Beitrag ID: **34**   Typ: **Talk**

MOLSIM: An Interoperable Ontology for Representing Biomolecular Simulation

Molecular dynamics (MD) simulations generate vast amounts of data foundational to structural biology, yet their value is often limited by inconsistent metadata and software-specific formats that create isolated data silos. To address this "semantic gap," we introduce MOLSIM, an interoperable ontology designed to formalize the description of atomistic biomolecular simulations and enhance the implementation of FAIR (Findable, Accessible, Interoperable, Reusable) principles. Developed in adherence to Open Biological and Biomedical Ontologies (OBO) Foundry principles, MOLSIM prevents redundancy by systematically reusing terms from established ontologies such as ChEBI and the Unit Ontology. A core feature of MOLSIM is its software-agnostic labeling, which resolves ambiguities in simulation metadata; for instance, mapping disparate keywords like 'ntt=1' in AMBER and 'tcoupl' in GROMACS to a unified Berendsen Thermostat class. The ontology was constructed using a Large Language Model (LLM)-assisted workflow, employing LLM to extract technical terms from software manuals, followed by rigorous expert curation. Currently comprising approximately 2,000 terms, MOLSIM enables simulation data to be structured as Knowledge Graphs. This allows for the seamless integration of MD metadata with external open knowledge bases such as Wikidata, UniProt, and the PDB, providing the necessary semantic granularity to support next-generation community repositories.

ONLY WORKSHOPS - Proposed interaction format

Alternative Track

1. Harmonisation of Metadata: Closing Semantic Gaps

ONLY WORKSHOPS - Tentative audience

ONLY WORKSHOPS - Maximum number of participants

ONLY WORKSHOPS - Special technical requirements

Autoren: KRANZ, Angela (FZ-Jülich); USADEL, Björn (FZ-Jülich); MUSYAFFA, Fathoni (FZ-Jülich); DÖRPHOLZ, Hannah (FZ-Jülich); GOHLKE, Holger (FZ-Jülich); BONUS, Michele (HHU-Düsseldorf); GENTILE, Rocco (HHU-Düsseldorf)

Vortragende(r):   MUSYAFFA, Fathoni (FZ-Jülich)

Sitzung Einordnung: TALK SESSION

Track Klassifizierung: HMC Conference 2026 Track Topics: 3. Ontology-Driven Metadata Harmonization: Closing Semantic Gaps