The effective use of data in materials research

Tim Mueller

Johns Hopkins University

Rapid advances in information technology have made it possible to generate, analyze, and distribute large data sets of material properties. One of the great challenges in materials research is to effectively make use of these data sets to accelerate the design and development of new materials. To this end, I will discuss how machine learning techniques can be used to develop model Hamiltonians capable of accurately evaluating the energies of millions of structures per minute for multi-nanometer nanoparticles. I will also demonstrate how informatics can be used to rapidly generate highly efficient k- point grids, addressing a longstanding problem in computational materials research. Our group has developed a publicly accessible k-point grid server backed by a database of hundreds of thousands of k- point grids, and we estimate that for a broad range of computational methods the use of this server should nearly double the speed of electronic structure calculations on crystalline materials.