Finding the needle in the haystack: Materials discovery through high-throughput ab initio computing and data mining

Geoffroy Hautier

Université catholique de Louvain

Essential materials properties can now be assessed through ab initio methods. When coupled with the exponential rise in computational power, this predictive power provides an opportunity for large-scale computational searches for new materials. We can now screen thousands of materials by their computed properties even before the experiments. This computational paradigm allows experimentalists to focus on the most promising candidates, and enable researchers to efficiently and rapidly explores new chemical spaces.

In this talk, I will present the challenges as well as opportunities in materials discovery in high-throughput ab initio computing using examples from transparent conducting materials. I will especially highlight computational predictions which have been followed by experimental synthesis and characterization. In addition to allowing the ability to navigate through a large volume of materials data to identify promising compounds, high-throughput computing also offers unprecedented data mining opportunities to detect new relationships between chemistry, structures, and properties. I will illustrate examples of these relationships through our recent work in crystal structure descriptors and automatic local environment identification, which merge traditional solid-state chemistry and materials science concepts through modern informatics.

The impact of high-throughput computing is multiplied when the generated data is shared with free and easy access. I will finish my talk by presenting the Materials Project (http://www.materialproject.org), a collaborative project which precisely targets such a data dissemination.