Making scientific data accessible: the NOMAD Laboratory

C. Draxl[1]

[1] Physics Department and IRIS Adlershof, Humboldt-Universität zu Berlin
[2] Fritz Haber Institute of the Max Planck Society, Berlin

Knowledge and understanding of materials is based on their characterization by a variety of measured and/or computed properties. This includes structural features, mechanical and thermal behavior, electronic and magnetic properties, the response to light, and more. For the computational side, the NOMAD Laboratory (Novel Materials Discovery) Centre of Excellence (https://nomad-coe.eu) has created a data infrastructure to collect, share, and keep data available for at least ten years. The variety of data uploaded to the NOMAD Repository (https://nomad-repository.eu) allows for in-depth explorations and discoveries.

The NOMAD Encyclopedia (https://encyclopedia.nomad-coe.eu) is a graphical user interface that allows for making millions of calculations accessible and tangible. The NOMAD Analytics Toolkit (https://analytics-toolkit.nomad-coe.eu) provides a collection of examples and tools to demonstrate how materials data can be turned into knowledge. This platform hosts "interactive codes" in terms of data-science notebooks that present the application of data-analytics models. They concern topics like crystal-structure prediction, property prediction, error estimates, classification of materials, and more.

Prerequisite for all this is a FAIR [1,2,3] data infrastructure, where normalized data in unified, formats and reliable metadata [4] are key.
I will demonstrate NOMAD's capabilities, and discuss our current efforts to extend the data archive by experimental data. An issue, in this context, is understanding errors in both experiment and theory; related to this, assigning error bars and trust levels to high-throughput screening results, as obtained by density-functional theory. For the latter, I will discuss our recent efforts to learn fully converged results of complex materials from errors of calculations of simple materials [5].

References

[1] M. D. Wilkinson et al., Sci. Data 3, 160018 (2016).

[2] C. Draxl and M. Scheffler, MRS Bulletin 43, 676 (2018).

[3] C. Draxl and M. Scheffler, J. Phys. Mater. (2019); DOI: https://iopscience.iop.org/article/10.1088/2515-7639/ab13bb

[4] L. M. Ghiringhelli, C. Carbogno, S. Levchenko, F. Mohamed, G. Hus, M. Lüder, M. Oliveira and M. Scheffler, npj Comput. Mater. 3, 46 (2017).

[5] C. Carbogno, et al., in preparation.