presenter: LYPOVA, Iryna (TUM/FRM II)

Session Classification: Poster

Contribution ID: **101** Type: **Poster**

Persistent Sample Identification with an IGSN Service at Kiel University

The capture and curation of sample metadata is a crucial element of FAIR data collection, ensuring a consistent record of investigated physical systems. This necessitates the presence of a persistent sample identifier to link sample information with experimental data and electronic laboratory notebooks. The International Generic Sample Number (IGSN) is one such identifier, which is issued by DataCite and adheres to a common metadata standard with general Digital Object Identifiers (DOIs). The ability for strong integration with the DOI ecosystem makes IGSNs an excellent choice for use in DAPHNE4NFDI.

Kiel University hosts a service for registration of IGSN identifiers at the university level. It enables the integration of a content review process in the workflow and generates the necessary landing pages for URL resolution of the persistent identifiers. A dedicated DAPHNE4NFDI repository has been established and registrations from a number of members are already included. Former restrictions on the use by other DAPHNE4NFDI partners have recently loosened, and a request form and template for an IGSN registration at Kiel University are available.

We will discuss the current status and best practices for IGSN registration with this service, and give advice on how to get started with IGSN registrations in an individual capacity to adhere to the FAIR principles in publications.

Primary authors: HAYEN, Nicolas (Kiel University); PETERSEN, Thorge (Kiel University Computing Center)

Co-authors: Dr ASHTIANI, Ali (Kiel University); Dr MURPHY, Bridget (Kiel University); Dr KOBUS, Julia (Kiel University); Dr JORDT, Philipp (Kiel University); HÖVELMANN, Svenja (Kiel University); PAUL-STÜVE, Thilo (Kiel University Computing Center)

Presenter: HAYEN, Nicolas (Kiel University)

Session Classification: Poster

Contribution ID: **102** Type: **Poster**

Standardized Metadata Provision in the Communication Protocol SECoP

The Sample Environment Communication Protocol (SECoP) provides a generalized way for controlling measurement equipment –with a special focus on sample environment (SE) equipment [1,2]. In addition, SECoP holds the possibility to transport SE metadata in a well-defined way.

SECoP is designed to be
- simple to use,
- inclusive concerning different control systems and control philosophies and
- self-explaining providing a machine readable description of all available data and metadata.

The recently completed HMC project SECoP@HMC [3] focused on the standardized provision of metadata for typical SE equipment at large scale facilities (photons, neutrons, high magnetic fields) and on standardized metadata storage. The fact that SECoP is self-explaining and machine-readable favours the automated interpretation of data and metadata. With the latest definition of SECoP, we were able to integrate the use of vocabularies or glossaries.

With the ongoing development of SECoP and the provision of several tools for its easy implementation, a complete standardized system for controlling SE equipment and collecting and saving SE metadata is available and usable for experimental control systems. This approach can be applied to other research areas as well.

[1] K. Kiefer, et al. (2020). An introduction to SECoP –the sample environment communication protocol. Journal of Neutron Research, 21(3-4), pp.181–195
[2] https://github.com/sampleenvironment/secop
[3] https://helmholtz-metadaten.de/de/inf-projects/secopathmc

Primary author: KIEFER, Klaus (Helmholtz-Zentrum Berlin)

Co-authors: ZAFT, Alexander (Forschungszentrum Jülich); PETTERSSON, Anders (European Spallation Source); KLEMKE, Bastian (Helmholtz-Zentrum Berlin); FAULHABER, Enrico (Technische Universität München / FRM II); BRANDL, Georg (Forschungszentrum Jülich); GUENTHER, Gerrit (Helmholtz-Zentrum Berlin); KOTANSKI, Jan (Deutsches Elektronen Synchrotron DESY); ROSSA, Lutz (Helmholtz-Zentrum Berlin); UHLARZ, Marc (Helmhotz-Zebtrum Dresden Rossendorf); ZOLLIKER, Markus (Paul Scherrer Institut); EKSTRÖM, Niklas (European Spallation Source); BRAUN, Peter (Helmholtz-Zentrum Berlin); KRACHT, Thorsten (Deutsches Elektronen Synchrotron DESY)

Presenter: KIEFER, Klaus (Helmholtz-Zentrum Berlin)

Session Classification: Poster

Contribution ID: **103** Type: **Poster**

An ontology for Photon and Neutron Experimental Techniques (PaNET)

In Germany alone, approximately 5500 scientists generate 28 petabytes of data each year using Photons and Neutrons Experimental Techniques (PaNET). Such a large volume of data needs to be catalogued in a meaningful way to ensure Findable, Accessible, Interoperable and Reusable (FAIR) data [1]. To adhere to the FAIR data principle, an ontology of Photon and Neutron (PaN) experimental techniques (PaNET) has been developed over the years. The ontology includes terminologies related to different PaNETs and groups them according to functional dependence, experimental probe, purpose, and experimental physical process. The goal of this ontology is to facilitate consistent PaNET terminologies by providing global persistent identifiers, community-agreed labels and synonyms, and human-readable definitions, annotations and references [2]. The current state of the developed ontology includes several tags, which can be found online [3]. Further contributions to expand the ontology are planned in this work. These contributions are expected to complement and further develop the current state of ontology.

[1] URL:https://doi.org/10.5281/zenodo.8040606
[2] URL:https://doi.org/10.5281/zenodo.4806026
[3] URL: https://expands-eu.github.io/ExPaNDS-experimental-techniques-ontology/webvowl/index.html

Primary authors: MAJUMDAR, Arnab (Helmholtz Zentrum hereon); BUSCH, Sebastian (GEMS at MLZ, Helmholtz-Zentrum Hereon, Germany)

Presenter: MAJUMDAR, Arnab (Helmholtz Zentrum hereon)

Session Classification: Poster

Contribution ID: **104** Type: **Poster**

MLZ-ELN as a technical solution for an electronic lab notebook

MLZ-ELN is now fully containerized and available as an easy-to-install instance running behind an nginx web server. It includes new features such as the Typesense search engine (https://typesense.org/), Centrifugo (https://centrifugal.dev/) as a powerful alternative to Python websockets, a newer Angular version (Angular v17) and SSO integration with Keycloak 25.
MLZ-ELN has been running 24/7 as a prototype at MLZ (https://mlz-garching.de/) since July 2024, connected to Virtual Panda and an X-ray residual stress diffractometer.
It is also integrated into the MLZ proposal system , with user authentication/authorization handled by the MLZ SSO service https://idm.mlz-garching.de.

Primary author: BAUDISCH, Josef (MLZ TUM)

Co-authors: PEDERSEN, Björn (MLZ, TU München, Garching); WANG, Wojian

Presenter: BAUDISCH, Josef (MLZ TUM)

Session Classification: Poster

Contribution ID: **105** Type: **Poster**

Activities of the HZB Research Data Management Team

The Research Data Management (Forschungsdatenmanagement, FDM) team at HZB is currently involved in a wide variety of projects to benefit both our own institute and the wider PaN community. We are active within both the DAPHNE4NFDI consortium and the ICAT Project, and are a cornerstone of the Helmholtz ROCK-IT project. We are developing our own software solutions – such as the NeXusCreator file convertor and the SEPIA sample database –as well as contributing to the community development of catalysis ontologies and NeXus definitions. We are embracing the use and recognition of Peristent Identifiers such as DOIs, and are driving progress towards high-level sample classification. We are responsible for data publications and for providing elog and electronic lab notebook (ELN) functionality; these services are available to all researchers at our facility.

Primary authors: VELAZQUEZ SANCHEZ, Ana (Helmholtz-Zentrum Berlin für Materialien und Energie); Dr PEREZ PONCE, Hector (Helmholtz-Zentrum Berlin für Materialien und Energie); GÖRZIG, Heike (Helmholtz-Zentrum Berlin für Materialien und Energie); RIAL, Katherine (Helmholtz-Zentrum Berlin für Materialien und Energie); LEWERENZ, Marcus (Helmholtz-Zentrum Berlin für Materialien und Energie); KRAHL, Rolf (Helmholtz-Zentrum Berlin für Materialien und Energie)

Presenters: GÖRZIG, Heike (Helmholtz-Zentrum Berlin für Materialien und Energie); RIAL, Katherine (Helmholtz-Zentrum Berlin für Materialien und Energie)

Session Classification: Poster

Contribution ID: **106** Type: **Poster**

AIXtal –A Web and now Cross-platform Refinement Platform for First-Time Users

AIXtal is an accessible refinement platform designed specifically for first-time users. It includes a simplified user interface with reduced complexity, tutorials, tooltips and AI helpers for assistance. Currently, AIXtal supports GSAS-II and Fullprof (undergoing a refactoring) refinement of powder X-ray diffraction data.

The platform provides visualization of structural, experimental and refinement data, including graphical background selection. The user is supported by on-demand assistance on refinement parameters, either through short tooltips or an AI guide that semi-quantitatively assesses the last refinement result/parameter and recommends next refinement steps. For more in-depth investigation, the user can also access a compact view (web or PDF) of the refinement parameters, including esd's. If unfavorable refinement steps are indicated, the user can rely on the history function to jump back to any point in time. AIXtal also allows the user to export the refinement results in GPX format for further analysis with GSAS-II, if required. The platform AIXtal v1 with this functionality is in alpha state and was tested with students of chemistry lab course in April/May 2024.

AIXtal v2 is currently being developed based on the experience from AIXtal v1 and in order to broaden the use case scenarios. With multiple modes of operation, adding an expert mode with full refinement control to the simplified interface, it opens up for future developments of a more universal GUI for Rietveld programs also addressing other common, related crystallographic issues, e.g., a space group prediction. Furthermore, the web interface now runs on WebAssembly, which promises performance gains, and moreover now runs not only in the WebBrowser but also natively on Windows and Linux to allow a local installation as well. The GSAS-II implementation has also been extended to support refinement and visualization of multi-dimensional (angular and wavelength dispersive) data sets.

Current status, AI-helpers and features of AIXtal as well as its technical design principles are shown.

Primary author: NACHTIGALL, Noah

Co-author: HOUBEN, Andreas (RWTH Aachen)

Presenter: NACHTIGALL, Noah

Session Classification: Poster

Contribution ID: **107** Type: **Talk**

SciCat Integration at MLZ –Infrastructure and Live Demo

Tuesday 25 March 2025 10:00 (15 minutes)

This presentation provides an overview about the current status of SciCat at MLZ and the underlying infrastructure. Data acquisition and metadata capture are decoupled based on the RabbitMQ message broker. Information from various sources, such as the user office system, sample environment and the instrument are aggregated in the Networked Instrument COntrol System and transmitted in messages to our central Kubernetes cluster. Messages can be processed by secondary user services, e.g. the ingestor. The ingestor populates the SciCat catalogue with relevant metadata information from the experiment. A live demonstration using a virtual instrument will conclude the talk.

Primary authors: ZAFT, Alexander (Forschungszentrum Jülich); FELDER, Christian (Forschungszentrum Jülich GmbH)

Co-author: Dr TRAGESER, Christian (JCNS)

Presenters: ZAFT, Alexander (Forschungszentrum Jülich); FELDER, Christian (Forschungszentrum Jülich GmbH)

Session Classification: Metadata, Catalogues and Software (TA1, TA2 and TA3)

Contribution ID: **108** Type: **Poster**

X-ray Photon Correlation Spectroscopy (XPCS) at DESY

Efficient data management and analysis are essential for large-scale experiments at synchrotron facilities. At P10, DESY, the SciCat platform has been implemented to provide a FAIR (Findable, Accessible, Interoperable, Reusable) data management solution for SAXS/XPCS experiments. The successful installation of the SciCat ingestor and a test instance marks significant progress, with plans to integrate metadata ingestion, implement a comprehensive metadata schema, and deploy SciCat.
In parallel, the Xana 2.0 software development aims to modernize XPCS data analysis by introducing Python-based modules for standardized file formats like HDF5 and NeXus. This upgrade addresses the limitations of Xana 1.0, which relies on outdated libraries, by supporting robust metadata handling and parallelized I/O for large datasets. Together, SciCat and Xana 2.0 enhance data accessibility, management, and analysis, ensuring that P10 DESY remains at the forefront of experimental research while adhering to FAIR data principles.

Primary author: RAZA, Agha Mohammad (University of Siegen)

Co-authors: LEONAU, Aliaksandr; TOSSON, Amir; GUTT, Christian (Universität Siegen); UNGER, Frederik (University of Siegen); DARGASZ, Michelle (Uni Siegen); ANTHUPARAMBIL, Nimmi Das (University of Siegen); GAUTAM, Randeer Pratap (University of Siegen); Dr ÖZTÜRK, Özgül (University of Siegen)

Presenters: RAZA, Agha Mohammad (University of Siegen); LEONAU, Aliaksandr; TOSSON, Amir

Session Classification: Poster

Contribution ID: **109** Type: **Talk**

Towards AI-assisted high-throughput workflow for structural analysis of thin films

Tuesday 25 March 2025 10:30 (15 minutes)

The compositional optimization of new materials necessitates the high-throughput screening of a multitude of compositions, which must be investigated to elucidate the non-linear and non-monotonic structure-property-composition dependencies [1]. In this regard, data-driven material science enables researchers to accelerate the identification of new materials with desired properties for specific applications by efficiently exploring vast material spaces. Such high-throughput data-driven studies comprise two key elements: the combinatorial preparation of suitable sample libraries spanning wide compositional ranges, and the high-throughput screening of the structure and properties of the synthesized samples [2].
Surface-sensitive X-ray scattering methods available at modern X-ray sources, e.g. Grazing-Incidence Wide-Angle Scattering (GIWAXS), comprise a convenient tool for structural investigations of thin films with high screening rates and ultimate resolution. At the same time, the determination of the crystalline structure from the collected GIWAXS patterns represents a significant bottleneck of this approach as it is time- and resource-consuming, and frequently requires additional input from other methods. This makes it an ideal use case within TA3 in DAPHNE4NFDI.
In this work, we present a data analysis workflow for high-throughput and time-resolved *in situ* structural studies of thin films using GIWAXS. The workflow includes the following steps: data conversion and correction, metadata incorporation, Bragg peak detection, fitting and indexing [3-5]. We also present a universal data format used at all stages and a GUI software for convenient visualization of the final and intermediate results at any stage.

References:
1. A. Ludwig, npj Comput. Mater. 5 (2019), 70, https://doi.org/10.1038/s41524-019-0205-0
2. J. M. Gregoire, L. Zhou, J. A. Haber, Nat. Synth. 2 (2023), 493, https://doi.org/10.1038/s44160-023-00251-4
3. C. Völter et al., J. Appl. Cryst. (2025), in print
4. V. Starostin et al., Synchrotron Radiat. News 35 (2022), 21, https://doi.org/10.1080/08940886.2022.2112499
5. V. Starostin et al., npj Comput. Mater. 8 (2022), 101, https://doi.org/10.1038/s41524-022-00778-8

Primary authors: Dr LAPKIN, Dmitry (Universität Tübingen); Mr ABUKAEV, Ainur (Universität Tübingen); Mr VÖLTER, Constantin (Universität Tübingen); Mr ROMODIN, Mikhail; Mr HYLINSKY, Maik (Universität Tübingen); Mr MUNTEANU, Valentin (Universität Tübingen); Dr KNESCHAUREK, Ekaterina (Universität Tübingen); Dr STAROSTIN, Vladimir (Universität Tübingen); Dr HINDERHOFER, Alexander (Universität Tübingen); Prof. SCHREIBER, Frank (Universität Tübingen)

Presenter: Dr LAPKIN, Dmitry (Universität Tübingen)

Session Classification: Metadata, Catalogues and Software (TA1, TA2 and TA3)

Contribution ID: **110** Type: **Use case flash talk (Mon)**

UC8: Neutron TOF diffraction

Monday 24 March 2025 15:39 (7 minutes)

The university group at RWTH Aachen specializes in neutron TOF powder diffraction method development, primarily driven by the new concepts of the neutron time-of-flight diffractometer POWTEX, developed in collaboration with Forschungszentrum Jülich at FRM-II/MLZ in Garching. Unfortunately, no free neutrons were available in 2024. Within the DAPHNE project, we are therefore spreading our methods to other neutron TOF diffractometers while generally aiming to allow a broader, more sustainable applicability of the new developments. This overall goal splits into the following tasks, addressing the different steps in the workflow from raw data to scientific result: 1. multidimensional data reduction using Mantid, 2. derive and include fundamental instrument description in NeXuS data files, 3. multidimensional Rietveld test-cases, 4. AIXtal, a web and cross-platform Rietveld platform for (not only) first-time users, 5. AI tools for structure solution and profile refinement of powder diffraction data.

