

Eighth FAIRmat Users Meeting

June 16–17, 2026

Center for the Science of Materials Berlin (CSMB)

Humboldt-Universität zu Berlin, Germany



Schedule Overview

Tuesday, June 16	Wednesday, June 17
	9:00–10:15 Invited Talks 2.049
	10:15–10:45 Coffee Break Foyer
	10:45–12:00 Invited Talks 2.049
12:00–13:30 Registration & Lunch Foyer	12:00–13:30 Lunch & Farewell Foyer
13:30–14:00 Welcome & NOMAD Updates 2.049	
14:00–16:00 Tutorial 1 2.049	14:00–16:00 Tutorial 2 1.021
16:00–16:30 Coffee Break Foyer	
16:30–18:30 Expert Panel 2.049	16:30–18:30 Tutorial 3 1.021
18:30–20:30 Community Exhibit & Dinner Foyer	

The detailed program can be found below and on the [webpage](#).

Welcome and NOMAD Updates



FAIRmat and NOMAD: Recent Updates and Emerging Use Cases

Lukas Pielsticker

Time: 13:30–14:00; Room: 2.049

In this overview talk, I will present recent developments across FAIRmat and the NOMAD ecosystem, highlighting the continued growth of the platform as a research data infrastructure for materials science. FAIRmat now supports a broad range of research activities, from individual projects and laboratory workflows to institutional data management and collaborative research environments.

The presentation will showcase recent advances across the ecosystem, including improvements to data management, metadata handling, workflow support, data exploration, and deployment options. Through selected examples, I will demonstrate how these developments help researchers and organizations manage research data more effectively while enabling FAIR and reproducible research practices.

The talk will conclude with an outlook on the next phase for FAIRmat and NOMAD, highlighting emerging capabilities, expanding application areas, and the growing role of FAIR research data in data-driven and AI-based scientific discovery.

Hands-on Tutorials



Getting Started with NOMAD for Materials-Science Research Data Management

Siamak Nakhaie

Time: 14:00–16:00; Room: 2.049

In this hands-on workshop, you will learn how to efficiently search, filter, and explore data in NOMAD. We will guide you through managing, uploading, and sharing your research data, as well as steps to publish datasets with DOIs to guarantee accessibility and reusability. You will also gain practical experience using NOMAD's ELN functionality to document your research. Please bring your laptop to participate in the interactive exercises.



Deploying NOMAD Oasis with Kubernetes

Lauri Himanen

Time: 14:00–16:00; Room: 1.021

Many existing NOMAD Oasis installations rely on Docker Compose, which has long been the recommended and most widely used deployment approach. While this setup works well for smaller environments, it can become limiting when scaling to larger workloads or more complex infrastructures.

In this workshop, I will introduce an alternative deployment strategy using Kubernetes. This approach enables more flexible scaling of computational workloads across multiple nodes, while also providing robust tools for service lifecycle management, monitoring, and resilience.

The session will include a live demonstration based on a local development setup that participants can replicate on their own laptops. Building on this, we will explore how the same deployment can be extended to a cloud environment, running across multiple virtual machines to achieve a scalable and production-ready system.

By the end of the workshop, attendees will have a practical understanding of how Kubernetes can be used to deploy and manage NOMAD Oasis in both local and cloud-based scenarios.



Parsing in NOMAD: from Raw Files to Structured Entries

Lev Ginzburg and Sarthak Kapoor

Time: 16:30–18:30; Room: 1.021

Parsing in NOMAD is a central component of automated data ingestion, enabling the conversion of raw files into structured entries. This tutorial presents several parsing scenarios that differ in their level of automation and editability of resulting entries, and outlines recommended practices for parser design. The hands-on session will guide participants through the main stages of parser development, from schema creation to parser testing.

Participants are requested to bring a laptop with an IDE (e.g. Visual Studio Code), Git, Python, and a GitHub account.

Expert Panel: Deploying RDM Infrastructure across Research Organizations Using NOMAD Oasis



Joseph Rudzinski

Time 16:30–18:30; Room: 2.049

This session is aimed at decision-makers who are evaluating, planning, or operating Research Data Management (RDM) infrastructure. Using NOMAD Oasis as a concrete example, the discussion will focus on the strategic, organizational, and practical realities of deploying RDM platforms. It will also provide those new to NOMAD Oasis with a grounded introduction to how such systems are being adopted and operated across real institutional settings.

The panel brings together speakers with experience of RDM adoption across a range of institutional contexts, from individual research groups to cross-institutional networks and industry collaborations. Together, they will offer perspectives on how organizations approach RDM adoption under different scientific, organizational, and resource constraints.

The discussion will focus on questions that matter most in practice: how deployments are scoped and governed, what investments are required, how sustainability is maintained over time, and how institutions balance centralized coordination with distributed access and autonomy. Technical details will be addressed only where they directly inform strategic or organizational decisions. The session will combine short flash talks, a moderated panel discussion, and open audience Q&A to encourage exchange between organizations at different stages of their RDM journey.

