Practical Data Sharing for Molecular Simulation

Jacob Boes

June 9, 2016

All molecular simulation occurs at a computer. Consequently, all of the procedures and results are at some point recorded on the computer in a machine readable format. Thoroughly documenting this information promises a multitude of benefits, not only to the community, but also to the author:

1. Eliminating reproducibility issues
2. Reducing time spent solving the same problems
3. Recording of own work organized by manuscript
4. Allowing for data to be used in other applications and accessed through data mining

Despite the potential advantages, such high-level data sharing remains uncommon. This is potentially due to the perceived difficulty of generating the required documentation. Here we illustrate practical methods, already implemented in the Kitchin group[1], for streamlining the incorporation of high-level data sharing into the manuscript writing process. In this way, we hope to obtain the advantages listed above with little additional, or potentially even less effort than current methods.

1. Kitchin, John R. "Data Sharing in Surface Science." ScienceDirect, 16 May 2015.