Utilizing our Mantid routine PowderReduceP2D for high-pressure data collected from SNAP@SNS, ORNL, USA combined with the derived and iterated fundamental instrumental parameters of SNAP allowed us to test-case the multidimensional Rietveld refinement on high-pressure neutron TOF powder diffraction data of a PbNCN sample. While the details were recently published in [1], it is important to note three things: At first, it is very tedious to collect instrumental parameters from various sources while they should actually be part of the data file. At second, a detector coverage of only 1.3 sr (SNAP) vs. ≈ 9 sr (POWTEX) already allows to do multidimensional Rietveld, which is remarkable and underlines the general applicability. While the data reduction steps, with the exception of the one- or multi-dimensional binning, were as similar as possible, the scientific result of the multi-dimensional refinement does differ significantly from the conventional, one-dimensional Rietveld refinement.

Based on these results, we created a sample nexus file containing data fields for the fundamental instrumental parameters. This was presented and discussed at a workshop with the SNS diffraction department which also allowed us to collect the view from the facility perspective. Shaping the future of the Rietveld diffraction software was also recognized as a common interest, especially since the existing software has been mostly around for a long time and the future of the method needs to be clarified.

AIXtal v1 was developed as prototypic, modern web-platform, which allows first-time users to utilize the Rietveld method while partly hiding the complexity of the existing tools GSAS-II and FullProf for the case of conventional X-ray diffraction. The web interface now runs on WebAssembly, which promises performance gains, and moreover now runs not only in the WebBrowser but also natively on Windows, Linux to allow a local installation as well. The higher-performance GUI shall allow to process multidimensional neutron data in the future. While multidimensional Rietveld (GSAS-II 2D) is available as worker already, we only recently started working on GUI and plotting features for this case. The integration of AI methods into AIXtal, e.g. to predict the space group symbol or to support the refinement process, shall ease the structure solution and refinement from powder dat, not only for the unexperienced user.

[1] Y. Meinerzhagen, K. Eickmeier, P.C. Müller, J. Hempelmann, A. Houben, R. Dronskowski, J. Appl. Crystallogr. 2024, 57, 1436–1445, doi:10.1107/s1600576724007635.

Primary author: HOUBEN, Andreas (RWTH Aachen)

Co-authors: NACHTIGALL, Noah; DAS, Shubhayu (RWTH Aachen University, Institute of Inorganic Chemistry); MEINERZHAGEN, Yannick (RWTH Aachen)

Presenter: HOUBEN, Andreas (RWTH Aachen)

Session Classification: Use-cases (flash talks)

Contribution ID: **111** Type: **Poster**

Everybody talks about research data management - But where is the science?

Newton, when seeing further, stood on the shoulder of giants,[1] not on the piles of his predecessors‘crufty research data. Nearly 300 years later, Richard Feynman coined the term “cargo cult science”for research that appears scientific but has no scholarly contribution nor impact.[2] With all the current hype about research data management, we tend to forget that science is about gaining knowledge, not accumulating research data to train large-language models and other AI tools. Data is not insight, but its prerequisite at best. And sharing data is a highly non-trivial concept resting on many implicit assumptions often not fulfilled.[3] We present a number of strategies and tools that, if used competently, will help us to improve the quality of our research and ultimately contribute to science and scholarship. First come tools for capturing all relevant information during data acquisition,[4] and electronic lab notebooks.[5] A framework for scientific data analysis that provides a gap-less and complete protocol of each step and relieves the user of actual programming [6] is a huge step forward. This is complemented by a larger (local) infrastructure consisting of persistent and unique IDs (PID, UID), a repository for “warm”research data, lab management, and knowledge base.[7] All these strategies and tools focus on the individual scientists, as only they can potentially ensure the urgently required quality of data and results that underpin scientific insight. Eventually, we need to teach [8,9] young researchers early on what science is all about and why properly handling research data is a prerequisite for scholarly contribution. “At stake is the future of scholarship.”[3]

References

[1] I. Newton, letter to Robert Hooke, February 5th, 1676

[2] R. P. Feynman, Cargo cult science, *Eng. Sci.* 1974, **37**(7), 10.

[3] C. L. Borgman, *Big Data, Little Data, No Data.* MIT Press, Cambridge MA 2015.

[4] B. Paulus, T. Biskup, Towards more reproducible and FAIRer research data: documenting provenance during data acquisition using the Infofile format, *Digit. Discov.* 2023, **2**, 234.

[5] M. Schröder, T. Biskup, LabInform ELN: A lightweight and flexible electronic laboratory notebook for academic research based on the open-source software DokuWiki, *ChemRxiv* 2023, doi:10.26434/chemrxiv-2023-2tvct.

[6] J. Popp, T. Biskup, ASpecD: A modular framework for the analysis of spectroscopic data focussing on reproducibility and good scientific practice, *Chem. Meth.* 2022, **2**, e202100097

[7] T. Biskup, LabInform: A modular laboratory information system built from open source components, *ChemRxiv* 2022, doi:10.26434/chemrxiv-2022-vz360

[8] https://www.till-biskup.de/de/lehre/programmierkonzepte/.

[9] https://www.till-biskup.de/de/lehre/forschungsdatenmanagement/.

Primary author: Dr BISKUP, Till (Physikalisch-Technische Bundesanstalt (PTB))

Presenter: Dr BISKUP, Till (Physikalisch-Technische Bundesanstalt (PTB))

Session Classification: Poster

Contribution ID: **112** Type: **Use case flash talk (Mon)**

UC10: Diffraction (small and wide angle) & Spectroscopy

Monday 24 March 2025 15:53 (7 minutes)

Within Use Case 10 the **ICSP@FAU group** works towards an automated, generic and FAIR data workflow for research institutes handling data from various sources. Key components of the workflow are a local NOMAD Oasis1, which is hosted by a local computer center and the utilization of the NeXus2 standard wherever possible. We present how we utilize and customize our NOMAD Oasis to fit the FAIR requirements and support our scientific work from data collection to publication.
Key metadata requirements for biological samples studied within the Bio-SAXS part of Use Case 10 are identified by learning from the automatic data and metadata pipeline at EMBL's SAXS beamtime (P12), PETRA III and the integrated SASBDB3 database, a federated database designed for biological SAXS and SANS data, offering curated, searchable experimental data along with relevant metadata. Yet, at most beamlines and especially lab sources no automatic metadata pipeline exists and metadata is often collected manually and separately from the raw data. Here, we present our efforts at **CAU, EMBL and Uni Siegen** to increase findability and accessibility by establishing data handling protocols including the creation of NeXus files and generation of persistent sample identifiers (IGSN) on an exemplary data publication of Bio-SAXS data4 collected at BL2, DELTA.
Progress on automated feature detection in 2D image data at **Uni Tübingen** will be presented by machine learning, focussing here on grazing incidence wide angle X-ray Scattering (GIWAXS) 5-7. The peak detection was already demonstrated to be able to track and process scattering features during in-situ synchrotron experiments6. Additionally, a labelled data set for benchmarking peak detection algorithms was published5.

References
1) https://nomad-lab.eu/nomad-lab/nomad-oasis.html
2) https://www.nexusformat.org/
3) E. Valentini et al. Nucleic Acids Research, 2015, Vol. 43, Database issue D357–D363.
4) S. C. Hövelmann et al. IUCrJ, 2024, 11, 486-493.
5) C. Völter et al. J. Appl. Cryst. (2025), in print
6) V. Starostin et al. Synchrotron Radiation News 35 (2022) 21
7) V. Starostin et al. npj Comput Mater 8 (2022) 101

Primary authors: HINDERHOFER, Alexander; DALLMANN, Johannes (Friedrich-Alexander-Universität Erlangen-Nürnberg, Institute for Crystallography and Structural Physics); HÖVELMANN, Svenja (Kiel University)

Co-authors: Dr HAKIM, Bishoy (FAU); MURPHY, Bridget; GUTT, Christian (Universität Siegen); Dr LAPKIN, Dmitry (Uni Tübingen); Prof. SCHREIBER, Frank (Uni Tübingen); Mr GÖTZ, Klaus (ICSP at Uni FAU); Mrs DARGASZ, Michelle (Uni Siegen); SCHNEIDER, Thomas (Embl); UNRUH, Tobias (FAU)

Presenter: DALLMANN, Johannes (Friedrich-Alexander-Universität Erlangen-Nürnberg, Institute for Crystallography and Structural Physics)

Session Classification: Use-cases (flash talks)

Contribution ID: **113** Type: **Task Area status update (Tue)**

Status Update TA2 ((Meta)Data Repositories and Catalogues)

Tuesday 25 March 2025 09:30 (15 minutes)

Task Area 2 (TA2) aims to create raw- and curated open (Meta)data repositories and catalogs that align with FAIR principles across PaN radiation sources, universities, and other research institutions. The focus is on developing corresponding services to ensure FAIR data sustainability and achieve comprehensive usability of the data for the global scientific community in PaN Research. The repositories will not only contain raw and processed data but also detailed documentation of all processing, analysis steps, and sample descriptions. This approach enhances transparency, improving the quality, trustworthiness, and reusability of datasets used in scientific publications. In this presentation, we will highlight the key achievements and deliverables of the current project status including an outlook to future developments planned for DAPHNE 2.0 as contributed by the TA2 members.

Primary author: BUSCH, Sebastian (GEMS at MLZ, Helmholtz-Zentrum Hereon, Germany)

Co-authors: Dr HAKIM, Bishoy (FAU); UNRUH, Tobias (FAU)

Presenter: BUSCH, Sebastian (GEMS at MLZ, Helmholtz-Zentrum Hereon, Germany)

Session Classification: Metadata, Catalogues and Software (TA1, TA2 and TA3)

Contribution ID: **114** Type: **Poster**

Key Insights on disseminating Multidimensional Rietveld Refinement

Our group specializes in the development of neutron powder diffraction methods, primarily utilizing the neutron time-of-flight diffractometer POWTEX, which was developed in collaboration with Forschungszentrum Jülich at FRM-II in Garching. As part of the DAPHNE4NFDI project, we are expanding these methodologies to enhance their applications, with a strong emphasis on sustainability.

We are actively engaged in researching multidimensional data reduction and Rietveld refinement techniques that are applicable not only to POWTEX but also to large-area detector TOF diffractometers such as POWGEN and SNAP at SNS, ORNL in the USA, as well as future TOF diffractometers at ESS. Recent advancements have allowed for efficient multidimensional data reduction using the Mantid software. Furthermore, we have incorporated multidimensional refinement routines into a customized version of GSAS-II, which is currently undergoing refactoring and testing with real samples.

Our successful trials have already demonstrated the application of these methods on the POWTEX detector operated at POWGEN (SNS, ORNL), as published [1]. However, to extend these methodologies to other instruments effectively, the availability of instrument-specific information is crucial; unfortunately, this information is often not readily provided.

To address this issue, we propose creating a nexus file that serves as a centralized repository for all metadata related to an experiment—including essential instrument-specific details. This nexus file could also encompass information about any related files necessary during data reduction and versioning of algorithms used for this process. By ensuring that refinement software is properly adjusted, separate files for instrument parameters would become superfluous, thereby streamlining the analysis of measurement data.

Building on these concepts, we have developed a sample nexus file that incorporates some of the aforementioned features. It contains fundamental instrument parameters along with links to related files. These links can then be utilized during data reduction in Mantid through a slightly modified version of our algorithm PowderReduceP2D. Additionally, this algorithm automatically generates instrument parameter files that would otherwise need to be downloaded separately.

This integrated approach enhances both efficiency and accessibility in neutron powder diffraction research while supporting our commitment to sustainability and broadening the applicability of our methods across various instruments.

[1] Houben, Andreas; Meinerzhagen, Yannick; Nachtigall, Noah; Jacobs, Philipp; Dronskowski, Richard (2023): POWTEX visits POWGEN. In J Appl Cryst 56 (3), pp. 633–642. DOI: 10.1107/S1600576723002819.

Primary author: Dr MEINERZHAGEN, Yannick (RWTH Aachen)

Co-author: HOUBEN, Andreas (RWTH Aachen)

Presenter: Dr MEINERZHAGEN, Yannick (RWTH Aachen)

Session Classification: Poster

Contribution ID: **115** Type: **Use case flash talk (Mon)**

UC6: Soft Matter and Liquid Interfaces X-ray Reflectivity Diffraction & Spectroscopy

Monday 24 March 2025 15:25 (7 minutes)

Within the soft matter and liquid interfaces X-ray reflectivity Use Case 6, we develop a FAIR data pipeline for X-ray reflectivity at beamline P08, PETRA III. This includes automating electronic lab notebooks (ELNs) [1], metadata ingestion from the control system and IGSN creation for samples. Collaborating with DESY beamline scientists, experimental control group, and IT, we are implementing automated metadata ingestion into SciCat, based on the proposed DAPHNE4NFDI metadata schema [2]. Here, we will present the current status of the data and metadata integration and the recently introduced PaN Reflectivity Database [3,4], aggregating high-quality photon and neutron reflectometry data. Additionally, we will show integration of machine learning solutions [5,6,7] for X-ray and neutron reflectometry, collaborating with large-scale facilities to including also for feedback to the instruments, including so-called closed loop experiments [8] and expanding research to include hard materials and magnetism.

References
[1] P. Jordt et. al., Supplementary Information for Publication: Specifications for Electronic Laboratory Notebooks (ELN) in the Photon and Neutron Community. Zenodo (2025)
[2] W. Lohstroh et al., DAPHNE4NFDI - Draft recommendations on metadata capture and specifications (1.0). Zenodo (2024)
[3] https://public-data.desy.de
[4] https://sisyphos.desy.de
[5] A. Greco et al., J. Appl. Cryst. 55, 362-369 (2022).
[6] V. Munteanu et al., J. Appl. Cryst. 57, 456-469 (2024).
[7] V. Starostin et. al., Sci. Adv. (2025), in print
[8] L. Pithan et al., J. of Synchr. Rad. 30, 1064 (2023)

Primary authors: Ms HOEVELMANN, Svenja (CAU Kiel); KOBUS, Julia (Kiel University); JORDT, Philipp (Kiel University); HAYEN, Nicolas; MURPHY, Bridget; HINDERHOFER, Alexander; LAPKIN, Dmitry (Universität Tübingen); SCHREIBER, Frank (Universität Tübingen)

Presenter: Ms HOEVELMANN, Svenja (CAU Kiel)

Session Classification: Use-cases (flash talks)

Contribution ID: **116** Type: **Poster**

Data Management Concept for Heinz Maier-Leibnitz Zentrum (MLZ)

The data management concept for MLZ will be presented in this poster contribution. Various services work together, from initial proposal submission to final publication. The instrument control software NICOS delivers the beamline data to additional services like the metadata catalog SciCat, Electronic Laboratory Notebook (ELN) services designed for use at MLZ, and additional analysis tools, e.g., JupyterHub. An important requirement for the data management concept is to allow easy integration of all (Lab) instruments, including those not operated by MLZ's control software NICOS. For this purpose, a message queuing system is used at MLZ, allowing loosely coupled data acquisition systems and follow-up services such as ingestion into the metadata catalog SciCat. It is mandatory to allow easy integration of all (Lab)instruments to work at MLZ, including those not operated by NICOS. This poster will also showcase further current activities at MLZ.

