Efficient Workflow for Machine Learning-Guided Discovery of New Liquid Organic Hydrogen Carriers

Hassan Harb, Sarah N. Elliott, Ian T. Foster, Logan Ward, Stephen J. Klippenstein, Larry A. Curtiss, Rajeev Surendran Assary

Artificial intelligence methods can be used to screen, with unprecedented efficiency, very large chemical spaces for materials with various potential applications. In this study, we present a multi-stage, multi-fidelity screening workflow designed to identify new liquid organic hydrogen carriers (LOHCs). The workflow first screens chemical databases containing, in aggregate, billions of molecules to identify structures that are similar to, based on Tanimoto similarity index scores, a set of known LOHCs. Each identified molecule is then given a score based on multi-objective criteria developed from physiochemical and structural properties of known LOHCs. Finally, the highest-scoring molecules are selected for accurate quantum mechanical calculations. Using this workflow, we identified new LOHCs from databases containing billions of molecules.