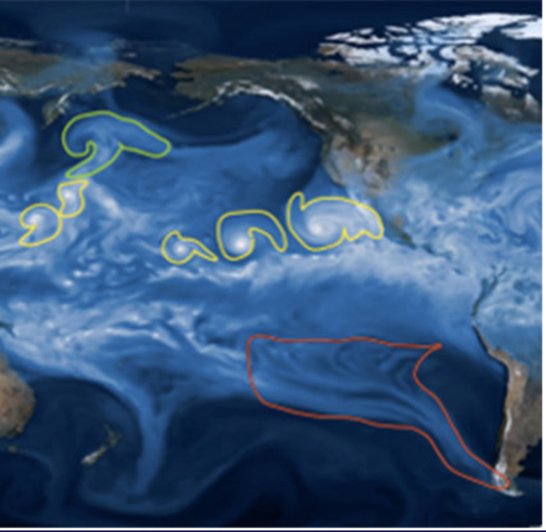
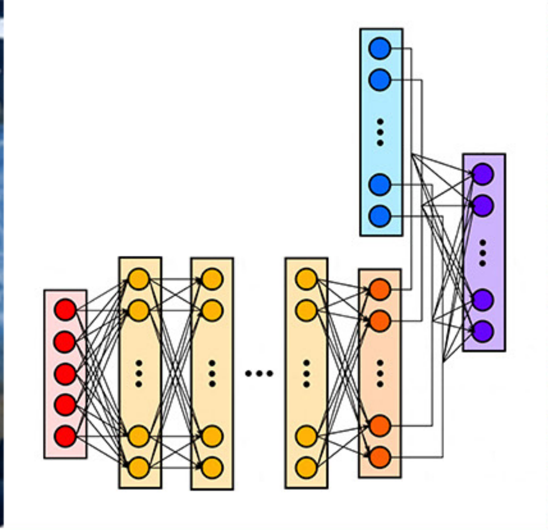
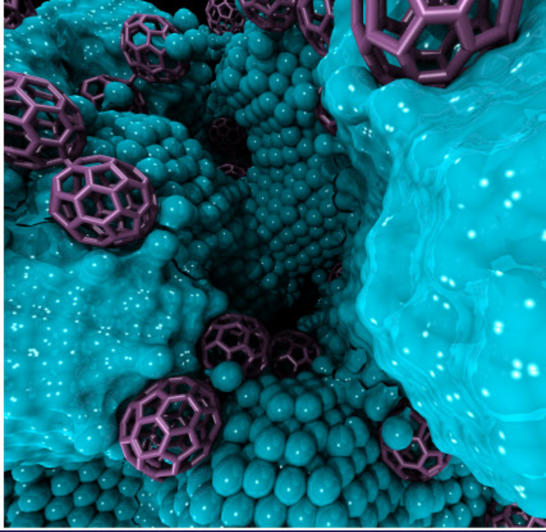
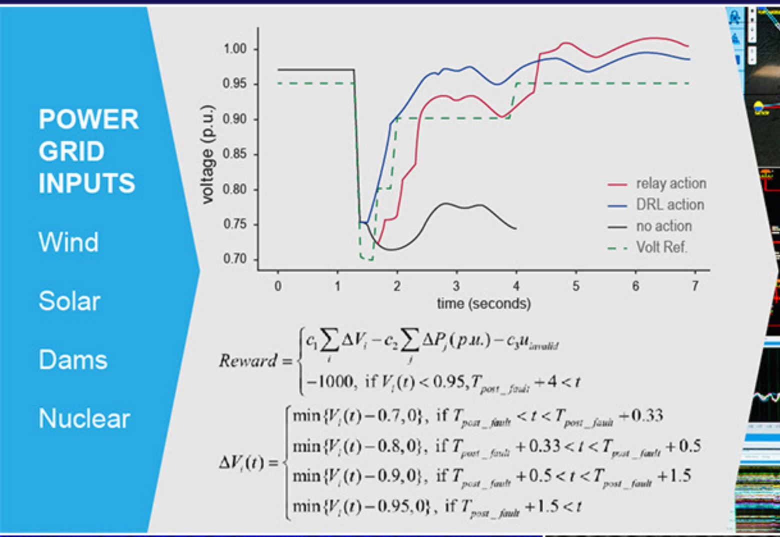
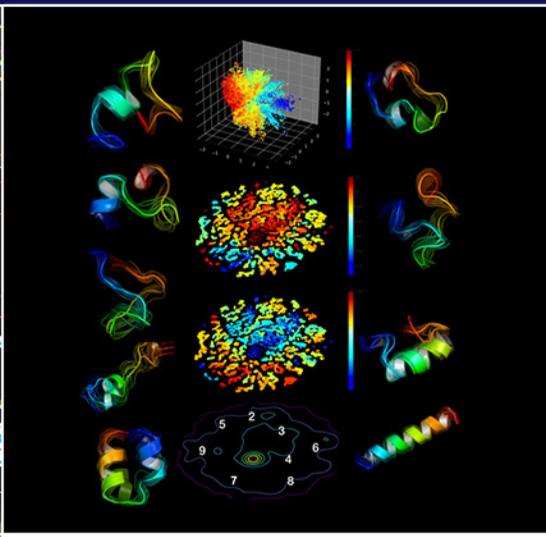
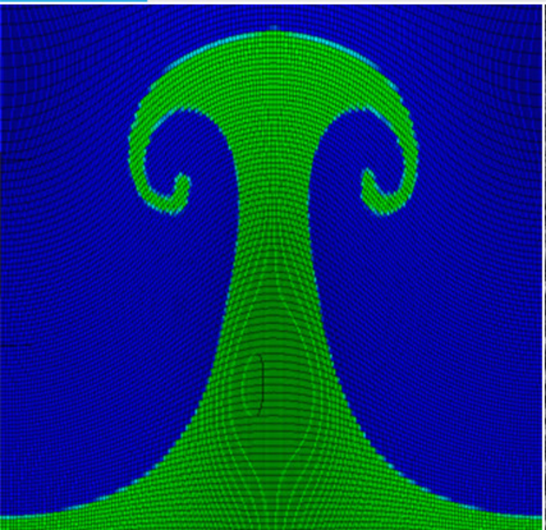
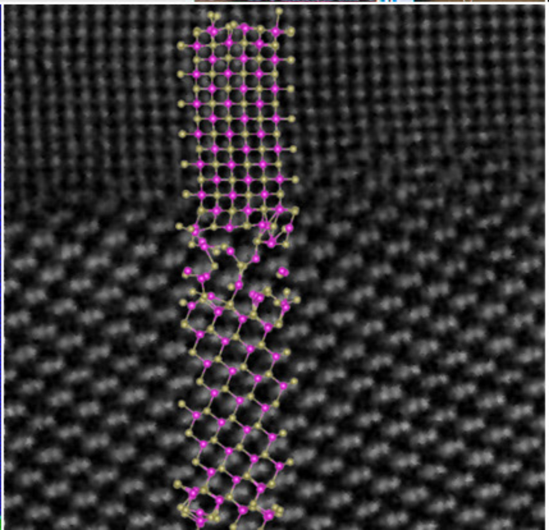