Primary author: Dr TRAGESER, Christian (JCNS)

Co-authors: ZAFT, Alexander (Forschungszentrum Jülich); FELDER, Christian (Forschungszentrum Jülich GmbH)

Presenter: Dr TRAGESER, Christian (JCNS)

Session Classification: Poster

Contribution ID: **117** Type: **Poster**

Data driven models for structure determination from powder diffraction measurements

The structure determination pipeline of a polycrystalline sample from its powder diffraction measurement has several adjunct steps, which are as follows: background removal, peak finding, indexing, database matching, extinction and space group(s) determination, atom type and position determination, and finally Rietveld refinement. Experienced users employ various well-established algorithms—available in software tools such as EXPO2014, Match!, GSASII, and FullProf—at different stages of the pipeline. At each step, they must apply their expertise to interpret results, evaluate candidate solutions, and make informed decisions to advance the structure determination process effectively. Our research focuses on (1) leveraging data-driven models as alternatives to various steps in the aforementioned pipeline and (2) developing a broader universal AI framework to directly propose candidate solutions. In the first direction, we have achieved state-of-the-art results in symmetry prediction tasks, with our model outperforming well-known indexing algorithms, such as those in EXPO2014, when used independently. We analyze our models in terms of robustness to noise and experimental effects, generalizability, and scalability, with a focus on their applicability to our broader goal of developing an automated AI for structure solution. From a user-focused software standpoint, we aim to integrate our most robust models into our AIXtal platform, as AI helpers.

Primary author: DAS, Shubhayu (RWTH Aachen University, Institute of Inorganic Chemistry)

Co-author: HOUBEN, Andreas (RWTH Aachen)

Presenter: DAS, Shubhayu (RWTH Aachen University, Institute of Inorganic Chemistry)

Session Classification: Poster

Contribution ID: **118** Type: **Poster**

SciCat - Metadata management for DESYs photon science experiments

The goal is clear: DESY photon scientists require a robust tool to manage metadata throughout the entire data lifecycle—from acquisition to publication and long-term archiving. SciCat, a scientific catalog, was chosen to fulfill this need.

During the initial phase, we at DESY could make crucial deployment and user interaction experiences which was presented at the last annual meeting. Over the past year, DESY could then proceed focusing on the actual benefit of SciCat to the users: being able to offer unique and persistent digital object identifiers (DOIs) of DESY data. The challenge was to find the best operational solution for DESY out of the many possible ways, purposes and scopes.

The scope we deal with concerns users at PETRA and FLASH beamline experiments who require a catalogue that provides a handle on datasets that are still under the embargo period and its (meta) data access is - at that stage - not yet public.

This contribution highlights efforts made in the last year towards this DOI minting service. A large part went and still goes into development work, before we could proceed. We now work towards a concrete setup of that service at DESY where automated elements and typical manual steps like data curation have their role in that service. The service, its prototype version, the current status and what we aim for in production will be presented in our roadmap.

While working on the implementation of this service for DESY, every SciCat user, enduser or site-admin, within DAPHNE4NFDI or worldwide will also benefit from an extended and updated documentation of SciCat capabilities –where a DAPHNE4NFDI member has taken the lead on. A preview of the updated documentation will be given too.

Primary author: KWEE-HINZMANN, Regina

Co-authors: SCHLUENZEN, Frank; KHOKHRIAKOV, Igor (DESY); REPPIN, Johannes (DESY); PITHAN, Linus (DESY, FS-EC); KOEHLER, Martin (Deutsches Elektronen-Synchrotron DESY); Dr WETZEL, Tim (Deutsches Elektronen-Synchrotron DESY)

Presenter: KWEE-HINZMANN, Regina

Session Classification: Poster

Contribution ID: **119** Type: **Facility update (Wed)**

Present and future of DAPHNE4NFDI activities at HZDR

Wednesday 26 March 2025 09:55 (15 minutes)

Within DAPHNE4NFDI TA1, HZDR is developing tools for metadata capture with the aim to facilitate automatic processing of that metadata in the data management chain at HZDR. These tools (ShotSheet, SimulationLogger) are currently in testing at productive environments (TRL 7). Although the tools were developed for use cases at HZDR, they are also designed to be used elsewhere, e.g. at other partners of DAPHNE4NFDI. Therefore, the applications will be enhanced in terms of configurability and prepared for a first software publication.
Within DAPHNE4NFDI TA2, SciCat has been included into the data management chain at HZDR (currently on TRL 6) in order to provide means to expose scientific metadata to external search portals like the PANOSC data portal. So far, the data and software publication repository system of HZDR (RODARE, TRL 9) was connected to external portals within the EOSC like B2FIND or EUDAT, mainly employing bibliographic metadata for findability. SciCat enables now to also search with scientific metadata like experiment classification terms (e.g. PANET ontology) and many more. Metadata entries in SciCat can originate from the above-mentioned sources of DAPHNE4NFDI TA1 as well as from the ELN system at HZDR (MediaWiki, TRL 9). The transfer from ELN is on TRL 7, the transfer from the developed apps on TRL 2.
All these implementations and cross-connections are prototypes in order to demonstrate on actual data the functioning and to learn best practice before widening the scope for all scientific fields represented at HZDR. For the future of DAPHNE4NFDI, we plan to implement metadata standards (schemas) at the start of the data management chain. If standards exist, they should shape the metadata entry into the according form and determine data transfer and curation pipelines into the metadata repositories. A particular concern of ours is to make the underlying metadata standard in catalogues such as SciCat visible and, in particular, filterable in the future. Increasing the visibility of metadata standards in the repositories will promote the use of these standards and help to develop them further. Here we can utilize former and other projects like HELIPORT, HELPMI, NAPMIX or Semantic-X-Lab. HELIPORT allows to keep control of involved resources in an experiment, but also schemas, metadata catalogs, toolchains etc. HELPMI and NAPMIX started to define a metadata standard in the field where the metadata capturing tools have been developed, and these metadata standard drafts should be developed further. Semantic-X-Lab is a current project to extract metadata from various portals and related systems to create a comprehensive knowledge graph to reveal connections that were not previously obvious.

Primary authors: TIPPEY, Kristin (HZDR); SCHLENVOIGT, Hans-Peter (HZDR); KLUGE, Thomas (HZDR); PAPE, David (Helmholtz-Zentrum Dresden - Rossendorf); VOIGT, Martin (HZDR); GRUBER, Thomas (HZDR); KNODEL, Oliver (Helmholtz-Zentrum Dresden-Rossendorf)

Presenter: SCHLENVOIGT, Hans-Peter (HZDR)

Session Classification: Facility Updates

Contribution ID: **120** Type: **Poster**

Data & Metadata in the Helmholtz ROCK-IT Project

For a ROCK-IT thermal catalysis experiment, the (meta)data lifecycle in the facility begins with the initial proposal and extends through the entire investigation, culminating in the archiving and cataloguing of the (meta)data. At all stages, coherence between data and metadata is vital. We are using a combination of established systems and new initiatives to achieve a comprehensive framework aligned with FAIR principles.

Primary authors: GAUR, Abhijeet (KIT); VELAZQUEZ SANCHEZ, Ana (HZB); ROTHKIRCH, Andre (DESY); AGGARWAL, Anjali (DESY); BRUENDERMANN, Erik (KIT); GUENTHER, Gerrit (HZB); Dr PEREZ PONCE, Hector (HZB); GÖRZIG, Heike (HZB); GRUNWALDT, Jan-Dierk (KIT); GETHMANN, Julian (KIT); RIAL, Katherine (HZB); KRAHL, Rolf (HZB); GRUBER, Thomas (HZDR); Dr KONRAD, Uwe (HZDR); Mr BINIYAMINOV, Vitaly (KIT)

Presenters: VELAZQUEZ SANCHEZ, Ana (HZB); RIAL, Katherine (HZB)

Session Classification: Poster

Contribution ID: **121** Type: **Poster**

Developing an all-in-one ELN and RDM tool for university scale facilities

Research data management (RDM) on the scale of university facilities in contrast to largescale facilities has different requirements and challenges. First and foremost are the need to catalogue data from a large variety of different techniques and instruments as well as reliably document the work of frequently changing researchers and projects. For this reason, the ICSP@FAU group within use case 10 works towards an automated, generic and FAIR data workflow, handling data from various sources. This workflow is inherently connected to an electronic lab notebook (ELN) system via the use of a local NOMAD Oasis1.

We are actively developing a Nomad Oasis plugin package that enables the automatic processing of instruments located at our chair2 as well as generic measurement files. For the data pipelines, we are utilizing the NeXus3 standard wherever possible, since NeXus aims to already achieve the A (accessibility) I (interoperability) and R (reusability) principles of FAIR data. This enables us to work in close cooperation with the developers of FAIRmat developing the python converter package pynxtools4.

To meet the need of a stable system that is available without the need of constant maintenance by faculty personnel we are working together with the FAU Competence Center for Research Data and Information (CDI). Every time a stable release of a Nomad Oasis environment is achieved by our work, it will be included in the Nomad instances hosted by the CDI. On this basis, the CDI will install a Nomad Oasis environment enabling other institutes of the FAU to easily request their own instances. The Nomad Oasis ecosystem will be maintained and kept up-to-date by the CDI.

The poster will present how we utilize and customize our NOMAD Oasis to fit the FAIR requirements and support our scientific work from data collection to publication. It will detail our previous achievements in developing these systems, including a metadata PostgreSQL database3, our configuration of the nomad oasis4 and plugin development5 as well as the implementation at our chair. Furthermore, we will present our roadmap for developments planned in the future.

References

1) https://nomad-lab.eu/nomad-lab/nomad-oasis.html
2) https://www.icsp.nat.fau.eu/list-equipment/
3) https://www.nexusformat.org/
4) https://github.com/FAIRmat-NFDI/pynxtools
5) https://codebase.helmholtz.cloud/DAPHNE4NFDI/catacore
6) https://github.com/ICSP-LKS/nomad_icsp
7) https://codebase.helmholtz.cloud/DAPHNE4NFDI/elno

Primary authors: DALLMANN, Johannes (Friedrich-Alexander-Universität Erlangen-Nürnberg, Institute for Crystallography and Structural Physics); GÖTZ, Klaus (ICSP at Uni FAU)

Co-authors: Dr HAKIM, Bishoy (FAU); UNRUH, Tobias (FAU)

Presenters: DALLMANN, Johannes (Friedrich-Alexander-Universität Erlangen-Nürnberg, Institute for Crystallography and Structural Physics); GÖTZ, Klaus (ICSP at Uni FAU)

Session Classification: Poster

Contribution ID: **122** Type: **Task Area status update (Tue)**

Status Update TA3 (Infrastructures for Data and Software Reuse)

Tuesday 25 March 2025 10:15 (15 minutes)

We will give an overview on activities within Task Area 3, with a focus on achievements, both by individual partners and on the global level, within the period since the last user meeting.
This part will cover software developments and maturation, as well as the establishment of a DAPHNE4NFDI software catalogue and activities to harmonise good software development and deployment practice.
The report will also give an outlook on future plans and challenges, affecting the remaining duration of the current DAPHNE4NFDI project, and visions for the next round (”DAPHNE-2.0”), in which we will cover topics such as implementation of machine learning methods, software integration into workflows and remote computing approaches.

Primary authors: DALL’ANTONIA, Fabio (European XFEL GmbH); SCHREIBER, Frank (Universität Tübingen)

Co-author: HINDERHOFER, Alexander

Presenters: DALL’ANTONIA, Fabio (European XFEL GmbH); SCHREIBER, Frank (Universität Tübingen)

Session Classification: Metadata, Catalogues and Software (TA1, TA2 and TA3)

Contribution ID: **123** Type: **Poster**

Reflectorch: A machine learning application for reflectometry data analysis

We present a machine learning approach for automatized analysis of X-ray (XRR) and neutron reflectivity (NR) data that. The method utilizes prior knowledge to regularize the training process over larger parameter spaces. [1-2] We demonstrate the effectiveness of our method in various scenarios, including multilayer structures with box model parametrization and a physics-inspired special parametrization of the scattering length density profile for a multilayer structure. In contrast to previous methods, our approach scales favorably when increasing the complexity of the inverse problem, working properly even for a several layer multilayer model and an N-layer periodic multilayer model with up to 20 open parameters. We will also discuss autonomous experiments enabled by machine-learning-based online data analysis in synchrotron beamline environments. 3

1 V. Munteanu, V. Starostin, A. Greco, L. Pithan, A. Gerlach, A. Hinderhofer, S. Kowarik, F. Schreiber, Neural network analysis of neutron and X-ray reflectivity data incorporating prior knowledge
J. Appl. Cryst. 57 (2024) 456

2 V. Starostin, M. Dax, A. Gerlach, A. Hinderhofer, Á. Tejero-Cantero, F. Schreiber, Fast and reliable probabilistic reflectometry inversion with prior-amortized neural posterior estimation
Sci. Adv. (2025), in print

3 L. Pithan, V. Starostin, D. Marecek, L. Petersdorf, C. Völter, V. Munteanu, M. Jankowski, O. Konovalov, A. Gerlach, A. Hinderhofer, B. Murphy, S. Kowarik, F. Schreiber, Closing the loop: Autonomous experiments enabled by machine-learning-based online data analysis in synchrotron beamline environments
J. Synchrotron Rad. 30 (2023) 1064

Primary authors: HINDERHOFER, Alexander; MUNTEANU, Valentin (Universität Tübingen); LAPKIN, Dmitry (Universität Tübingen); SCHREIBER, Frank (Universität Tübingen)

Presenter: HINDERHOFER, Alexander

Session Classification: Poster

Contribution ID: **124** Type: **Poster**

mlgid: Machine learning assisted high-throughput workflow for GIWAXS data analysis

Due to the lack of phase information, determining the physical parameters of multilayer thin films from measured neutron and X-ray reflectivity curves is, on a fundamental level, an underdetermined inverse problem. This so-called phase problem poses limitations on standard neural networks, constraining the range and number of considered parameters in previous machine learning solutions. To overcome this, we present an approach that utilizes prior knowledge to regularize the training process over larger parameter spaces. [1-3] We demonstrate the effectiveness of our method in various scenarios, including multilayer structures with box model parameterization and a physics-inspired special parameterization of the scattering length density profile for a multilayer structure. By leveraging the input of prior knowledge, we can improve the training dynamics and address the underdetermined (”ill-posed”) nature of the problem.

