Ontologies in Computational Materials Science

Abstract for FAIR-DI conference

3-5 June, 2020

Maja-Olivia Lenz, Luca M. Ghiringhelli, Carsten Baldauf, Matthias Scheffler

May 6, 2020

With the tremendous increase in the amount of data in materials science, new ways to store and annotate data are necessary to ensure fulfilling 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. This way data is represented in a machine-readable fashion which opens up new application possibilities, for example in interdisciplinary research, to increase the reusability of data, or simply asking complex questions. The European Materials and Modeling Ontology, EMMO (`http://emmo.info/`), is an attempt to provide a standard representational framework for the physical sciences. However, appropriate ready-to-use domain ontologies are so far completely lacking in the field of materials science. There are several large databases for computational material data each adopting their own meta data schemes for data annotation. The largest is the NOMAD Repository that has most other relevant databases in the field included. Furthermore, the NOMAD Archive provides a normalized form of these data independent of their source using the NOMAD Metainfo [1] as metadata schema. The NOMAD Metainfo includes a number of relations between concepts and therefore goes beyond the simple metadata picture. We are converting it to the ontology format OWL and demonstrate how this enables connecting multiple sources of knowledge. Within the NOMAD ecosystem, we have created a Crystal Structure Ontology (CSO) in order to represent material, in particular crystal solids, as well as a Materials Properties Ontology (MPO) that semantically describes concepts used by materials scientists. One of the more complex concepts is for example the electronic band structure and its relations to other properties. We demonstrate a first application of this NOMAD ontology triad (Metainfo Ontology, CSO and MPO) and showcase interoperability with external ontologies. As an outlook we discuss ideas how to connect with experimental data through ontologies.

[1] L. M. Ghiringhelli *et al.*, npj Comput. Mater. 3, **46** (2017).