

Ontologies in Computational Materials Science: The NOMAD experience

Maja-Olivia Lenz-Himmer^{1,2}, Luca M. Ghiringhelli^{1,2,3}, Carsten Baldauf¹,
and Matthias Scheffler^{1,2}

¹ Fritz Haber Institute of the Max Planck Society, Berlin, Germany

² Physics Department and IRIS Adlershof, Humboldt-Universität zu Berlin, Germany

³ ghiringhelli@fhi-berlin.mpg.de, <https://th.fhi.mpg.de/>

Key Words: *Ontology, hierarchical metadata, electronic-structure codes, force-field codes.*

With the tremendous increase in the amount of data in materials science, new ways to store and annotate data are necessary to fulfill the FAIR principles – and to do efficient, good, and new science. Consequently, ontologies have been of increased interest as they do not only allow storing and annotating but also semantically linking data even across domains.

The Novel-Materials Discovery (NOMAD) Repository is the largest database in materials science and provides a normalized, source-independent form of these data in the NOMAD Archive using the NOMAD MetaInfo [1] as metadata schema. The NOMAD MetaInfo includes a number of relations between concepts and therefore already goes beyond the simple metadata concept. We have converted it to an ontology and extended it to increase semantics based on the European Materials and Modeling Ontology (EMMO).

Furthermore, within the NOMAD ecosystem, we have created an ontology collection covering materials structures and properties in a more general semantic way. We demonstrate how this may enable connecting multiple sources of knowledge and allow for semantic searches.

REFERENCES

[1] L. M. Ghiringhelli *et al.*, Towards efficient data exchange and sharing for big-data driven materials science: metadata and data formats, *npj Comput. Mater.* **3**, 46 (2017). DOI: 10.1038/s41524-017-0048-5