1 V. Munteanu, V. Starostin, A. Greco, L. Pithan, A. Gerlach, A. Hinderhofer, S. Kowarik, and F. Schreiber.
Neural network analysis of neutron and X-ray reflectivity data incorporating prior knowledge
J. Appl. Cryst. 57 (2024) 456

2 V. Starostin, M. Dax, A. Gerlach, A. Hinderhofer, Á. Tejero-Cantero, and F. Schreiber.
Fast and reliable probabilistic reflectometry inversion with prior-amortized neural posterior estimation
Sci. Adv. (2025), in print

3 L. Pithan, V. Starostin, D. Marecek, L. Petersdorf, C. Völter, V. Munteanu, M. Jankowski, O. Konovalov, A. Gerlach, A. Hinderhofer, B. Murphy, S. Kowarik, and F. Schreiber.
Closing the loop: Autonomous experiments enabled by machine-learning-based online data analysis in synchrotron beamline environments
J. Synchrotron Rad. 30 (2023) 1064

Primary authors: HINDERHOFER, Alexander; MUNTEANU, Valentin (Universität Tübingen); LAPKIN, Dmitry (Universität Tübingen); SCHREIBER, Frank (Universität Tübingen)

Presenter: HINDERHOFER, Alexander

Session Classification: Poster

Contribution ID: **125** Type: **Talk**

Research Data Management Workflows for Multidimensional Characterization Techniques with NOMAD

Tuesday 25 March 2025 11:35 (10 minutes)

We present an overview of research data management workflows for multidimensional characterization techniques using the NOMAD platform. Our approach focuses on efficiently handling large-volume datasets, particularly in HDF5 format, and developing specialized NeXus application definitions for emerging characterization methods. We demonstrate how cloud-based analysis tools can be seamlessly integrated into the entire workflow, illustrated by examples from electron microscopy and multidimensional photoelectron spectroscopy. By leveraging customized JupyterLab environments and desktop-based tools in the cloud, this strategy supports efficient, advanced analyses that remain tightly integrated with the NOMAD data infrastructure, eliminating the need to relocate large datasets and enhancing data shareability. Furthermore, we show how instrument inventories and sample metadata can be linked to specific measurements, enabling robust traceability and a comprehensive history throughout the entire experimental lifecycle.

Primary author: MARQUEZ PRIETO, Jose Antonio (Humboldt University of Berlin)

Co-author: MANSOUR, Ahmed (Physics Department and IRIS Adlershof, Humboldt-Universität zu Berlin, Germany)

Presenter: MARQUEZ PRIETO, Jose Antonio (Humboldt University of Berlin)

Session Classification: Talks - Collaboration with/in NFDI

Contribution ID: **126** Type: **Poster**

Advancing FAIR principles for MHz-XPCS experiments at EuXFEL within the framework of the DAPHNE4NFDI project

X-ray Photon Correlation Spectroscopy (XPCS) is an experimental technique used to study protein and molecular dynamics on length scales ranging from angstroms to micrometers. The unique properties of the pulses from the European X-ray Free Electron Laser (EuXFEL), such as exceptional transverse coherence and megahertz (MHz) repetition rate, have made it possible to investigate the collective diffusive dynamics of such systems with (sub-) microsecond temporal resolution. Although XPCS has become a routine technique at storage-ring sources, its implementation at EuXFEL is still in progress.

We aim at developing the MHz-XPCS at EuXFEL as a routine technique guided by the FAIR principles, the implementation of which is actively progressing within the DAPHNE4NFDI project. In our poster, we discuss possible strategies to achieve this goal, covering stages from initial near-online data processing to the uploading of the refined "Top Level"data on the dedicated DAPHNE4NFDI web platform.

Primary author: Dr LEONAU, Aliaksandr (UniSiegen / EuXFEL)

Co-authors: TOSSON, Amir (Universität Siegen); RAZA, Agha Mohammad (Universität Siegen); ÖZTÜRK, Özgül (Universität Siegen); DARGASZ, Michelle (Universität Siegen); ANTHUPARAMBIL, Nimmi Das (Universität Siegen); UNGER, Frederik (Universität Siegen); GAUTAM, Randeer Pratap (Universität Siegen); WRIGLEY, James (European XFEL); TEODORO, Matheus (European XFEL); CARINAN, Cammille (European XFEL); ROSCA, Robert (European XFEL); DALL'ANTONIA, Fabio (European XFEL GmbH); MÖLLER, Johannes (European XFEL); JO, Wonhyuk (European XFEL); BRAUSSE, Felix (European XFEL); MADSEN, Anders (European XFEL); GELISIO, Luca (European XFEL); GUTT, Christian (Universität Siegen)

Presenter: Dr LEONAU, Aliaksandr (UniSiegen / EuXFEL)

Session Classification: Poster

Contribution ID: **127** Type: **Poster**

Concept for a generic export and import mechanism for electronic lab books based on OpenSemanticLab/MediaWiki

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 poster 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.
Currently Mediawiki pages of a certain object type contain a specific set of properties and could be exported directly. But to apply a certain schema for the export a mapping to a new set of properties can be configured which describe the export schema. This can be done for several export and object types. When creating a new object type the interconnection to existing object types is defined like in a graph. 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. On the other hand with that configuration the opposite direction for importing is possible.

Primary author: GRUBER, Thomas (HZDR)

Co-authors: PAPE, David (Helmholtz-Zentrum Dresden - Rossendorf); VOIGT, Martin (HZDR); KNODEL, Oliver (Helmholtz-Zentrum Dresden-Rossendorf)

Presenter: GRUBER, Thomas (HZDR)

Session Classification: Poster

Contribution ID: **128** Type: **Poster**

Multiplexing TAS measurements assisted by active learning

E. Vitale, M. Ganeva, A. Schneidewind, W. Lohstroh

Three-axis spectrometers (TAS) are versatile instruments to study inelastic neutron scattering. They allow high energy resolution investigations of fundamental excitations across various energy and momentum coordinates. However, traditional TAS methods are limited by point-by-point measurement in reciprocal space, which can be time-consuming and less effective for rapid kinetic studies. A promising method to more efficient TAS measurements is the multiplexing technique, such as Multiplexing-PUMA at MLZ. This method allows simultaneous measurements across multiple (Q, Ef)-channels, enabling broader reciprocal space mapping while maintaining high data quality.
The ARIANE (Artificial Intelligence-Assisted Neutron Experiments) approach has proven effective in optimizing single-point TAS measurements 1. By applying machine learning techniques, it improves data acquisition efficiency. Building on this foundation, we propose an innovative extension of the ARIANE framework for multiplexing measurements in TAS. This extended approach will employ an active learning algorithm to dynamically identify regions of interest in the (Q,ω)-space. It will suggest and measure 11 different locations, provided it complies with the physical constraints of the PUMA multiplexing setup. This will significantly enhance information gain and overall experimental efficiency, maximizing the utilization of neutron beam time for users.

Primary author: VITALE, Eugenio

Co-authors: Dr SCHNEIDEWIND, Astrid (JCNS); Dr GANEVA, Marina (JCNS); Dr LOHSTROH, Wiebke (TUM)

Presenter: VITALE, Eugenio

Session Classification: Poster

Contribution ID: **129** Type: **Talk**

DECTRIS CLOUD : Collaborative, Scalable, and Insightful Analysis of Scientific Data

Tuesday 25 March 2025 16:45 (15 minutes)

Modern photon science experiments generate vast amounts of high-resolution data, necessitating scalable computational solutions to ensure reproducibility, accelerate analysis workflows, and facilitate collaborative research endeavors. DECTRIS CLOUD offers a high-performance platform where scientists can deploy, share, and collaboratively use software, enhancing data curation through advanced visualization and analysis tools that facilitate better decision-making.

By integrating on-demand computing power, automated data pipelines, and secure cloud storage, DECTRIS CLOUD supports streamlined data analysis, reducing processing time while ensuring consistency across experiments. Researchers can share fully configured computational environments using DECTRIS CLOUD's container-based approach, ensuring reproducibility by eliminating discrepancies between local setups. The platform's real-time collaboration tools further enhance teamwork, enabling researchers to share data, monitor experiments, and refine analyses together, fostering user support in photon and neutron large-scale facilities.

This presentation will illustrate how DECTRIS CLOUD accelerates scientific discovery by providing an efficient, transparent, and collaborative platform for data analysis. We will demonstrate how cloud-enabled environments powered by DECTRIS CLOUD's advanced data curation and collaborative analysis capabilities can serve as an elastic extension of local infrastructure, improving efficiency, enhancing data integrity, and fostering transparent, collaborative research in large-scale scientific investigations.

Primary author: LEROY, Ludmila (DECTRIS)

Co-authors: Dr BURIAN, Max (DECTRIS); Dr LARSEN, Camilla (DECTRIS)

Presenter: LEROY, Ludmila (DECTRIS)

Session Classification: Outreach and Internationalization (TA4 and TA5)

Contribution ID: **130** Type: **Poster**

XRR: Automated Reduction and Analysis & Public Data

We present an automated data reduction and analysis pipeline for X-ray reflectometry (XRR), designed for the LISA instrument at the P08 beamline at DESY. It is adapted from the development of solid surface XRR AI analysis \cite{1,2,3} to be used for liquid surfaces and interfaces. Using AI-based models, it enables real-time processing and analysis of reflectivity measurements. By providing immediate feedback during experiments, it enhances measurement efficiency and supports informed decision-making. As a Python package, this pipeline will also make advanced analysis accessible to less experienced users, improving reproducibility and usability.

In parallel, we present the PaN Reflectivity Database, a resource compiling published photon and neutron reflectometry metadata to serve as a reference for high-quality reflectometry datasets. To facilitate data contributions, we provide a Streamlit-based upload tool where users can enter metadata for their datasets. Before submission, data owners confirm publication under the Creative Commons Attribution (CC-BY) License to ensure open accessibility. Following an internal curation process, both metadata and datasets become publicly available in the database.

By integrating AI-driven automation with open data initiatives, our work advances the accessibility, transparency, and reproducibility of XRR data analysis.

Primary authors: KOBUS, Julia; Dr WETZEL, Tim (Deutsches Elektronen-Synchrotron DESY); PITHAN, Linus (DESY, FS-EC); TIRUMALAI NALLAM CHAKRAVARTY, Parthasarathy (DESY); MURPHY, Bridget

Presenter: KOBUS, Julia

Session Classification: Poster

Contribution ID: **131** Type: **Poster**

DAMNIT-web: Transitioning to a Modern Interface for Automated Experiment Overview

DAMNIT provides scientists with an advanced alternative to traditional spreadsheets by automating the collection, processing, and display of experimental data in an interactive table. Initially delivered with a PyQt desktop application hosted on our internal cluster and accessed via SSH, DAMNIT faced challenges in terms of ease-of-use and accessibility.

We are now transitioning to a modern web-based interface that leverages a specialized technology stack: FastAPI serves as the backend-for-the-frontend, GraphQL manages system communication, and React with Redux powers the frontend. This new platform is engineered for high performance; it loads data on demand (such as through infinite table pagination) and efficiently downsamples heavy data for plotting, providing immediate feedback even when working with large databases. Early testing within our facility is already providing valuable feedback that is driving our user-centered development process.

A key new feature is the integration of OAuth and Keycloak, which secures login and proposal access for authorized users. While DAMNIT is currently deployed on our internal network, we are actively developing a strategy to extend access via the facility's public infrastructure, enabling remote usage.

This innovation represents a significant step forward in experimental data management, merging the reliability of our proven desktop solution with the enhanced accessibility, performance, and modern features of a web platform.

Primary author: CARINAN, Cammille (European XFEL)

Co-authors: TEODORO, Matheus (European XFEL); GELISIO, Luca (European XFEL); ROSCA, Robert; MICHELAT, Thomas (European XFEL)

Presenter: TEODORO, Matheus (European XFEL)

Session Classification: Poster

Contribution ID: **132** Type: **Poster**

RefXAS: XAS reference database under DAPHNE4NFDI

Abstract: Under the DAPHNE4NFDI consortium, RefXAS has been established as a comprehensive open-access reference database for X-ray absorption spectroscopy (XAS), addressing the need for high-quality reference data with well-documented metadata. The platform provides a structured web interface for data submission, automated metadata handling, and quality control to ensure compliance with predefined standards. Metadata categories such as "Sample,""Spectra,""Instrument,"and "Bibliography"facilitate data reusability, enhancing the comparability of experimental and simulated spectra.

Recent advancements include the introduction of a standardised download package, improved filtering mechanisms, and the transition to an institutional VM, ensuring long-term accessibility and alignment with FAIR principles. Additionally, automated querying, beamline registration, and enhanced data visualisation have been implemented to improve usability. Future efforts will focus on integrating the NeXus format to enhance interoperability and support machine learning applications (LLM).

References

1 A. Gaur, S. Paripsa, F. Förste, D. Doronkin, W. Malzer, C. Schlesiger, B. Kanngießer, E. Welter, J.-D. Grunwaldt and D. Lützenkirchen-Hecht, Proc Conf Res Data Infrastr 1 (2023) "CoRDI 2023"

2 Paripsa, S.; Gaur, A.; Förste, F.; Doronkin, D.E.; Malzer, W.; Schlesiger, C.; Kanngiesser, B.; Welter, E.; Grunwaldt, J.-D.; Lützenkirchen-Hecht, D.
"RefXAS: an open access database of X-ray absorption spectra."
J. Synchrotron Rad. (2024) 31, 1105-1117 DOI: 10.1107/S1600577524006751

3 Paripsa, S.; Gaur, A.; Förste, F.; Doronkin, D.E.; Malzer, W.; Schlesiger, C.; Kanngiesser, B.; Welter, E.; Grunwaldt, J.-D.; Lützenkirchen-Hecht, D.
"RefXAS: an open access database of X-ray absorption spectra –improvements and outlook."
Proc. 15th International Conference on Synchrotron Radiation Instrumentation (SRI 2024), J. Phys. Conf. Ser. (2025) accepted.

Primary authors: PARIPSA, Sebastian (University of Wuppertal); GAUR, Abhijeet (ITCP, KIT); FÖRSTE, Frank (TUB); DORONKIN, Dmitry (Karlsruhe Institute of Technology); MALZER, Wolfgang (TUB); SCHLESIGER, Christopher (TUB); KANNGIESSER, Birgit (TUB); WELTER, Edmund (DESY); GRUNWALDT, Jan-Dierk (KIT); Prof. LUETZENKIRCHEN-HECHT, Dirk (Bergische Universität Wuppertal)

Presenters: PARIPSA, Sebastian (University of Wuppertal); GAUR, Abhijeet (ITCP, KIT)

Session Classification: Poster

Contribution ID: **133** Type: **Poster**

Annual Meeting DAPHNE4NFDI for Synchrotron Bio-imaging: the LMU activities

Advances in synchrotron imaging and computed tomography (CT) enable non-invasive, high-resolution 3D visualization of samples, including biological tissues, surpassing standard methods. This innovation generates massive data volumes, requiring tailored solutions for data capture, management, storage, and repositories for processed data and analysis code. These solutions enhance transparency and maximize data reuse in biomedical research. We have been collaborating closely with scientists and IT experts from various synchrotron facilities, primarily at ESRF and PETRA III. We have established a consistent vocabulary and compiled a comprehensive list of parameters applicable across different imaging beamlines. The results of this work will be shared with other imaging beamlines and facilities. The LMU team, together with the tomography-user teams of Göttingen University and HZB, has finalized a metadata table and its description for generic (bio)synchrotron imaging and CT experiments for the white paper. Currently, the implemented or in-progress metadata recording is focused mainly on storing parameters up to the data acquisition phase. In collaboration with ESRF IT experts, we are currently discussing the extension of this list to include comprehensive post-processing metadata, at least up to CT data reconstruction, which is a common component of all CT experiments. Additionally, ESRF does not provide detailed sample metadata information. We are in the process of discussing with the ESRF IT team to integrate the IGSN link with ICAT to connect the corresponding experiments. The subsequent steps of image processing, data analysis, and quantification require significant collaboration within the community, as each application demands specific processing pipelines and tools. We aim to identify the common components of this workflow and propose a suitable metadata scheme. As part of our commitment to TA3, we are developing an image processing workflow tailored for the analysis of biological tissues, incorporating machine learning-based models for segmentation, analysis, and quantification. This processing pipeline will be shared with the broader user community (https://github.com/hfahad/U-NET-Synchrotron_CT-image_segmentation.git). Through collaboration among user groups, scientists, and IT experts from synchrotron facilities, the LMU team is working to establish a general FAIR data workflow for synchrotron imaging and CT experiments. The LMU team has finalizing the installation process of SciCAT on the LMU machine with coordinated efforts from LMU IT Services, the Leibniz Supercomputing Centre (LRZ), and the Helmholtz Centre Dresden-Rossendorf. Currently, SciCAT is undergoing further processing and will be initially available for our group and subsequently for our collaborators. The purpose of this implementation is to install and interface SciCAT with the existing catalog system, enhancing overall data management.

