Contribution ID: **71**   Type: **Talk**

Taming Data from Simulations and Experiments for FAIR Science - From Scattered Files to Structured Insights with NOMAD

Saturday 20 September 2025 14:55 (12 minutes)

Simulations and experimental measurements in physics and materials science routinely generate terabytes of heterogeneous data, much of which remains inaccessible beyond the original study. The widespread use of proprietary or program-specific file formats, combined with the failure to supply relevant metadata, reduces the reproducibility and extendability of available data, thereby violating the FAIR principles of Findability, Accessibility, Interoperability, and Reusability.

NOMAD (nomad-lab.eu), the open-source data infrastructure developed within the German NFDI consortium FAIRmat, addresses these challenges by implementing a consistent and extensible framework for FAIR data management across computational and experimental materials science.[1,2] Recent developments expand its scope beyond electronic-structure simulations and experimental techniques to include classical molecular dynamics, free energy calculations, and multiscale modeling. Its structured, plugin-based architecture for custom parsers, metadata schemas, and workflow management facilitates provenance tracking across various data sources, allowing, e.g., for the representation of force fields, molecular topologies, and hierarchical systems. Integrated tools coupled with an extensive API provide convenient access to curated, AI-ready datasets for secondary analysis.

[1] Scheidgen, M. et al., JOSS 8, 5388 (2023).
[2] Scheffler, M. et al., Nature 604, 635-642 (2022).

Primary author: MOHR, Bernadette (FAIR)

Presenter: MOHR, Bernadette (FAIR)

Session Classification: Parallel

Track Classification: Other topic/subfield (please comment in the box below)