**Autonomous Acquisition of Composition-Temperature Phase Diagrams via Active Factorization of X-ray Diffraction Data**

**B. DeCost**, National Institute of Standards and Technology, USA

Unsupervised Machine learning techniques have become instrumental in the reduction and analysis of data from many combinatorial experiments. This has enabled important work in high throughput X-ray diffraction mapping of composition spread thin films, where material synthesis is massively parallelized and the preliminary analysis has become quasi-real-time thanks to advances in cluster analysis techniques. However, exhaustive high throughput characterization is sometimes inefficient in that only a small subset of measurements provide unique information regarding the location of phase boundaries. We apply active learning algorithms to adaptively select diffraction measurements to identify the location of phase boundaries while avoiding these effectively redundant measurements. This expands the scope of materials systems accessible in a single measurement session, and has the added benefit of lowering the barrier of inspecting and interpreting the raw data by actively filtering for the most relevant data. After each measurement, we update our estimated phase map by re-clustering the diffraction data, and use a Bayesian non-parametric model to extrapolate the cluster assignments to the full design space. Subsequent measurements are chosen on the basis of the confidence of the Bayesian non-parametric model of the phase map, so that the automated system continually seeks to decrease its global uncertainty over the estimated phase diagram. We have applied this system to map ternary and pseudo-ternary phase diagrams at room temperature, and to map composition-temperature phase diagrams for binary metal-insulator transition materials. Our system is nearly an order of magnitude more efficient than dense measurements for these systems, even when restricting the sampling order for the composition-temperature diagrams to monotonic increases in temperature. Active sampling promises to enable efficient exploration of chemical systems with more than three components, with high (or adaptive) resolution in composition and temperature. Finally, I will discuss preliminary work on a semi-mechanistic modeling approach that aims to increase the robustness of the system to alloying effects such as diffraction peak broadening and shifting.