BASIC RESEARCH NEEDS FOR Scientific Machine Learning

Core Technologies for Artificial Intelligence



Prepared for U.S. Department of Energy Advanced Scientific Computing Research

U.S. DEPARTMENT OF ENERGY

Scientific Machine Learning & Artificial Intelligence

Developing the mathematical and scientific computing foundations for accelerated research insights

Scientific Machine Learning is a core part of Artificial Intelligence and a computational technology that can be trained, with scientific data, to augment or automate human skills. Across the Department of Energy, Scientific Machine Learning has the potential to transform science and energy research by harnessing DOE investments in massive data from scientific user facilities, software for predictive models and algorithms, high-performance computing platforms, and the national workforce. The versatile and crosscutting nature of machine learning provides strong motivation for formulating a prioritized research agenda to maximize its capabilities and scientific benefits for the DOE. At the January 2018 Basic Research Needs workshop, major machine learning opportunities and grand challenges were viewed through the lens of applied mathematics and scientific computing research. The workshop report identifies six Priority Research Directions for Scientific Machine Learning. The full report is available via https://doi.org/10.2172/1478744.

Scientific Machine Learning & Artificial Intelligence

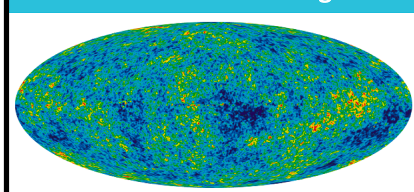
Scientific progress will be driven by

- **Massive data: sensors, simulations, networks**
- **Predictive models and adaptive algorithms**
- **Heterogeneous high-performance computing**

Trend: Human-AI collaborations will transform the way science is done.

