A.J. Medford

Bio: A. J. Medford is an Associate Professor in the School of Chemical & Biomolecular Engineering at the Georgia Institute of Technology. His research focuses on the intersection of data science, electronic structure theory, and surface science, with contributions including machine-learned exchange–correlation functionals for density functional theory, the application of neural networks to transient kinetic data, mechanistic insight into nitrogen conversion reactions, and high-throughput materials discovery for catalysis and separations. He received his B.S. in Textile Engineering from North Carolina State University, was a Fulbright Fellow at the Technical University of Denmark, and earned his Ph.D. in Chemical Engineering from Stanford University. He is a recipient of the 2023 ACS "Early Career in Catalysis" award and developed the course "Data Analytics for Chemical Engineers," now a core course in Georgia Tech's "Data Science for the Chemical Industry" certificate program, for which he has earned several teaching awards including the Georgia Tech "Junior Faculty Teaching Excellence" and "Excellence in Online Teaching" awards.

Title: Large Scale Datasets and Machine Learning for Direct Air Capture: The Open DAC Project and Beyond

Abstract

Direct air capture (DAC) with porous adsorbents has the potential to aid large-scale decarbonization, but identifying useful sorbents for capturing CO_2 from humid air remains a formidable challenge given the vast chemical space of candidate materials such as metal–organic frameworks (MOFs). This talk surveys how AI and machine learning, powered by large, high-fidelity computational datasets, are reshaping the discovery pipeline for DAC sorbents using the Open DAC (ODAC) project as a central example. The earlier Open DAC 2023 (ODAC23) dataset established the approach with roughly 38 million density functional theory (DFT) calculations of CO_2 and H_2O adsorption across more than 8,000 MOFs. The new Open DAC 2025 (ODAC25) dataset comprises nearly 60 million DFT single-point calculations for CO_2, H_2O, N_2, and O_2 adsorption in more than 15,000 sorbent structures, introducing chemical and configurational diversity through functionalized MOFs, high-energy GCMC-derived placements, and synthetically generated frameworks. ODAC25 substantially improves both the accuracy of DFT calculations and the treatment of flexible MOFs relative to ODAC23. Alongside the datasets, we release state-of-the-art machine learning interatomic potentials, evaluate them on adsorption energy and Henry's law coefficient predictions, and show how these models can be integrated into classical molecular simulation workflows to move toward economical and scalable DAC processes. The talk will cover these developments and briefly outline challenges and opportunities in the path from atomistic predictions to process-relevant sorbent screening.