TOWARDS ONTOLOGIES FOR DATABASES INTEROPERABILITY

Gian-Marc Rignanese

UCLouvain, Institut de la Matièere Condensée et des Nanosciences (IMCN),
Chemin des Ètoiles 8, Louvain-la-Neuve 1348, Belgium
email: gian-marco.rignanese@uclouvain.be
website: https://perso.uclouvain.be/gian-marco.rignanese

Key Words: *Materials Databases, Application Programming Interface, Ontologies.*

In the last decades, a number of materials databases have become available online (see e.g. Ref. [1] for an extensive, yet inevitably incomplete, list). In many cases, these can be accessed via a graphical web interface which targets a "low-throughput" human usage but that is not very well suited for a systematic "high-throughput" computational approach. In fact, in order to take full advantage of modern data-analytics techniques, it is essential that these databases also become accessible through an application programming interface (API) as it is already the case for some of them [2-5]. It would actually be even more beneficial to have access to information originating from multiple databases as they often cover different material families and properties. Nonetheless, retrieving data from multiple databases is difficult since the available APIs are different from one database to another.

In order to overcome these problems, the OPTIMADE API was developed. It was designed so that it can be implemented without significant changes to the established back-end code, and, furthermore, adopting the API is straightforward for the end user. The OPTIMADE specification version 1.0.0 was released on 1 July 2020 [6]. It is supported by leading databases such as AFLOW, the Materials Cloud, the Materials Project, NOMAD, OQMD, ... Currently, the returned properties comprise both mandatory information about the structure (such as the elements the lattice vectors), as well as optional and database-specific information prefixed with the database name (e.g., _aflow_).

In this talk, I will outline some key features of the API specification. I will illustrate its usage through some examples. Finally, I will discuss how it would benefit from more fundamental work on an ontology for materials databases.

REFERENCES

[1] J. Hill *et al.*, MRS Bulletin **41**, 399-409 (2016).
[2] R.H. Taylor *et al.*, Comput. Mater. Sci. **93**, 178-192 (2014).
[3] S.P. Ong *et al.*, Comput. Mater. Sci. **97**, 209–215 (2015).
[4] F. Rose *et al.*, Comput. Mater. Sci. **137**, 362–370 (2017).
[5] L. Talirz *et al.*, Sci. Data **7,** 299 (2020).
[6] C.W. Andersen *et al.*, zenodo, https://doi.org/10.5281/zenodo.4195051