Primary authors: FAHAD, Hafiz Muhammad (LMU); Prof. COAN, Paola (Ludwig Maximilian University of Munich University in Germany)

Presenter: FAHAD, Hafiz Muhammad (LMU)

Session Classification: Poster

Contribution ID: **134** Type: **Use case flash talk (Mon)**

UC3 and UC4: X-ray absorption spectroscopy & X-ray emission spectra, RIXs etc.

Monday 24 March 2025 14:54 (7 minutes)

In the frame of DAPHNE4NFDI, an X-ray absorption spectroscopy (XAS) reference database called RefXAS has been set-up where users are provided with well curated XAS reference spectra along with related metadata fields and online processing tools for visualizing the data. The developed online procedure enables users to submit a raw dataset along with its associated metadata via a dedicated website for inclusion in the database. The published data at the database can be easily linked to the raw data available at other repositories. Quality criteria formulated for the uploaded reference data at the database make users aware of the usability of the data. These quality criteria, which are unique to RefXAS, are further employed for automatic quality check of the uploaded data which is then followed by manual curation at the interface. Different XAS data formats can be uploaded to RefXAS. The output data format consists of all the important metadata provided during upload including quality criteria, curation details and bibliographic details of the data.

Primary authors: GAUR, Abhijeet (ITCP, KIT); PARIPSA, Sebastian (University of Wuppertal)

Co-authors: KANNGIESSER, Birgit (TU Berlin); SCHLESIGER, Christopher (Technische Universität Berlin); Prof. LUETZENKIRCHEN-HECHT, Dirk (Bergische Universität Wuppertal); DORONKIN, Dmitry (Karlsruhe Institute of Technology); WELTER, Edmund (DESY); FÖRSTE, Frank (TUB); GRUNWALDT, Jan-Dierk; MALZER, Wolfgang (TUB)

Presenters: GAUR, Abhijeet (ITCP, KIT); PARIPSA, Sebastian (University of Wuppertal)

Session Classification: Use-cases (flash talks)

Contribution ID: **135** Type: **Poster**

Enhancing Scientific Research and FAIR Data Management Through a Cloud-Based Web Application Platform: FUSE (FAIR Unified Scientific Environment) for the XPCS Use Case in the DAPHNE4NFDI Project

The DAPHNE4NFDI (Data from Photon and Neutron Experiments for the National Research Data Infrastructure) initiative aims to revolutionize data management practices within the photon and neutron science communities by embedding the principles of FAIR data—Findability, Accessibility, Interoperability, and Reusability. Addressing the challenge of unifying diverse datasets and research findings within this specialized scientific field, the project leverages innovative solutions, including dynamic metadata frameworks, expansive data lakes, and user-centered platform designs. This paper discusses the creation of FUSE (FAIR Unified Scientific Environment), a platform that promotes effective research data sharing and utilization. Under the umbrella of the DAPHNE4NFDI project, the X-ray Photon Correlation Spectroscopy (XPCS) method serves as a representative use case to showcase the platform's capabilities. A key feature of FUSE is its ability to enable direct access to reduced/output scientific data from multiple sources, seamlessly integrating these with appropriate metadata items. This integration ensures standardized and comprehensive data descriptions, fostering enhanced interoperability, reproducibility, and data usability across the research community. Additionally, this paper delves into the architecture of FUSE, emphasizing its user-centric design, and explores its impact on facilitating collaborative scientific research. The implementation of experiment classes, dataset instances, Electronic Lab Notebooks (ELN), and sample instances are discussed to underscore their contributions toward achieving a comprehensive and efficient research data management system.

Primary authors: RAZA, Agha Mohammad (University of Siegen); LEONAU, Aliaksandr; TOSSON, Amir; GUTT, Christian (Universität Siegen); DARGASZ, Michelle (Uni Siegen); Dr OEZTUERK, OEzguel (University of Siegen)

Co-authors: UNGER, Frederik (University of Siegen); ANTHUPARAMBIL, Nimmi Das (University of Siegen); GAUTAM, Randeer Pratap (University of Siegen)

Presenter: TOSSON, Amir

Session Classification: Poster

Contribution ID: **136** Type: **Poster**

Use Case 6: Soft Matter and Liquid Interfaces X-ray Reflectivity Diffraction & Spectroscopy

Within the soft matter and liquid interfaces X-ray reflectivity Use Case 6, we develop a FAIR data pipeline for X-ray reflectivity at beamline P08, PETRA III. This includes automating electronic lab notebooks (ELNs) 1, metadata ingestion from the control system and IGSN creation for samples. Collaborating with DESY beamline scientists, experimental control group, and IT, we are implementing automated metadata ingestion into SciCat, based on the proposed DAPHNE4NFDI metadata schema 2. Here, we will present the current status of the data and metadata integration and the recently introduced PaN Reflectivity Database [3,4], aggregating high-quality photon and neutron reflectometry data. Additionally, we will show integration of machine learning solutions [5,6,7] for X-ray and neutron reflectometry, collaborating with large-scale facilities to including also for feedback to the instruments, including so-called closed loop experiments 8 and expanding research to include hard materials and magnetism.

References
1 P. Jordt et. al., Supplementary Information for Publication: Specifications for Electronic Laboratory Notebooks (ELN) in the Photon and Neutron Community. Zenodo (2025)
2 W. Lohstroh et al., DAPHNE4NFDI - Draft recommendations on metadata capture and specifications (1.0). Zenodo (2024)
3 https://public-data.desy.de
4 https://sisyphos.desy.de
5 A. Greco et al., J. Appl. Cryst. 55, 362-369 (2022).
6 V. Munteanu et al., J. Appl. Cryst. 57, 456-469 (2024).
7 V. Starostin et. al., Sci. Adv. (2025), in print
8 L. Pithan et al., J. of Synchr. Rad. 30, 1064 (2023)

Primary authors: HÖVELMANN, Svenja (Kiel University); KOBUS, Julia (Kiel University); JORDT, Philipp (Kiel University); HAYEN, Nicolas; MURPHY, Bridget; HINDERHOFER, Alexander; LAPKIN, Dmitry (Universität Tübingen); SCHREIBER, Frank (Universität Tübingen)

Presenter: HÖVELMANN, Svenja (Kiel University)

Session Classification: Poster

Contribution ID: **137** Type: **Task Area status update (Tue)**

Status Update TA1 (Managing Data Production)

Tuesday 25 March 2025 09:00 (15 minutes)

Presentation on the activities of TA1, covering
- status of the ELNs
- metadata capture at the beamlines (manual and automated)
- best practices for metadata capture
- Sample PIDs
- data formats.

Primary authors: LOHSTROH, Wiebke; MURPHY, Bridget; JORDT, Philipp (Kiel University)

Presenter: LOHSTROH, Wiebke

Session Classification: Metadata, Catalogues and Software (TA1, TA2 and TA3)

Contribution ID: **138** Type: **Facility update (Wed)**

Facility update MLZ

Wednesday 26 March 2025 10:45 (15 minutes)

Advances in neutron instrumentation and techniques offer new opportunities for researchers. At the same time there is an increasing demand to make measured data accessible to the wider community through improved research (meta)data- management, and for implementation of FAIR data principles by which data should be made Findable, Accessible, Interoperable and Reusable. The challenge is becoming even greater due to increasing data rates, multi-dimensional data sets and in-situ / operando experiments. New data management and analysis schemes are established, metadata capture for re-use with searchable catalogues is deployed, and on-the-fly data analysis and reduction are developed. Within the DAHNE4NFDI initiative such workflows along the whole data pipeline are exemplarily within use cases and systematically connected to prior and subsequent laboratory work. This presentation will give an overview of our activities at the Heinz Maier-Leibnitz Zentrum (MLZ) in Garching.

Primary authors: LOHSTROH, Wiebke; BUSCH, Sebastian (GEMS at MLZ, Helmholtz-Zentrum Hereon, Germany); SCHNEIDEWIND, Astrid (JCNS)

Presenter: LOHSTROH, Wiebke

Session Classification: Facility Updates

Contribution ID: **139** Type: **Poster**

Leveraging Large Language Models for Automated Outreach in DAPHNE4NFDI TA4

One of the main goals of TA4 is effective science communication and outreach for maximizing the impact of DAPHNE4NFDI. For managing the outreach activities with less time and sources we explore the integration of Large Language Models (LLMs) to automate key aspects of outreach in TA4 at DAPHNE4NFDI.
We propose an LLM-driven workflow for automatization of (1) event summarization, generating structured reports, social media posts, and news from workshops, meetings and conferences; and (2) social media monitoring and insights, employing AI-based analytics to optimize content strategies and track audience engagement.
Preliminary tests demonstrate that LLMs can generate targeted outreach messages, and provide data-driven insights into audience behavior. This approach serves as a model for integrating LLMs into research outreach, supporting more efficient science communication.

Primary author: Dr ÖEZTÜRK, Öezgül (University of Siegen)

Co-authors: Mr RAZA, Agha Mohammad (University of Siegen); Dr LEONAU, Aliaksandr (University of Siegen); Dr TOSSON, Amir (University of Siegen); Ms DARGASZ, Michelle (University of Siegen)

Presenter: Dr ÖEZTÜRK, Öezgül (University of Siegen)

Session Classification: Poster

Contribution ID: **140** Type: **Talk**

NFDI-MatWerk: Use-case-driven development of research data infrastructure for materials science and engineering

Tuesday 25 March 2025 11:15 (10 minutes)

NFDI-MatWerk represents the communities of materials science and engineering, diverse in material scales, methodologies and technologies. A supporting infrastructure is built upon exemplary infrastructure use cases (IUCs), representing individual domain areas. Two such IUCs will be presented, also highlighting their use of ontologies in the materials science context:

IUC04 aims to ensure a model-driven (guided probing) collection of the relevant experimental and computational data to construct so-called defect-phase diagrams and their post-processing according to newly established simulation protocols. Another focal point is automatic semantic annotation of the computational data and its upload to an ELN system.
IUC17 has the aim of developing semantic representations describing crystalline structures and crystalline defects and their temporal evolution and to test if these descriptions are well designed and applicable for different types of simulations, experiments and microscopy. Another aspect covered is aiding domain scientists in implementing ontologies in their everyday research.

The built infrastructure services and solutions are currently in roll-out for the broader materials science and engineering community. To empower the community for FAIR research data management, NFDI-MatWerk is currently building several central services for learning, connecting and contributing to a FAIR data future.

Primary authors: MOHRBACHER, Julia; KRUZIKOVA, Pavlina

Presenters: MOHRBACHER, Julia; KRUZIKOVA, Pavlina

Session Classification: Talks - Collaboration with/in NFDI

Contribution ID: **141** Type: **Poster**

THE EM GLOSSARY: A COMMUNITY RESOURCE PROVIDING HARMONISED TERMINOLOGY IN ELECTRON MICROSCOPY

With the EM glossary initiative, our goal is to harmonize terminology in the field of electron and ion microscopy so humans, devices and algorithms can exchange data easily and collaborate more efficiently and effectively.
We achieve this through the following measures:
1. moderation of a community collaboratively harmonising terminology,
2. provision of a machine-readable OWL artifact, 1
3. provision of a human-readable web app - the EM Glossary explorer; 2
4. support and counselling for application-level adoption of the resource.

The Helmholtz Metadata Collaboration (HMC) 3 is coordinating the EM glossary initiative. This includes a group of volunteers, with scientists from more than 23 institutions across Switzerland, Austria, and Germany, as well as representatives of the FAIRmat 4 and the NFDI-MatWerk 5 consortia, who meet every two weeks to discuss and achieve consensus on the definitions of terms commonly used in electron and ion microscopy. We aim to produce concise, unpacked definitions with rich annotations in accordance with semantic best practices. New participants are warmly welcomed.
Doing so we have, to date, generated harmonized definitions for more than 65 terms. These are provided as a machine accessible resource 1 in the web ontology language (OWL). The artifact offers stable domain-level semantics and is regularly updated based on progress in our community group. By aligning with, or adopting the artifact in application-level metadata, intra- and interdisciplinary interoperability is improved.
Terminology adopters include the NFDI consortia FAIRmat, and NFDI-MatWerk, the NFFA-Europe Pilot (NEP) 6, the Joint Lab 'Integrated Model and Data-driven Materials Characterization'(JL-MDMC) 7, and the research data management platform NOMAD 8. New adopters are encouraged and welcomed. We provide detailed support and counselling for potential adopters upon request.
In addition to technical adoption there is scientific value in the glossary itself: Easy browsing of the glossary is provided by the EM Glossary Explorer 2 to facilitate the use of the resource for technical writing or for teaching purposes.
In conclusion, through a consensus-based process we have produced a resource for the entire electron and ion microscopy community. Application-level semantics can be aligned with its machine-readable implementation which acts as a semantic glue within the field. In addition, the human-readable resource reduces ambiguity in communication within the field.
Interested in getting involved? Send an email to hmc-matter@helmholtz-berlin.de or hmc@fz-juelich.de to get in touch!
References:
1 EM Glossary OWL artifact base URI: https://purls.helmholtz-metadaten.de/emg/
2 EM Glossary Web Explorer: https://emglossary.helmholtz-metadaten.de/
3 Helmholtz Metadata Collaboration: https://helmholtz-metadaten.de/en
4 FAIRmat project: https://www.fairmat-nfdi.eu/
5 NFDI-Matwerk: https://nfdi-matwerk.de
6 NFFA-Europe Pilot : https://nffa.eu
7 JL-MDMC : https://jl-mdmc-helmholtz.de
8 NOMAD research data management platform: https://nomad-lab.eu/

Primary authors: MANNIX, Oonagh (HMC matter/HZB); HOFMANN, Volker

Co-authors: AZOCAR GUZMAN, Abril (IAS-9, FZJ); HEBERT, Cécile (LSME, IPHYS, Ecole Polytechnique Fédérale de Lausanne); SEDEQI, Mojeeb Rahman (Helmholtz-Zentrum Berlin für Materialien und Energie GmbH (HZB), Helmholtz Metadata Collaboration (HMC)); Dr OEZKAN, OEzlem (HMC); KONIJNENBERG, Peter (FZ Juelich IAS-9); SCHRÖDER, Rasmus (Univeristätsklinikum Heidelberg, Universität Heidelberg); AVERSA, Rossella (Karlsruhe Institute of Technology); BROCKHAUSER, Sandor (Humboldt-Universität zu Berlin); SANDFELD, Stefan

Presenter: MANNIX, Oonagh (HMC matter/HZB)

Session Classification: Poster

Contribution ID: **142** Type: **Use case flash talk (Mon)**

UC9: Tomography

Monday 24 March 2025 15:46 (7 minutes)

Tomography at neutron and photon sources is a ubiquitous tool with applications for a large variety of research domains. On the way to FAIR data, this use case is focusing on the metadata specification of tomography data from both photon and neutron sources, incl. the discussions for a common *nxs specification.

