Title: Statistical Thermodynamics as Paradigm for a Data Science

In this talk, we introduce the notions of information entropy (Kullbeck-Leibler divergence) and statistical internal energy function as purely statistical concepts in connection to a large, recurrent, set of data under a probabilistic model. We show a statistical thermodynamic structure emerges, a la P. W. Anderson, in the infinite limit of data {\it ad infinitum}. Through Legendre transform and its duality, we discuss a novel insight on the Second Law that is independent of time, as forcefully suggested by Lieb and Yngvason. We suggest a return to Boltzmann-Ehrenfest's ergodic hypothesis, not about mathematical models but on empirical big data.

[1] Lu, Z. and Qian, H. (2022) Phys. Rev. Lett. 128, 150603.
[2] Qian, H. (2022) Chem. Theory Comput. 18, 6421.