Users-Meet-Developers Exhibit

Time: 18:30–20:30; Room: Foyer

Presenter	Title of contribution
Abbas Gholami	<i>Scaling NOMAD Adoption Across Research Communities: Experiences from Supporting a Large Research Network</i>
Alexander Fuchs	<i>Controlling Experiments and Recording FAIR Data with NOMAD CAMELS</i>
Ana Velazquez	<i>Customized Digital Tools and DM Workflows towards Data-driven Catalytic Materials Discovery</i>
Bernadette Mohr	<i>FAIR Molecular Dynamics Data</i>
Dimitra Gkogkou	<i>Digital Governance for Printed Perovskite Solar Cells in PERSEUS Project</i>
Elaheh Saadat	<i>AI Automation for Parsing Data from Lab to NOMAD</i>
Erik Wutke	<i>Visual Analytics of Multimodal Data during Beamtime at the Liquid Metallet Source at HZB</i>
Jose Pizarro	<i>NERxiv: a Tool for LLM-based Automated Extraction and Standardization</i>
Lena Mittmann	<i>NOMAD Oasis at DTU</i>
Markus Kühbach	<i>Electron Microscopy, Atom Probe, and Microstructures in NOMAD</i>
Martin Kuban	<i>Data Analysis with the MADAS Framework</i>
Nathan Daelman	<i>The Calculations App in NOMAD v2</i>
Pavel Ondračka	<i>Pitfalls in Mining NOMAD for Machine-learning Potentials – Lessons from a Silicon Case Study</i>
Rubel Mozumder	<i>Experimental Data Conversion for NOMAD with pynxtools Plugins (e.g. pynxtools-spm)</i>
Sandor Brockhauser	<i>NOMAD at MPI CBS</i>
Santa Pile	<i>MaMMoS: Ontology-linked FAIR Workflows from Magnetometry Data to Sensors and Other Magnetic Devices Optimisation</i>
Šimon Kratochvíl	<i>Use of Open Data for Training Machine-learning Interatomic Potentials</i>
Sonal Ramesh Patel	<i>A Roadmap for NOMAD Adoption at BESSY II</i>

Invited Talks: NOMAD Applications and Use Cases



Data @Big Chemistry: Deploying NOMAD Oasis as a User Platform

Mark Driver

Time: 9:00–9:25; Room: 2.049

Big Chemistry is a national growth fund project from the Netherlands (<https://bigchemistry.nl>). Big Chemistry aims to establish an autonomous “RobotLab” that combines chemical robotics, artificial intelligence and intensive data collection. We aim to bring a Big Data Driven approach to Chemistry with a focus on applications in formulation science. We thus require a robust data management infrastructure. The consortium includes 5 academic institutions located across the Netherlands and a growing number of industrial partners. NOMAD Oasis was selected as part of our data management solution to enable the sharing of operationally FAIR data which is discoverable and browsable for scientists across the consortium. We present the process of deploying multiple NOMAD Oasis instances as a data platform, and the additional requirements that industrial collaborations place on data management, including additional cyber security considerations.



NOMAD Oasis for Electron Microscopy

Christoph T. Koch

Time: 9:25–9:50; Room: 2.049

Electron microscopes come in many flavors, with instruments ranging from table-top scanning electron microscopes to 4 or more meters tall (scanning) transmission electron microscopes with hundreds of electron optical elements in them. Likewise, the type of data they produce is very diverse, ranging from spectra to diffraction patterns and images, and 3D, 4D, 5D, or higher dimensional hypercubes of data, which are stored in different proprietary formats, depending on the instrument vendor.

In this presentation I will introduce the NOMAD-Oasis setup we run within the structure research and electron microscopy group at the physics department of HU Berlin. I will highlight the pynxtools-em software which allows us to convert data from different vendors to an open data standard, which NOMAD is able to display, independent of the dimensionality of the data. I will also show how the

NOMAD Remote Tools Hub (NORTH) is utilized to perform complex data processing tasks of data sets being several tens of GB in size, and incorporating high-performance computing, directly on the server – including tools which can be controlled using natural language.



NOMAD Adoption Across Research Communities

Joseph Rudzinski

Time: 9:50–10:15; Room: 2.049

FAIRmat and NOMAD provide a powerful foundation for FAIR and AI-ready research data management. However, translating these capabilities into real research workflows often requires additional integration, infrastructure, onboarding, and workflow adaptation efforts. In this talk, we present practical experiences from Glaide Data GmbH supporting NOMAD adoption across research groups, consortia, and emerging communities beyond FAIRmat's core domains. We discuss strategies for integrating NOMAD into existing research environments, evaluating and adapting community plugins and schemas, and accelerating adoption through practical implementation, infrastructure support, training, and onboarding. Through concrete examples, we highlight how these efforts can help extend the impact and reach of the FAIRmat ecosystem.



