Zack Ulissi
Carnegie Mellon University

"Enabling Catalyst Data Science with Active Optimization for Compositionally-Diverse Materials."

Abstract: Increasing computational sophistication and resources can enable a larger and more integrated role of theory in the discovery and understanding of new catalysts. Most catalyst studies start in a data-poor regime where the material of interest is unrelated to previous to studies (new structure, composition etc) or the computational methods are incompatible with previous studies (different exchange-correlation functionals, methods, etc). Efficient methods to quickly define, schedule, and organize DFT calculations are thus important and enable the application of optimal design of experiments approaches. In this presentation, I will discuss on-going work and software development to enable data science methods in catalysis. I will describe applications of our approach to ordered bimetallic alloy catalysts, with applications to several electrochemical catalyst discovery efforts including CO2 reduction, oxygen reduction, and water splitting chemistry. I will also discuss the methods and approaches we use to share data among group members and educate new students with the necessary skills to pursue these research directions (including statistics, machine learning, computer science, etc). Finally, I will discuss the transition from data-poor to data-rich regimes and our experiences when data-intensive deep-learning methods become more appropriate than simpler models based on chemical intuition.