Primary authors: LOHSTROH, Wiebke; GOERZIG, Heike (HZB); HAMMEL, Jörg (Hereon); ALFALOU, Abdel (TUM); SCHULZ, Michael (TUM); ZIESCHE, Ralf (HZB)

Presenter: ZIESCHE, Ralf (HZB)

Session Classification: Use-cases (flash talks)

Contribution ID: **143** Type: **Poster**

CDIF-4-XAS: Describing X-Ray Spectroscopy Data for Cross-Domain Use

The CDIF-4-XAS project - Describing X-Ray Spectroscopy Data for Cross-Domain Use, will enable new science by making it easier to access, combine and reuse XAS data across research infrastructures (RIs) and disciplines.
CDIF-4-XAS will enhance the interoperability and reusability of XAS data by applying the Cross-Domain Interoperability Framework (CDIF), a set of guidelines and practices for using domain-agnostic standards to support the interoperability and reusability of FAIR data, especially across domain and institutional boundaries. By embracing FAIR principles, the project aims to streamline the sharing of XAS data, thus enabling more efficient data integration across RIs and scientific domains, including life sciences, chemistry, and environmental sciences.
X-ray Absorption Spectroscopy (XAS) data is vital for many fields, but its specialised formats and metadata conventions hinder cross-domain use. Data sharing between applications, databases, and facilities is inefficient, leading to the loss of essential experimental information. With increasing data volumes and interdisciplinary collaborations, the need for a more interoperable solution becomes urgent.
The CDIF is a set of practical, implementation-level principles designed to improve data management practices within any community and lower the barriers to cross-domain data reuse. CDIF offers standards and methodologies for achieving different levels of interoperability necessary for reusing data across diverse domains. It is built around five core profiles that address the essential functions for implementing cross-domain FAIR principles.
Through CDIF-4-XAS, metadata and provenance for XAS data are standardised, enhancing the integration of XAS datasets into EOSC infrastructures. This promotes data reuse across different research domains, making XAS data more accessible and valuable.
Coordinated by CODATA, and involving a number of PaNOSC partners and key infrastructures, CDIF-4-XAS will unlock new possibilities for research by enabling seamless data combination and analysis across RIs. Researchers in energy, chemistry, and environmental sciences, among others, will benefit from easier access to XAS data. By implementing CDIF for XAS data (XAS-CDIF), the project tackles the interoperability and reusability challenges described above, and can better integrate the XAS data into the emerging EOSC infrastructures for cataloguing, integration and analysis. The benefits will be considerable: existing resources will become more available, and it will become easier and more efficient to use data across RIs and in other domains. Finally, XAS data is a key research product in the domain of energy research, thereby representing key strategic developments towards a sustainable future for the global society. XAS is an important characterization method for materials used in these fields and essential for developing new materials.
The project started in October 2024 and can present now an overview of standards, vocabularies (and ontologies), data formats and practices within the XAS area.

Primary authors: NIEVA DE LA HIDALGA, Abraham (Cardiff University); GREGORY, Arofan (CODATA); GÖRZIG, Heike (Helmholtz-Zentrum Berlin für Materialien und Energie); LIBORIO, Leandro (STFC); KUBIN, Markus (HMC, HZB); HEIKKURINEN, Matti (CODATA); AUSTIN, Patrick (STFC); KRAHL, Rolf (Helmholtz-Zentrum Berlin für Materialien und Energie); HODSON, Simon (CODATA)

Presenter: GÖRZIG, Heike (Helmholtz-Zentrum Berlin für Materialien und Energie)

Session Classification: Poster

Contribution ID: **144** Type: **Poster**

Exploring the Open Data Landscape: The HMC Dashboard on Open and FAIR Data

The HMC Dashboard on Open and FAIR Data in Helmholtz was designed to support various target groups in finding and monitoring open and FAIR research data produced by the Helmholtz Association. This poster will present our approach and how it relates to photon and neutron research.

The dashboard is an interactive tool developed by the Helmholtz Metadata Collaboration to monitor and analyze open and FAIR data practices within the Helmholtz Association of German research centers. It provides insights into which repositories research communities use to publish research data and allows for assessing trends and gaps therein. The dashboard visualizes data collected through a harvesting and evaluation pipeline, enabling users to explore interactive statistics on data publications from various perspectives, such as the institutional, the repository-level, or from an individual researcher's view. This tool serves as a means to engage various target groups, identify areas for improvement in data management practices, and guide efforts towards creating a unified FAIR data space within the research organization.

With this poster we invite the DAPHNE4NFDI community to discuss our apporach and how it can contribute to improved open and FAIR data practices in photon and neutron science.

Primary author: KUBIN, Markus (HMC, HZB)

Co-authors: PREUSS, Gabriel (HZB); SEDEQI, Mojeeb Rahman (Helmholtz-Zentrum Berlin für Materialien und Energie GmbH (HZB), Helmholtz Metadata Collaboration (HMC)); MANNIX, Oonagh (HMC matter/HZB); EHLERS, Pascal (Institute of Transportation Systems, German Aerospace Center (DLR))

Presenter: KUBIN, Markus (HMC, HZB)

Session Classification: Poster

Contribution ID: **145** Type: **Poster**

Key Insights from the HMC Data Professionals Survey 2024: Assessing Research Data Management Practices

The Helmholtz Metadata Collaboration (HMC) conducted a comprehensive Data Professionals Survey in 2024 to assess the current state of research data management practices across the Helmholtz Association of German Research Centers. As a number of German large scale facilities are operated by the Helmholtz Association, this study is of particular relevance to the photon and neutron communities represented at the DAPHNE4NFDI conference.

Our survey employed a mixed-methods approach, combining quantitative and qualitative data collection through a detailed questionnaire distributed to data professionals across multiple Helmholtz research centers. The analysis focused on identifying common themes, challenges, and opportunities in research data management, with particular attention to FAIR (meta)data practices, the impact of data policies, professional training and tools related to research data management.

The key findings of the survey provide valuable insights into the status quo, gaps, and needs of research data professionals within the Helmholtz Association. These results will inform HMC's strategy to develop targeted support and resources to enhance research data management practices across the organization.

In this poster presentation, we will highlight selected findings that are particularly relevant to the photon and neutron research community. Our findings are intended to contribute to the ongoing dialogue on improving data management practices and fostering collaboration among data professionals in the field.

Primary author: KUBIN, Markus (HMC, HZB)

Co-authors: LEMSTER, Christine (Geomar); SCHWEIKERT, Jan (Karlsruhe Institute of Technology (KIT)); STUCKY, Karl-Uwe (KIT); KULLA, Lucas (DKFZ); MANNIX, Oonagh (HMC matter/HZB); SHANKAR, Sangeetha (German Aerospace Center); Dr GERLICH, Silke (HMC)

Presenters: KUBIN, Markus (HMC, HZB); MANNIX, Oonagh (HMC matter/HZB)

Session Classification: Poster

Contribution ID: **146** Type: **Talk**

PUNCH4NFDI: Workflows with REANA

Tuesday 25 March 2025 11:25 (10 minutes)

The PUNCH4NFDI consortium has been working with developers at CERN to enable the use of the REANA workflow environment in a federated (and heterogeneous) infrastructure. While the access (Authentication) via the NFDI AAI is easy, the connection with different resource-management facilities required work. Within PUNCH4NFDI, we've connected the CoBALD/TARDIS resource manager (integrated now as module in recent REANA release) and also enabled the use e.g. of dCache storage space, albeit this is still ongoing. The use of storage resources via standard S3 protocol is available. Developing and deployment of more job controller interfaces, e.g. for SLURM clusters, is in the works.
During a ErUM data workshop, a workflow for processing of timepix3 raw data stream has been demonstrated together with HZB colleagues. The talk will give an overview of the REANA deployment within PUNCH4NFDI, and then provide some examples to discuss the advantages of a REANA component for standarized processing tasks of medium experiment data. This supports especially the "R" in FAIR.

Primary author: ENKE, Harry (Leibniz Institute for Astrophysics Potsdam (AIP))

Co-authors: Dr SACCHI, Elena (Leibniz Institute for Astrophysics Potsdam (AIP)); Dr KRONAST, Florian (HZB); Dr SCHNIZER, Pierre (HZB)

Presenter: ENKE, Harry (Leibniz Institute for Astrophysics Potsdam (AIP))

Session Classification: Talks - Collaboration with/in NFDI

Contribution ID: **147** Type: **Poster**

Improving research data management for samples: The SEPIA Sample Database for Metadata Storage and Exchange

SEPIA **(Sample Essentials, Persistent Identifiers & Attributes System)**:

The overall goal of the SEPIA project is to enrich the research data collected at the Helmholtz-Zentrum Berlin (HZB) by providing a better description of the sample being measured. The key for this endeavour is to track the full history of the sample and to collect all the information generated on the way. To this end, a database is created to identify the individual sample and allow referencing it unambiguously.

This poster will explore the architecture and functionalities of the SEPIA project, highlighting its role in improving data accessibility and interoperability among researchers and labs. We will discuss the potential impact of SEPIA on collaborative research efforts, data sharing practices, beamlines and the overall improvement of scientific research through better metadata management.

Primary authors: SEDEQI, Mojeeb Rahman (Helmholtz-Zentrum Berlin für Materialien und Energie GmbH (HZB), Helmholtz Metadata Collaboration (HMC)); KRAHL, Rolf (Helmholtz-Zentrum Berlin für Materialien und Energie); RIAL, Katherine (Helmholtz-Zentrum Berlin); GÖRZIG, Heike (HZB)

Presenters: SEDEQI, Mojeeb Rahman (Helmholtz-Zentrum Berlin für Materialien und Energie GmbH (HZB), Helmholtz Metadata Collaboration (HMC)); KRAHL, Rolf (Helmholtz-Zentrum Berlin für Materialien und Energie)

Session Classification: Poster

Contribution ID: **148** Type: **Poster**

Metadata collection and creation of sample identifier in the Bio-SAXS use case based on the concept of the Small Angle Scattering Biological Data Base

To support the principles of FAIR data, the Small Angle Scattering Biological Data Base (SASBDB) serves as a platform to make biological SAXS and SANS data more accessible and findable 1. Within the SASBDB framework, experimental data can be deposited alongside relevant experimental conditions, sample details, and instrument characteristics, ensuring comprehensive documentation for future reuse 2.

Using SASBDB as a benchmark, we identified essential metadata and information required to properly deposit entries in such a database. While some beamlines are already designed to save important metadata during data collection, many smaller or laboratory X-ray sources lack automated metadata recording, resulting in raw detector images being stored without crucial sample or instrument information. To address this challenge, we introduce a standardized metadata catalogue related to BioSAXS experiments, along with a method for sample identification using the generation of an International Generic Sample Number (IGSN). We present a use case specific workflow scheme that includes the creation of a Nexus file as an additional step that incorporates metadata and IGSN information in a standardized format.

1 E. Valentini et al. Nucleic Acids Research, 2015, Vol. 43, Database issue D357–D363

2 A. Kikhney et al. Protein Science, 2019, Vol. 29, Issue 1

Primary authors: DARGASZ, Michelle; HÖVELMANN, Svenja (Kiel University)

Co-authors: MURPHY, Bridget; GUTT, Christian (Universität Siegen)

Presenters: DARGASZ, Michelle; HÖVELMANN, Svenja (Kiel University)

Session Classification: Poster

Contribution ID: **149** Type: **Poster**

Labimotion as a customizable ELN for heterogeneous catalysis workflows

Due to the complex data structures and the unique needs of the community, the design of Electronic Lab Notebooks (ELNs) solutions in the field of heterogeneous catalysis and particularly when including research in large scale facilities1 is highly challenging. LabIMotion, an extension of the open-source ELN Chemotion2, is designed to fill the gap between the preparation, testing and characterization of heterogeneous catalytic materials. Within the ELN, the user can work with pre-defined workflows, but also free input fields for e.g. research plans as well as link and share data with other users. LabIMotion envisions direct links to metadata catalogues like SciCat3 (for synchrotron/neutron characterizations), PID solutions and advanced research data management tools like Adacta (for improved traceability of catalytic data, experimental setups, and related resources). In the presented work, the adaptability of LabIMotion in the catalysis field is demonstrated through examples such as Cu-based catalysts for methanol synthesis4 and noble metal-based emission control catalysts5.

DAPHNE4NFDI Deliverables / Categories:
2.2.1 White paper on metadata definition and known metadata
2.3.2 Preliminary specification for use cases minimum metadata

Topics to be (potentially) discussed/addressed in the course of the meeting:

1. Sample PID, DataCite services
2. ELN integration –Integration of Electronic log notebook

Cooperation partners:
Nicole Jung, Schirin Hanf, Anna Zimina, Arik Beck, Olaf Deutschmann
Karlsruhe Institute of Technology (KIT), Engesserstr. 20, Karlsruhe, D-76131.

1 https://www.tandfonline.com/doi/full/10.1080/08940886.2024.2432265
2 https://chemotion.net/docs/labimotion
3 https://scicatproject.github.io/
4 https://www.spp2080.org
5 https://www.trackact.kit.edu

Primary authors: MAURER, Florian (Karlsruhe Institute of Technology); GRUNWALDT, Jan-Dierk (KIT)

Co-authors: Ms SCHULTE, Mariam (Karlsruhe Institute of Technology); DOLCET, Paolo (Karlsruhe Institute of Technology)

Presenter: GRUNWALDT, Jan-Dierk (KIT)

Session Classification: Poster

Contribution ID: **150** Type: **Poster**

Catalysis use case in DAPHNE4NFDI at KIT

In the frame of the collaboration centre "TrackAct"(www.trackact.kit.edu), the priority program SPP2080 (www.spp2080.org) and ROCK-IT (www.rock-it-project.de), we have developed a workflow for catalysis experiments in DAPHNE4NFDI. Starting with sample preparation and basic characterization of the samples, documentation based on FAIR principle is needed by established tools such as LabIMotion for synthesis, catalytic performance using LabImotion (NFDI4Chem) / AdActa (NFDI4CAT) and characterization which is also documented in the electronic lab notebooks. Preferentially, the samples are published in data publications/DOIs.
Key to understanding catalysts are operando studies, often conducted at synchrotron radiation facilities1. For this purpose a workflow similar to the XAS use case has been set up. On-site we are improving the workflow using the catalysis beamline at KIT Light Source (CatAct), e.g. SciCat has been recently implemented based on collaboration in the DAPHNE4NFDI-initiative. We will discuss the present status using the power of the different tools and recent studies including Pt/CeO2 catalysts 2 and coated catalysts 3 for emission control as well as Cu/ZnO-ZrO2 catalysts for the production of green methanol 4. In all cases, in addition to the main paper, electronic supporting informations and data publications of the raw data have been published (e.g. on KITOpen, Zenodo 5) Particularly, the format of the data publications will be discussed and how the data are then uploaded into reference databases.