Customizing NOMAD Oasis for Thin-film Research in Sustainable Energy Materials

Diego Alejandro Garzon

Time: 10:45–11:10; Room: 2.049

Research at the Sadewasser Research Group focuses on sustainable energy materials for applications in photovoltaics and batteries. These research activities involve a broad range of deposition techniques, wet and dry laboratory processes, and advanced characterization methods, generating heterogeneous datasets that must be organized in a consistent and traceable manner.

This presentation describes the development of a custom NOMAD Oasis plugin created to support the organization of experimental activities within our group. The plugin was designed to reflect the structure of our laboratory practices and to facilitate the integration of data and metadata associated with thin-film

fabrication, processing, and advanced characterization across different research lines and applications.

The talk will discuss the practical process of translating laboratory procedures and experimental relationships into NOMAD schemas and plugin components, including approaches for metadata organization, representation of experimental processes, and adaptation of the Oasis environment to the day-to-day needs of experimental research.

In addition, the presentation will explore the use of AI-assisted programming tools during the implementation of the plugin. By integrating large language model-based coding assistants into the development process, it became possible to quickly extend custom NOMAD functionality with a relatively small development effort. A short live demonstration will illustrate how AI-assisted development can support the customization of NOMAD Oasis environments for specific experimental contexts.



Automating Research Workflows with NOMAD Actions

Ahmed Ilyas, Lev Ginzburg, and Sarthak Kapoor

Time: 11:10–11:35; Room: 2.049

NOMAD Actions provide a framework for executing complex or resource-intensive workflows within NOMAD. They are ideal for tasks that require special resources like GPUs, have long execution times, or need to interact with external systems, offering a more powerful alternative to standard normalizers.

Built as plugin entry points, Actions use Temporal.io to orchestrate workflows made up of reliable, self-contained steps. Temporal handles the hard distributed computing problems, such as fault tolerance, retries, and scaling, so that the plugin authors can focus on the core scientific logic. The resulting workflows can be triggered and inspected from the NOMAD GUI.

We illustrate Actions with examples from automated data processing and advanced ML workflows. These example Actions are already deployed in lab environments, supporting ML-assisted characterization, generative AI model sampling, and large-scale data transformation.



AI-Assisted Scientific Data Extraction

Sherjeel Shabih and Yaru Wang

Time: 11:35–12:00; Room: 2.049

Scientific progress increasingly depends on the ability to transform unstructured information into accessible, structured data. However, the rapid growth of scientific literature has made manual data extraction and curation a major bottleneck across disciplines. Recent advances in artificial intelligence offer new opportunities to automate this process and unlock knowledge at unprecedented scale.

This session will explore emerging AI-assisted approaches for scientific data extraction, focusing on multimodal workflows that convert diverse information sources into structured, machine-readable datasets. We will present recent developments in NOMAD that leverage large language models and domain-specific validation to extract scientific information from research publications, enabling the creation of continuously updated knowledge resources. In addition, we will showcase new capabilities that extend data extraction beyond traditional documents, including the use of speech and audio inputs as alternative pathways for capturing and structuring scientific knowledge.

Through examples from materials science and photovoltaics, we will discuss the opportunities, challenges, and limitations of AI-driven extraction systems, including issues of accuracy, validation, reproducibility, and integration with existing scientific infrastructures. The session aims to provide researchers with an overview of how AI can accelerate the transformation of scientific content into reusable data and support data-driven discovery in an era of rapidly expanding scientific output.

Venue

Center for the Science of Materials Berlin (CSMB)

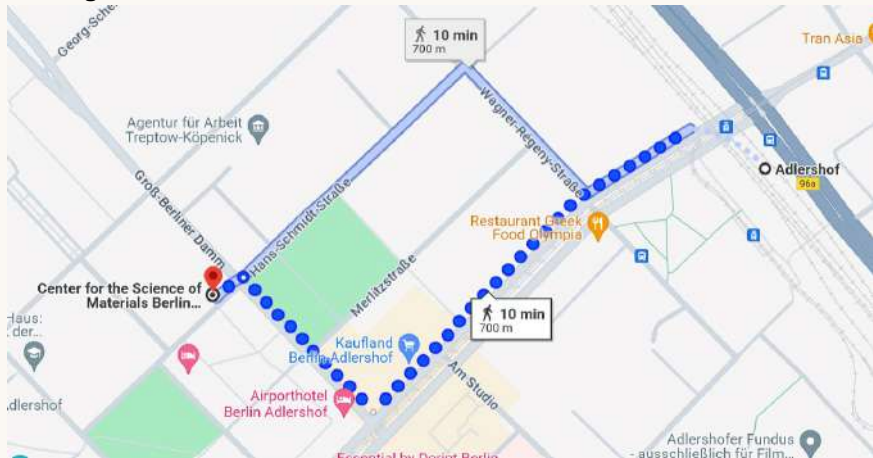
Humboldt-Universität zu Berlin

Zum Großen Windkanal 2,

12489 Berlin, Germany

Public transport: S Adlershof (Lines: S8, S85, S9, S45, S46)

Walking from S Adlershof:



Stay in Touch

General organizational questions:

- Natalia Slepenco: natalia.slepenco@physik.hu-berlin.de



