

Artificial Intelligence for data-driven materials science: achievements and perspectives

Luca M. Ghiringhelli

*Fritz-Haber-Institut der Max-Planck-Gesellschaft,
Faradayweg 4, D-14 195 Berlin-Dahlem, Germany*

The number of possible materials is practically infinite, while only few hundred thousands of (inorganic) materials are known to exist and for few of them even basic properties are systematically known. In order to speed up the identification and design of new and novel optimal materials for a desired property or process, strategies for quick and well-guided exploration of the materials space are highly needed. A desirable strategy would be to start from a large body of experimental or theoretical data, and by means of artificial-intelligence (AI) methods, to identify yet unseen patterns or structures in the data, and consequentially predictive (data-driven) models. This leads to the identification of maps (or charts) of materials where different regions correspond to materials with different properties. The main challenge on building such maps is to find the appropriate descriptive parameters (called descriptors) that define these regions of interest.

Here, I present novel methods for the AI-aided identification of descriptors and materials maps, tailored to work (also) with "small-data", and applied to the metal/insulator classification for binary materials, the prediction of novel topological insulators, of CO₂ activation on metal-oxide surfaces, and more. I focus on the (verified) predictive power of the learned maps, which goes beyond the mere interpolation of more "traditional" AI approaches, and analyze current and future challenges.

I also describe the infrastructure to perform such analyses online, within the framework of the Novel-Materials-Discovery (NOMAD) Laboratory.