Combining the power of high-throughput ab initio calculations and machine learning towards materials informatics

G.-M. Rignanese
Université Catholique de Louvain, Louvain-In-Neuve, Belgium

The progress in first-principles simulation codes and supercomputing capabilities have given birth to the so-called high-throughput (HT) ab initio approach, thus allowing for the identification of many new compounds for a variety of applications (e.g., lithium battery and photovoltaic). As a result, a number of databases have also become available online, providing access to various properties of materials, mainly ground-state though. Indeed, for more complex properties (e.g., linear or higher-order responses), the HT approach is still out of reach because of the required CPU time. To overcome this limitation, machine learning approaches have recently attracted much attention in the framework of materials design.

In this talk, I will review recent progress in the emerging field of materials informatics. I will present the concept of high-throughput ab initio calculations and illustrate it with an example on transparent conducting materials. I will then explain how the developments of this approach naturally lead to the use of machine learning. In particular, I will introduce the MODNet framework and its recent developments for predicting materials properties. Finally, I will show how high-throughput ab initio calculations and machine learning can be combined, giving a few recent examples.