Contribution ID: **124** Type: **TALK**

Using application-level ontologies in materials simulation workflows

Tuesday 5 November 2024 13:40 (20 minutes)

The microstructure of materials is characterized by crystallographic defects, which ultimately determine the material properties. In computational materials science, methods and tools are used to predict and analyze defect structures. The increase of computational power has led to the generation of large amounts of complex and heterogeneous data, increasing the need for the implementation of data-driven approaches. In the field of atomistic simulations, we currently face several challenges that impair data reusability: (1) to facilitate the understanding and use of computational samples (or atomic structures), well-described and harmonized metadata and data are crucial. However, most existing approaches focus on perfect crystal structures only, i.e. neglecting defects. (2) Calculations often involve a combination of different software tools and various file formats. This results in heterogeneous metadata which leads to a lack of semantic interoperability. (3) The workflow provenance that was used to set up a digital sample is frequently lacking.
To address these problems and facilitate data re-use in the field, we have developed the Computational Materials Sample Ontology (CMSO), an application-level ontology for material science computational samples. CMSO initially focuses on describing structures at the atomistic level [1]. The use of the CMSO ontology is complemented by the development of domain-level ontologies describing crystallographic defects [2] and atomistic simulation concepts.
Importantly, to aid domain scientists in implementing ontologies in their everyday research, we developed software tools for the automated annotation and identification of structural features. AtomRDF [3] provides a way for users to automatically annotate their data with ontologies and create application-level knowledge graphs. This improves the querying and findability of their research data. In addition, atomID [4] showcases the use of ontologies in identification processes frequently performed in materials simulations.
The here shown combination of controlled vocabularies and software tools for generating linked open data ensures interoperability between different file formats and software, while also offering the potential for data to be findable and reusable [5].

References
[1] https://purls.helmholtz-metadaten.de/cmso/
[2] https://github.com/OCDO/
[3] https://github.com/pyscal/atomRDF
[4] https://github.com/Materials-Data-Science-and-Informatics/atomID
[5] Wilkinson, M., Dumontier, M., Aalbersberg, I. et al., Sci Data, 2016, 3, 160018.

Please specify ”other”

In addition, please add 3 to 5 keywords.

domain ontology, materials science, semantic interoperability

Please specify ”other”

For whom will your contribution be of most interest?

Researchers

Please assign yourself (presenting author) to one of the following groups.

Scientists and technicians who maintain and operate research infrastructure for data generation

Primary author: AZOCAR GUZMAN, Abril (IAS-9, FZJ)

Co-authors: MENON, Sarath (Max-Planck-Institut für Eisenforschung GmbH); HOFMANN, Volker; Dr HICKEL, Tilmann (BAM); SANDFELD, Stefan

Presenter: AZOCAR GUZMAN, Abril (IAS-9, FZJ)

Session Classification: Session E1

Track Classification: Connecting research data: 6. Interoperable semantics at domain and application level