1 Sarma, B. B.; Grunwaldt, J.-D., CHIMIA 78 (5), 288–296 (2024). doi:10.2533/chimia.2024.288.
2 Gashnikova, D.; Maurer, F.; Bauer, M. R.; Bernart, S.; Jelic, J.; Lützen, M.; Maliakkal, C. B.; Dolcet, P.; Studt, F.; Kübel, C.; Damsgaard, C. D.; Casapu, M.; Grunwaldt, J.-D.
2024. ACS Catal. 14 (19), 14871–14886 (2024). doi:10.1021/acscatal.4c02077
3 Rang, F.; Delrieux, T.; Maurer, F.; Flecken, F.; Grunwaldt, J.-D.; Hanf, S., Small Science, 4 (12), 2400345 (2024). doi:10.1002/smsc.202400345
4 Schulte, M. L.; Truttmann, V.; Doronkin, D. E.; Baumgarten, L.; Nicolai, A.; Beltran, D. A. M.; Summ, F. J.; Kiener, C.; Warmuth, L.; Pitter, S.; Saraçi, E.; Grunwaldt, J.-D., Angew. Chem. Int. Ed., in press, e202423281 (2025). doi:10.1002/anie.202423281
5 For the KITOpen-raw data publications: https://doi.org/10.35097/ncym5afk9q64kz8x; https://doi.org/10.35097/0k5f6x56ux
https://zenodo.org/records/14054929

Primary authors: MAURER, Florian; GRUNWALDT, Jan-Dierk; Mr BINIYAMINOV, Vitaly (KIT / ITCP)

Co-authors: GAUR, Abhijeet (ITCP, KIT); Dr ZIMINA, Anna (KIT); Prof. HANF, Schirin (KIT); DELRIEUX, Tim (Karlsruhe Institut für Technologie)

Presenter: GRUNWALDT, Jan-Dierk

Session Classification: Poster

Contribution ID: **151** Type: **Use case flash talk (Mon)**

UC2: Correlation spectroscopy - XPCS

Monday 24 March 2025 14:47 (7 minutes)

The DAPHNE4NFDI initiative is transforming data management for photon and neutron science by implementing FAIR (Findable, Accessible, Interoperable, Reusable) principles across various experimental techniques. At the P10 beamline of DESY, we have deployed SciCat for SAXS/XPCS experiments, enhancing metadata ingestion and data accessibility. In parallel, the development of Xana 2.0 introduces modernized XPCS data analysis with HDF5/NeXus support and parallelized I/O for large datasets.

Expanding these efforts to MHz-XPCS at EuXFEL, we focus on establishing this as a routine technique while embedding FAIR principles. Within DAPHNE4NFDI, we are also developing FUSE (FAIR Unified Scientific Environment)—a cloud-based platform that integrates diverse research data, ensuring interoperability and accessibility for XPCS experiments.

Beyond data management, effective outreach is crucial for maximizing impact. In TA4 of DAPHNE4NFDI, we explore Large Language Models (LLMs) for automating science communication, including event summarization and social media engagement. This AI-driven approach optimizes outreach strategies, making research more accessible.

Our contributions from the University of Siegen demonstrate how advanced data solutions and AI-powered communication are shaping the future of photon and neutron science.

Primary authors: RAZA, Agha Mohammad (University of Siegen); LEONAU, Aliaksandr; TOSSON, Amir (University of Siegen); GUTT, Christian (Universität Siegen); Dr OEZTUERK, OEzguel (University of Siegen)

Presenter: RAZA, Agha Mohammad (University of Siegen)

Session Classification: Use-cases (flash talks)

Contribution ID: **152** Type: **Task Area status update (Tue)**

Status Update TA4 (Dissemination and Outreach)

Tuesday 25 March 2025 17:00 (15 minutes)

We will give an overview on activities in Task Area 4. The task area "Outreach and dissemination" aims at informing and educating the DAPHNE4NFDI community by sharing the results obtained from the other task areas and initiate exchange and work out case studies. The starting point is the homepage www.daphne4nfdi.de where the community is informed about upcoming events, about progress of the projects and about helpful tools. This is supported by further tools like LinkedIn or the Zenodo community efforts. Meeting highlights during the past year have been the SRI2024, the user and status meetings, the DPG conferences, strategy workshops as well events of coordinatged programmes such as DFG collaboration research centres, priority programs or graduate schools. With the help of highlights, especially use cases, the engagement of the community in the project is fostered. In addition, we have successfully conducted the DAPHNE4NFDI lecture series. Further important aspects are the development of teaching material for a FAIR data management, lecture material for DAPHNE4NFDI-related topics etc. TA4 further aims at exploring new exciting topics for workshops/trainings, how DAPHNE4NFDI activities can be integrated in summer schools etc. Joint outreach and dissemination strategies are discussed as it requires the attention and input input from all DAPHNE4NFDI participants. The contribution will round up with a view on the next years including "DAPHNE2.0".

Acknowledged contributors: Lisa Amelung, Svenja Hövelmann, Bridget Murphy, Özgül Öztürk, Christian Gutt, Amir Tosson, Frank Weber, Oliver Knodel, Paolo Dolcet and further TA4-participants

Primary authors: SCHNEIDEWIND, Astrid (JCNS); Dr TRAGESER, Christian (JCNS); GRUNWALDT, Jan-Dierk; Mr BINIYAMINOV, Vitaly (KIT / ITCP)

Presenters: Dr TRAGESER, Christian (JCNS); GRUNWALDT, Jan-Dierk; Mr BINIYAMINOV, Vitaly (KIT / ITCP)

Session Classification: Outreach and Internationalization (TA4 and TA5)

Contribution ID: **153** Type: **Talk**

Chemotion ELN and its extensions for interdisciplinary work: How to achieve both standardization and flexibility at the same time

Tuesday 25 March 2025 11:45 (10 minutes)

The electronic laboratory notebook (ELN) Chemotion was developed to meet the specific needs of scientists in the field of chemistry regarding documentation, process description, and data analysis. Due to the interdisciplinary nature of many research groups in chemistry, various modules have been created in the past, enabling the use of the ELN in related fields and supporting interdisciplinary work far beyond chemistry. The greatest challenge in developing these modules, which extend beyond the traditional applications of chemistry, lies in the standardization and structuring of new content. To achieve standardization and structuring, close collaboration between developers and domain experts is initially required, followed by implementation within the broader scientific community. This article presents highlights from recent months that exemplify the expansion of chemistry-specific software and infrastructure into related scientific disciplines and showcase opportunities for integrating custom content. Most of the developments are in use within the community of NFDI4Chem and could also offer benefits for scientists assigned to NFDI4Cat and DAPHNE4NFDI.

Primary authors: Ms LIN, Chia-Lin; Mr TARIQ, Hadi; Dr JUNG, Nicole (Karlsruhe Institute of Technology); Ms HUANG, Pei-Chi; Mr TREMOUILHAC, Pierre; Prof. BRÄSE, Stefan

Presenter: Dr JUNG, Nicole (Karlsruhe Institute of Technology)

Session Classification: Talks - Collaboration with/in NFDI

Contribution ID: **154** Type: **Poster**

Streamlining Data Validation and Standardization with JSON Schemas

Our digital labbook collects diverse user-submitted data, which demands a validation and standardization approach that is both flexible and robust. We leverage JSON Schemas as the backbone of our system to enforce data validity while accommodating a wide range of input formats. In this talk, we provide an primer on JSON Schemas and demonstrate how they are used to standardize user-facing endpoints. Attendees will gain insights into schema design, validation mechanisms, and practical strategies for managing heterogeneous data in dynamic environments.

Primary author: MOHR, Sebastian Bernd (Georg-August-Universität Göttingen)

Co-author: Dr OSTERHOFF, Markus (Röntgenphysik Göttingen)

Presenters: Dr OSTERHOFF, Markus (Röntgenphysik Göttingen); MOHR, Sebastian Bernd (Georg-August-Universität Göttingen)

Session Classification: Poster

Contribution ID: **155** Type: **Facility update (Wed)**

ESRF Data Strategy (Participant)

Wednesday 26 March 2025 10:10 (20 minutes)

The ESRF Data Strategy aims at fully exploiting the potential of the 4th generation ESRF-EBS towards a fully data-centric approach. More specifically, it includes more efficient tools for data processing and data analysis, further development of metadata for an increased usability of data sets together with improved automation workflows and AI exploitation of data, as well as increased data FAIRness resulting in an increased re-use of data sets.

I will present the ESRF data strategy implementation status, introducing in particular recent software developments and how they fit into a consistent approach encompassing the entire data life cycle from beamline control (BLISS), advanced detector management (Lima2), e-Logbook, workflow management including feedback loops (EWOKS), remote data analysis (VISA), data catalogue (DRAC).

I will also briefly discuss the present status of the involvement of the ESRF in DAPHNE4NFDI. The ESRF is going to coordinate the future Photon and Neutron EOSC node which might offer new opportunities for increased collaboration with DAPHNE4NFDI and German National photon and neutron sources.

Primary author: DAILLANT, Jean (ESRF)

Presenter: DAILLANT, Jean (ESRF)

Session Classification: Facility Updates

Contribution ID: **156** Type: **Talk**

SciCat: the way forward

Tuesday 25 March 2025 09:45 (15 minutes)

In this talk, we will present the SciCat project and its community, together with the core concepts adopted in the project. We will provide an update on the latest functionalities recently added, including the ones directly supported by the "DAPHNE4NFDI contribution for SciCat"effort. We will touch on the effort to improve documentation and to lower the barrier to adopt and deploy SciCat. Next, we will illustrate the project vision including the objectives to keep the project up to date with the latest technologies, improve user experience and foster innovation. We will conclude with few use cases illustrating best-practices on how to better support your data and metadata with SciCat.

Primary author: NOVELLI, Max (European Spallation Source)

Presenter: NOVELLI, Max (European Spallation Source)

Session Classification: Metadata, Catalogues and Software (TA1, TA2 and TA3)

Contribution ID: **157** Type: **Poster**

Industry Engagement: How DAPHNE4NFDI and Industry Can Mutually Benefit

Effective Research Data Management (RDM) and the adoption of FAIR principles are transforming the research landscape. But how can the DAPHNE4NFDI-industry partnership leverage its unique strengths to benefit from these developments?

This planned white paper will explore how structured RDM strategies at large-scale photon and neutron facilities, particularly through DAPHNE4NFDI, can bridge the gap between academia and industry. We examine the evolving role of the NFDI initiative, its collaboration with industry partners, and the potential for monetizing research data.

While the primary goal of FAIR data is to enhance accessibility and collaboration, and raising awareness of implementing RDM measures in PaN research. Here, we explore how well-managed data can create added value, unlock business opportunities, and drive financial returns for industry stakeholders. Additionally, we highlight the potential benefits of this collaboration for the NFDI, paving the way for further development of its best-practices. Through expert insights, discussion and case studies, we identify pathways to transform well-managed research data into a valuable, sustainable asset that benefits both research and industry.

Primary author: ASHTIANI ABDI, Ali

Co-authors: MURPHY, Bridget; GRUNWALDT, Jan-Dierk (KIT)

Presenter: ASHTIANI ABDI, Ali

Session Classification: Poster

Contribution ID: **160** Type: **Use case flash talk (Mon)**

UC5: Spectroscopy

Monday 24 March 2025 15:01 (7 minutes)

Inelastic scattering is a fundamental technique for probing lattice dynamics and magnetic excitations, serving as a cornerstone in materials science and condensed matter research. While it has historically been linked to neutron scattering, recent innovations in modern synchrotron facilities have facilitated the acquisition of complementary x-ray scattering data, which are vital for contemporary scientific investigations. This encompasses experiments conducted under high-pressure conditions on small single crystals and within materials that demonstrate substantial neutron absorption.
In use case 5, we aim to establish suitable metadata vocabularies specifically for Inelastic Neutron Scattering (INS) and Inelastic X-ray Scattering (IXS) techniques, alongside fostering the use of electronic laboratory notebooks (ELNs). This poster demonstrates our efforts in developing metadata schemas for a triple-axis spectrometer and provides an update on the status of the electronic lab notebook being developed at MLZ.

Primary author: TYMOSHENKO, Yuliia (KIT, IQMT)

Co-authors: SCHNEIDEWIND, Astrid (FZ Jülich); BAUDISCH, Josef (MLZ TUM); LOHSTROH, Wiebke; WEBER, Frank (Karlsruhe Institute of Technology)

Presenter: TYMOSHENKO, Yuliia (KIT, IQMT)

Session Classification: Use-cases (flash talks)

Contribution ID: **161** Type: **Use case flash talk (Mon)**

UC7: Ultrafast / Magnetic x-ray scattering

Monday 24 March 2025 15:32 (7 minutes)

Understanding ultrafast structural and magnetic dynamics in materials is essential for advancing fields such as spintronics, quantum materials, and laser-driven phase transitions. Time-resolved X-ray scattering techniques at free-electron lasers (XFELs) provide a powerful tool to investigate femtosecond-scale structural and spin-related phenomena in thin films and nanostructures. In particular, the combination of ultrashort infrared (IR) laser pulses as the pump and X-ray free-electron laser (FEL) pulses as the probe enables the direct observation of transient changes in magnetic order and lattice structure on timescales shorter than 100 fs.

In a typical pump-probe experiment, an intense IR laser pulse excites the sample, inducing electronic and structural dynamics, while a delayed FEL pulse probes the resulting changes via X-ray diffraction or resonant magnetic scattering. By systematically varying the time delay between the pump and probe, it is possible to capture the evolution of magnetization, spin textures, and lattice distortions with femtosecond temporal resolution. This approach has been instrumental in revealing light-induced phase transitions, demagnetization dynamics, and the coupling between electronic, magnetic, and structural degrees of freedom.

Our investigation utilizes the European XFEL (Hamburg) and FERMI@ELETTRA (Italy) for time-resolved X-ray scattering experiments to study ultrafast dynamics. Elettra's current Data Policy based on the FAIR principles, which the experimental data and metadata of peer-reviewed experiments are stored in an online catalogue, accessible to registered users after a three-year embargo period. For the detailed analysis, we consider experimental data collected by several scientific groups over several years. Collecting all data in the same format for comparison is a critical step in our research.

To enhance data processing and workflow efficiency, we are also working on Data And Metadata iNspection Interactive Thing (DAMNIT) for the HED beamline at EuXFEL. DAMNIT aims to provide an automated pipeline designed to streamline data organization, processing, and metadata handling for ultrafast X-ray experiments, ensuring robust data handling, metadata extraction, and rapid feedback for experiment optimization. Special consideration is given to the specific requirements of pump-probe experiments, where data acquisition must be synchronized with the optical laser and collected when the shutter is open. Therefore, the system is optimized to address these needs effectively. By integrating DAMNIT with our workflows, we aim to improve data accessibility, reproducibility, and compliance with FAIR principles within DAPHNE4NFDI.

Primary authors: GUTT, Christian (Universität Siegen); Dr OEZTUERK, OEzguel (University of Siegen)

Presenters: GUTT, Christian (Universität Siegen); Dr OEZTUERK, OEzguel (University of Siegen)

Session Classification: Use-cases (flash talks)

Contribution ID: **162** Type: **Facility update (Wed)**

Facility update from European XFEL

Wednesday 26 March 2025 09:20 (20 minutes)

Progress of work by colleagues from European XFEL focused on TA3 while affecting the other task areas as well. Developments of the DAMNIT system for near-online extraction and presentation of metadata and orchestration of data processing/analysis pipelines included the establishment of a web-frontend version, the preparation of an infrastructural change for a centralized database, and the prototyping of a data export interface, with the pioneering use case of communication with the Open-XPCS platform from Siegen university.

Other advancements presented will be the expanded treatment of PaNET terms for experiment techniques that are now also integrated to the data acquisition systems of our scientific instruments, the implementation of data management plans and their link to the Electronic Lab Notebook and other services at our facility, and finally work on a metadata collection, conversion and validation workflow for raw crystallographic diffraction data, which is expected to be of high interest to the DAPHNE4NFDI community.

