



The Department of Chemical Engineering Presents:

## Johannes Hachmann<sup>1,2,3</sup>

(1) Department of Chemical and Biological Engineering,  
University at Buffalo, The State University of New York, Buffalo, New York, United States.  
(2) Computational and Data-Enabled Science and Engineering Graduate Program,  
University at Buffalo, The State University of New York, Buffalo, New York, United States.  
(3) New York State Center of Excellence in Materials Informatics, Buffalo, New York, United States.

# Machine Learning the Structure-Property Relationships that Define Chemistry

### Abstract:

In this presentation, we will demonstrate how modern data science can be used to learn and understand the structure-property relationships at the heart of chemistry. We will discuss how such an understanding as well as data-derived prediction models that encapsulate the relationships can be used to accelerate the discovery process in chemistry and enable the rational design of complex molecular systems with tailored property profiles. We will also show how machine learning can advance molecular modeling and simulation techniques that are traditionally used for the prediction of chemical behavior. The results of these physics-based modeling approaches can be calibrated, augmented, or even replaced by data-derived models at a fraction of the computational cost. Finally, we will introduce our software ecosystem for data-driven *in silico* research that we employ in studies such as the development of new high-refractive index polymers. It consists of four loosely connected program suites: *ChemLG* allows us to enumerate compound space and create molecular candidate libraries; *ChemHTPS* provides an automated platform for the virtual high-throughput screening of these compound libraries; *ChemBDDDB* offers a database and data model template for the massive information volumes created by data-intensive projects; and *ChemML* is a machine learning and informatics toolbox for the validation, analysis, mining, and modeling of such data sets.

### Biosketch:

Johannes Hachmann is an Assistant Professor of Chemical Engineering at the University at Buffalo (UB), a Core Faculty Member of the UB Computational and Data-Enabled Science and Engineering graduate program, and a Faculty Member of the New York State Center of Excellence in Materials Informatics. He earned a Dipl.-Chem. degree (2004) after undergraduate studies at the universities of Jena and Cambridge, M.Sc. (2007) and Ph.D. (2010) degrees in Chemistry from Cornell University, and he conducted postdoctoral research at Harvard University before joining the UB faculty in 2014. The research of the Hachmann Group fuses (*first-principles*) molecular and materials modeling with virtual high-throughput screening and modern data science (i.e., the use of database technology, machine learning, and informatics) to advance a data-driven discovery and rational design paradigm in the chemical and materials disciplines. One of the centerpieces of the group's efforts is the creation of an open, general-purpose software ecosystem for the data-driven design of chemical systems and the exploration of chemical space. This work was recognized with a 2018 NSF CAREER Award.

**Wednesday, January 23, 2018**

**3:25 PM**

**Goergen Hall Room 101**