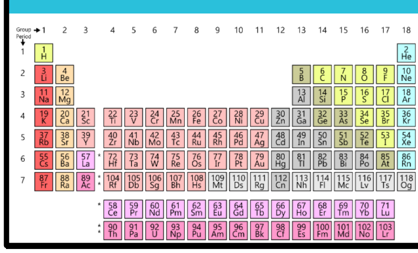
EXEMPLARS OF SCIENTIFIC ACHIEVEMENT

Cosmic Microwave Background

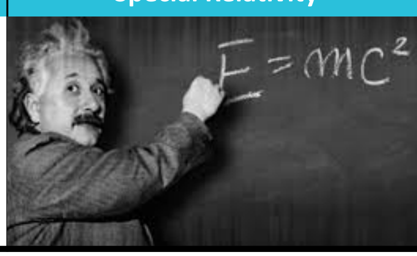


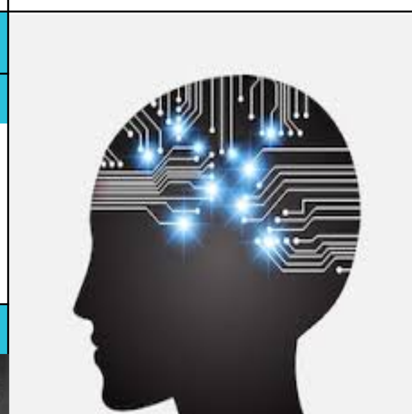
DNA Structure

Periodic Table of the Elements



Special Relativity





Human-AI insights enabled via scientific method, experimentation, & AI reinforcement learning.

U.S. DEPARTMENT OF ENERGY | Office of Science

DOE Applied Mathematics Research Program
Scientific Machine Learning Workshop (January 2018)

Priority Research Directions

1. **Domain-aware scientific machine learning for leveraging scientific domain knowledge**

   Key question: *How can domain knowledge be effectively incorporated into scientific machine learning methods?*

   Incorporating domain knowledge may dramatically reduce data requirements, accelerate training and prediction, and improve the accuracy and defensibility of scientific machine learning models. Progress will require new mathematical methods that can account for physical principles, symmetries, uncertainties, and constraints.

2. **Interpretable scientific machine learning for explainable and understandable results**

   Key question: *What is the right balance between the use of increasingly complex machine learning models versus the need for users to understand the results and derive new insights?*

   The increased integration of domain knowledge into scientific machine learning methods may improve the interpretability of these methods. Advances will require developing new exploration and visualization approaches to interpret and debug complex models using domain knowledge.

3. **Robust scientific machine learning for stable, well-posed, and efficient methods**

   Key question: *How can efficient scientific machine learning methods be developed and implemented to ensure the results are not unduly sensitive to perturbations in training data and model selection?*

   Scientific machine learning models are subject to uncertainty in their general form, internal structure, and associated parameters. To be considered reliable, new approaches must reach the same level of rigor expected of mainstream scientific computing algorithms. Research will focus on scientific machine learning methods and implementations that are stable, well-posed, and efficient.

4. **Data-intensive scientific machine learning for automated scientific inference and data analysis**

   Key question: *What novel approaches can be developed for reliably finding signals, patterns, or structure within high-dimensional, noisy, or uncertain input data?*

   Scientific machine learning has the potential to reveal valuable information hidden in massive amounts of scientific data from experiments, observations, simulations, and other sources.

5. **Machine learning-enhanced modeling and simulation for predictive scientific computing**

   Key question: *What are the barriers and potential advantages to using scientific machine learning in developing predictive computational models and adaptive algorithms?*

   Scientific machine learning has the potential to improve the fidelity of reduced-order or sub-grid physics models, automate computational steering, and optimize parameter tuning within multiscale scientific simulations.

6. **Intelligent automation and decision-support for the management and control of complex systems**