Primary authors: DALL'ANTONIA, Fabio (European XFEL GmbH); CARINAN, Cammille (European XFEL); GELISIO, Luca (European XFEL); DO CARMO TEODORO, Matheus

Presenter: DALL'ANTONIA, Fabio (European XFEL GmbH)

Session Classification: Facility Updates

Contribution ID: **163** Type: **Talk**

Base4NFDI

Tuesday 25 March 2025 11:55 (10 minutes)

Base4NFDI is the joint initiative of all 26 NFDI consortia and supports the development of basic services as common, interoperable solutions. This flashtalk is a short recap of two presentations given in advance on February 19 (https://doi.org/10.5281/zenodo.14894128) and February 25 (https://doi.org/10.5281/zenodo.14930141). It will quickly introduce the project along with four of its currently funded basic services:
PID4NFDI - a centralized infrastructure service for managing persistent identifiers within NFDI
DMP4NFDI - our basic service for data management plans (DMPs) and software management plans (SMPs) across NFDI
TS4NFDI - A Cross-Domain Terminology Service for NFDI
KGI4NFDI - Enabling Knowledge Graph Infrastructure for NFDI

Primary authors: TATSCHECK, Jana; RETTBERG, Najla

Presenter: RETTBERG, Najla

Session Classification: Talks - Collaboration with/in NFDI

Contribution ID: **164** Type: **Poster**

DECTRIS CLOUD - Scalable research, collaborative discovery, and reliable results

DECTRIS Cloud is a web-based platform designed to accelerate scientifi c research by providing seamless integration of custom and sharable software, ready-to-use workfl ows, and powerful collaboration tools. Researchers can develop and share tailored software, while leveraging a comprehensive set of pre-built workfl ows and tools. This signifi cantly reduces setup time, allowing scientists to focus on generating insights and producing reproducible results, instead of managing complex confi gurations.

The platform facilitates real-time collaboration, enabling researchers to eff ortlessly share data, software, and workspaces with just a click, fostering effi cient co-creation across global teams. Researchers maintain full control over isolated workspaces, ensuring both security and adaptability to the specifi c needs of their projects. By off ering unparalleled autonomy and eliminating computational limitations, DECTRIS CLOUD empowers scientists to explore new scientifi c frontiers, collaborate with ease, and accelerate research.

Primary authors: LARSEN, Camilla (DECTRIS); LEROY, Ludmila (DECTRIS); BURIAN, Max (DECTRIS)

Presenters: LARSEN, Camilla (DECTRIS); LEROY, Ludmila (DECTRIS); BURIAN, Max (DECTRIS)

Session Classification: Poster

Contribution ID: **165** Type: **Talk**

Workshops - Collaboration in NFDI - Round 1

Tuesday 25 March 2025 12:15 (45 minutes)

Table 1: Physical Sciences in NFDI
Table 2: DAPHNE4NFDI integration to Base4NFDI (and vice versa)
Table 3: Quality insurance for public data bases & Quality control for curated reference databases
Table 4: LLMs
Table 5: Data processing/analysis workflows

Session Classification: Workshops/round tables - Collaboration with/in NFDI

Contribution ID: **166** Type: **Talk**

Workshops - Collaboration in NFDI - Round 2

Tuesday 25 March 2025 13:00 (45 minutes)

Table 1: Physical Sciences in NFDI
Table 2: DAPHNE4NFDI integration to Base4NFDI (and vice versa)
Table 3: Quality insurance for public data bases & Quality control for curated reference databases
Table 4: Machine learning
Table 5: Data processing/analysis workflows

Session Classification: Workshops/round tables - Collaboration with/in NFDI

Contribution ID: **167** Type: **Facility update (Wed)**

Facility update from European Spallation Source (ESS) (Participant)

Wednesday 26 March 2025 10:30 (15 minutes)

The European Spallation Source (ESS) aims to fully leverage the potential of its state-of-the-art neutron facility through a data-centric approach. This includes advanced tools for data acquisition, processing, and analysis, enhanced metadata frameworks for improved data usability, increased automation in workflows, AI-driven data insights, and strengthened adherence to FAIR principles to maximize data reusability.
Recent software developments supporting this strategy span the entire data life cycle, including the SciCat data portal, remote data analysis platforms, and automated workflow solutions. Additionally, ESS is actively engaged in European data initiatives, including collaborations within the EOSC framework and PaNOSC, fostering synergies for open science and data interoperability.
I will present the envisioned ESS data journey, highlighting our software components and their integration across the entire data life cycle, from proposal to publication.

Primary author: BOLMSTEN, Fredrik (European Spallation Source (ESS))

Presenter: BOLMSTEN, Fredrik (European Spallation Source (ESS))

Session Classification: Facility Updates

Contribution ID: **168** Type: **Talk**

Presentation of round tables and further instructions for workshop

Tuesday 25 March 2025 12:05 (10 minutes)

Presenter: AMELUNG, Lisa (DAPHNE4NFDI | DESY)

Session Classification: Talks - Collaboration with/in NFDI

Contribution ID: **169** Type: **Poster**

DAPHNE4NFDI: DAta from PHoton and Neutron Experiments for NFDI

Come and see the most recent overview of work of the consortium DAPHNE4NFDI!
Note: This poster was shown at the DFG Interim report symposium in Bonn on 14th of February 2025.

Thanks to all of you who have contributed!

This work was supported by the consortium DAPHNE4NFDI in the context of the work of the NFDI e.V. The consortium is funded by the DFG - project number 460248799. In-kind support by all partners is gratefully acknowledged.

Primary authors: MURPHY, Bridget; SCHNEIDEWIND, Astrid (FZ Jülich); LOHSTROH, Wiebke; GUTT, Christian (Universität Siegen); AMELUNG, Lisa (DAPHNE4NFDI | DESY); BUSCH, Sebastian (GEMS at MLZ, Helmholtz-Zentrum Hereon, Germany); GRUNWALDT, Jan-Dierk (KIT)

Presenters: MURPHY, Bridget; SCHNEIDEWIND, Astrid (FZ Jülich)

Session Classification: Poster

Contribution ID: **170** Type: **Use case flash talk (Mon)**

UC1: X-ray imaging

Monday 24 March 2025 14:40 (7 minutes)

Advances in synchrotron imaging and computed tomography (CT) enable non-invasive, high-resolution 3D visualization of samples, including biological tissues, surpassing standard methods. This innovation generates massive data volumes, requiring tailored solutions for data capture, management, storage, and repositories for processed data and analysis code. These solutions enhance transparency and maximize data reuse in biomedical research. We have been collaborating closely with scientists and IT experts from various synchrotron facilities, primarily at ESRF and PETRA III. We have established a consistent vocabulary and compiled a comprehensive list of parameters applicable across different imaging beamlines. The results of this work will be shared with other imaging beamlines and facilities. The LMU team, together with the tomography-user teams of Göttingen University and HZB, has finalized a metadata table and its description for generic (bio)synchrotron imaging and CT experiments for the white paper. Currently, the implemented or in-progress metadata recording is focused mainly on storing parameters up to the data acquisition phase. In collaboration with ESRF IT experts, we are currently discussing the extension of this list to include comprehensive post-processing metadata, at least up to CT data reconstruction, which is a common component of all CT experiments. Additionally, ESRF does not provide detailed sample metadata information. We are in the process of discussing with the ESRF IT team to integrate the IGSN link with ICAT to connect the corresponding experiments. The subsequent steps of image processing, data analysis, and quantification require significant collaboration within the community, as each application demands specific processing pipelines and tools. We aim to identify the common components of this workflow and propose a suitable metadata scheme. As part of our commitment to TA3, we are developing an image processing workflow tailored for the analysis of biological tissues, incorporating machine learning-based models for segmentation, analysis, and quantification. This processing pipeline will be shared with the broader user community (https://github.com/hfahad/U-NET-Synchrotron_CT-image_segmentation.git). Through collaboration among user groups, scientists, and IT experts from synchrotron facilities, the LMU team is working to establish a general FAIR data workflow for synchrotron imaging and CT experiments. The LMU team has finalizing the installation process of SciCAT on the LMU machine with coordinated efforts from LMU IT Services, the Leibniz Supercomputing Centre (LRZ), and the Helmholtz Centre Dresden-Rossendorf. Currently, SciCAT is undergoing further processing and will be initially available for our group and subsequently for our collaborators. The purpose of this implementation is to install and interface SciCAT with the existing catalog system, enhancing overall data management.

Primary authors: FAHAD, Hafiz Muhammad (LMU); COAN, Paola (Ludwig Maximilian University of Munich University in Germany)

Presenter: FAHAD, Hafiz Muhammad (LMU)

Session Classification: Use-cases (flash talks)

Contribution ID: **171** Type: **Talk**

Award Ceremony - Best Poster and Talk

Tuesday 25 March 2025 17:30 (10 minutes)

Contribution ID: **172** Type: **Facility update (Wed)**

Facilities: Welcome at HZB

Monday 24 March 2025 14:05 (20 minutes)

This talk will give an short overview what is brewing at HZB: it will cover the light source BESSY II and its ongoing uprade progam next to a short introduction of BESSY III. Then it will address the its data side:
data taking and processing illustrated by projects running at HZB.

Primary authors: RIES, Markus (Helmholtz-Zentrum Berlin); SCHNIZER, Pierre

Presenters: RIES, Markus (Helmholtz-Zentrum Berlin); SCHNIZER, Pierre

Session Classification: Facility Updates

Contribution ID: **173** Type: **not specified**

Welcome by DAPHNE4NFDI

Monday 24 March 2025 14:00 (5 minutes)

Presenters: KANNGIESSER, Birgit (TU Berlin); SCHLESIGER, Christopher (Technische Universität Berlin); GÖRZIG, Heike (HZB); KRAHL, Rolf (Helmholtz-Zentrum Berlin für Materialien und Energie); AMELUNG, Lisa (DAPHNE4NFDI | DESY)

Contribution ID: **174** Type: **Talk**

Status Update TA5

Task Area 5. The task area "External communication and policy"aims at defining common data policies, use cases and pilot workflows and standardised best practices with the aim of agreeing upon common standards. TA5 also encompasses cooperation with the other NFDI consortia, which are connected either by similar scientific questions and/or by issues of data management. Here we will also seek to coordinate and communicate with European user organisations, such as ESUO and ENSA, and the consortia of facilities, such as LEAPS and LENS. Participation in European projects (PaNOSC, etc.) that are closely linked to DAPHNE is envisaged, through partners or representatives of DAPHNE who are already active as observers.
After an intense discussion process within the European photon and neutron user communities and a joint meeting of their representing associations ENSA and ESUO, the paper „FAIR data –the photon and neutron communities move together towards open science"could be finished in April 2024. The publication by IUCrJ 01/2025 is well acknowledged by more than 1760 visits up to now.
On national level, the data policies of DESY and MLZ were adapted along with FAIR and open science principles, ready to be agreed by the director's boards.
The exchange within NFDI is an ongoing process, with activities of the consortia "Physics related sciences in NFDI"by frequent meetings, continuous exchange and a jointly organized series of related talks. DAPHNE4NFDI was invited to join similar activities of the Engineering sciences in NFDI, e.g. a joint contribution to the CoDRDI2025 conference. The data policies of DESY and MLZ were adapted along with FAIR and open science principles, ready to be agreed by the director's boards. Interactions with other consortia –FAIRmat, 4Chem, 4Cat, and PUNCH are successfully established and now continued along topical and scientific issues.
The cooperation with Helmholtz infrastructures could be sustained and will be extended.
Within the EOSC Federation, Daphne4NFDI will be engaged not only as a part of NFDI as a national node, but also in the PaNOSC topical node selected out of 13 example nodes –continuing the established cooperation on European level.

Presenters: SCHNEIDEWIND, Astrid (FZ Jülich); MURPHY, Bridget; GUTT, Christian (Universität Siegen)

Session Classification: Outreach and Internationalization (TA4 and TA5)

Contribution ID: **175** Type: **Poster**

Fair data initiatives at the Hereon Synchrotron beamlines

The Helmholtz-Zentrum Hereon is actively contributing to the FAIR data initiatives within the DAPHNE4NFDI project by implementing data management strategies across its beamlines. In the context of Task Area 1: Managing Data Production, the following efforts are included. A key component of this efforts is the Metadata Collector, which has overcome stability issues and is now successfully deployed at P07, with deployment at P05 micro ongoing as necessary user interface modifications are being implemented. To enhance metadata consistency, work is currently focused on synchronizing metadata schemas between P05 micro, P05 nano, and P07. Additionally, efforts are planned to establish a common metadata acquisition framework across beamlines by defining a shared ontology, metadata schema, and data format, ensuring interoperability and streamlined data management. Furthermore, regarding electronic logbooks, the Scilog system is scheduled for testing at P03 during the next beamtime. Future developments are required to enable automatic data ingestion into Scilog across other beamlines. In Task Area 2: (Meta)data Repositories and Catalogues, the Metadata Catalog is another significant advancement, with Scicat already deployed at P05, incorporating multiple datasets, and plans in place for extension to other beamlines. As part of Task Area 3: Infrastructure for Data and Software Reuse, Pydidas was developed to facilitate X-ray diffraction data analysis. Released as an open-source tool, Pydidas is designed for use both at the facility and the user's home institute, offering flexibility for a variety of diffraction experiments. It enables fast data reduction and analysis during experiments, providing rapid feedback on measurement progress and enhancing the efficiency of data analysis workflows. These initiatives reinforce Hereon's commitment to improving FAIR data principles, within the DAPHNE4NFDI framework, fostering a more efficient and standardized approach to data management.

Primary authors: LIPPMANN, Otto; HAMMEL, Jörg (Hereon)

Presenter: LIPPMANN, Otto

Session Classification: Poster

Contribution ID: **176** Type: **Task Area status update (Tue)**

Status Update TA5 (External Communication and Policy)

Tuesday 25 March 2025 17:15 (15 minutes)

Task Area 5. The task area "External communication and policy"aims at defining common data policies, use cases and pilot workflows and standardised best practices with the aim of agreeing upon common standards. TA5 also encompasses cooperation with the other NFDI consortia, which are connected either by similar scientific questions and/or by issues of data management. Here we will also seek to coordinate and communicate with European user organisations, such as ESUO and ENSA, and the consortia of facilities, such as LEAPS and LENS. Participation in European projects (PaNOSC, etc.) that are closely linked to DAPHNE is envisaged, through partners or representatives of DAPHNE who are already active as observers.
After an intense discussion process within the European photon and neutron user communities and a joint meeting of their representing associations ENSA and ESUO, the paper „FAIR data –the photon and neutron communities move together towards open science"could be finished in April 2024. The publication by IUCrJ 01/2025 is well acknowledged by more than 1760 visits up to now. On national level, the data policies of DESY and MLZ were adapted along with FAIR and open science principles, ready to be agreed by the director's boards.
The exchange within NFDI is an ongoing process, with activities of the consortia "Physics related sciences in NFDI"by frequent meetings, continuous exchange and a jointly organized series of related talks. DAPHNE4NFDI was invited to join similar activities of the Engineering sciences in NFDI, e.g. a joint contribution to the CoDRDI2025 conference. The data policies of DESY and MLZ were adapted along with FAIR and open science principles, ready to be agreed by the director's boards. Interactions with other consortia –FAIRmat, 4Chem, 4Cat, and PUNCH are successfully established and now continued along topical and scientific issues.
The cooperation with Helmholtz infrastructures could be sustained and will be extended.
Within the EOSC Federation, Daphne4NFDI will be engaged not only as a part of NFDI as a national node, but also in the PaNOSC topical node selected out of 13 example nodes –continuing the established cooperation on European level.

Primary authors: SCHNEIDEWIND, Astrid (JCNS); MURPHY, Bridget; GUTT, Christian (Universität Siegen)

Presenter: SCHNEIDEWIND, Astrid (JCNS)

Session Classification: Outreach and Internationalization (TA4 and TA5)