Data-Driven Materials Science

Jacqueline M. Cole

Ray Dolby Centre, Cavendish Laboratory, Department of Physics, University of Cambridge, J. J. Thomson Avenue, Cambridge, CB3 0US. United Kingdom.

This lecture presents the development and application of AI software tools, machine-learning (ML) and language models that are enabling data-driven materials chemistry. The talk describes how to source high-quality data to extract and deploy experimental data to (a) train ML models; (b) realise data-driven materials discovery; (c) create energy-efficient domain-specific language models for materials chemistry to feed agentic AI and help democratise AI for the materials research community; (d) support the advances of self-driving laboratories via AI-automated materials characterisation.