Contribution ID: **71**   Type: **not specified**

Foundation Models for Accelerated Discovery

Wednesday, 11 May 2022 14:00 (50 minutes)

AI is having a transformational impact for accelerating discovery. Massive volumes of scientific data, which are continuously growing due to tireless efforts from many scientific communities of discovery, are enabling data-driven AI methods to be developed at ever increasing scale and applied in novel ways to breakthrough bottlenecks in the scientific method, and to speed up the discovery process. Examples include using AI models to assist in knowledge extraction and reasoning over large repositories of scientific publications, creating AI surrogate models to predict the output of simulations and speed up complex simulation campaigns, training AI generative models to create novel hypotheses – such as to enable de-novo design of molecules by leveraging data about known chemicals and their properties, and developing AI models that can predict chemical reactions and automate synthesis and experimentation. At the same time, foundation models have emerged as a powerful new development in AI that will make further impact on accelerating scientific discovery. Foundation models learn "universal representations"from massive-scale data, typically using unsupervised or self-supervised training methods, with the goal to enable and simplify a diversity of downstream tasks. Prominent examples of foundation models are the large-language models trained from massive corpora of text that have been driving the state-of-the-art for natural language processing. In this talk, we review how foundation models work and discuss how they can learn effective representations for scientific discovery. We show examples how foundation models apply for challenges such as materials discovery and drug development. We discuss the potential for foundation models to have a key role for a broader set of scientific challenges and drive further impact of AI for accelerating discovery.

Presenter: SMITH, John R. (IBM Fellow and IBM Research global lead for Discovery Technology Foundations)

Session Classification: Industry Session