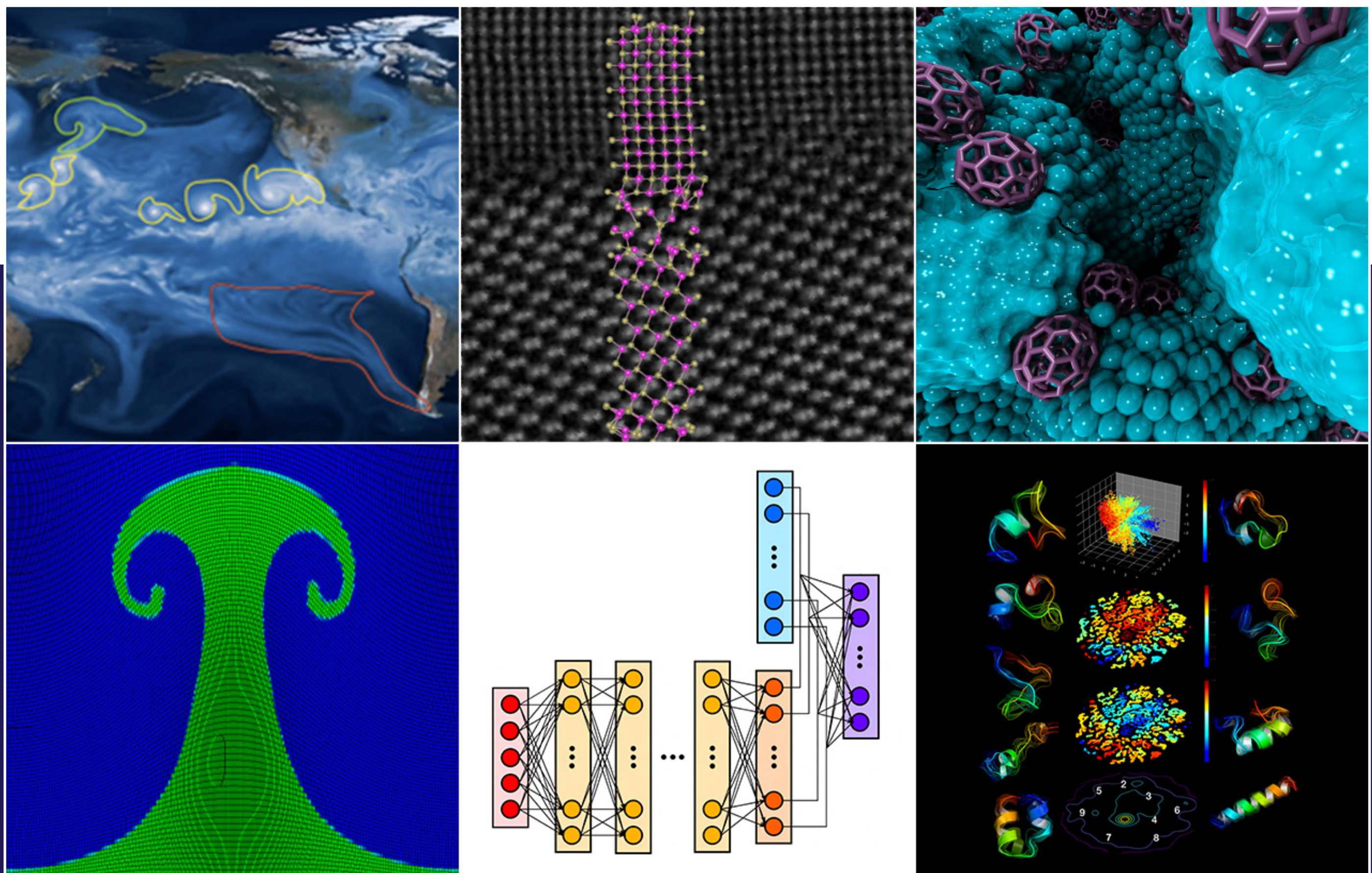
   Key question: *What are the challenges in developing scientific machine learning for decision-support and automation of complex systems and processes?*

   Scientific machine learning has widespread use in improving the operational capabilities of scientific user facilities, communication networks, power grids, or other sensor-equipped infrastructures and complex processes.

Summary

Scientific Machine Learning & Artificial Intelligence will have broad use and transformative effects across the Department of Energy. Accordingly, the Basic Research Needs workshop report has identified six Priority Research Directions (PRDs). The first three PRDs describe **foundational research themes** that are common to the development of all Scientific Machine Learning methods and correspond to the need for domain-awareness (PRD #1), interpretability (PRD #2), and robustness (PRD #3). The other three PRDs describe **capability research themes** and correspond to the three major use cases of Scientific Machine Learning for massive scientific data analysis (PRD #4), machine learning-enhanced modeling and simulation (PRD #5), and intelligent automation and decision-support of complex systems (PRD #6).

The PRDs provide a sound basis for a coherent, long-term research and development strategy in Scientific Machine Learning & Artificial Intelligence. Over the last decade, DOE investments in applied mathematics have laid the groundwork for the type of basic research that will underpin key advances in the six PRDs. Such advances will build on the work from leading researchers in optimization, linear algebra, high-performance solvers and algorithms, multiscale modeling and simulation, complex systems research, uncertainty quantification, and the new basic research areas that will emerge from the pursuit of transformative technologies.



U.S. DEPARTMENT OF ENERGY | Office of Science