DOE BER Earth and Environmental Systems Science Division (EESSD) Call for White Papers to Advance an Integrative Artificial Intelligence Framework for Earth System Predictability: AI4ESP

White Paper Purpose, Structure, and Submittal

Purpose

Submitted white papers will be used to inform the design of three sequential workshops (conducted in 2021-2022) focused on answering the following overarching question of: How can DOE directly leverage artificial intelligence (AI) to engineer a substantial (paradigm-changing) improvement in Earth System Predictability?

Structure of white papers

White papers should be prepared using the following outline and may be up to a maximum of 3 pages long (12-point font, not including the optional Suggested Partners and References sections). Use of the template provided is optional.

1. **Title**
2. **Authors/Affiliations:** List in order of largest contribution
3. **Focal Area(s):** One or two sentences only
4. **Science Challenge:** Short statement describing the area addressed by the white paper
5. **Rationale**: Description of the research needs/gaps, the barriers to progress, and the justification for and benefits associated with the proposed approach
6. **Narrative**: Scientific and technical description of the opportunities and approach; activities that will advance the science; and specific field, laboratory, model, synthesis, and/or analysis examples
7. **Suggested Partners/Experts (Optional)**: Laboratory and university partners or experts in the field who may be able to present a related webinar or plenary presentation at a workshop
8. **References (Optional)**

Authors are limited to two submissions as lead author but may participate as a co-author in other submissions. Teaming is encouraged to reduce the reviewing workload. Multi-institutional responses are welcome; however, a clear lead who can speak authoritatively on the white paper contents should be identified. [Note: Protected information should not be included in white papers, but instead should be shared directly with the appropriate U.S. Department of Energy (DOE) program manager(s).]

Submittal

White papers must be submitted as PDF files by **5:00 p.m. EST on February 15, 2021**, using this Google Form. After the submission date, white papers will be posted publicly on a website (active Feb 2021) and be made available in advance of the workshops for review by participant and general public engagement.

Questions: Prospective authors are welcome to seek clarification on any part of this announcement through the AI4ESP slack workspace #white-paper-call-questions slack channel.

Background

Throughout its history, the U.S. Department of Energy (DOE) has tackled some of the world's most difficult scientific and technical challenges. In response to changing scientific needs, DOE's national laboratories manage some of the most sophisticated facilities and observatories on the planet; develop multi-scale, multi-physics Earth system models (ESMs); and apply artificial intelligence and machine learning, advanced visualization, and cutting-edge computing assets in innovative ways to solve scientific challenges. Major scientific breakthroughs are seeded when individual and distinct investments can be integrated in novel ways, for example, where facilities, models, experiments, and artificial intelligence (AI) are partnered.

Nearly a decade ago, DOE recognized that an acceleration was needed in the transition of basic science into new predictive capabilities to meet the needs of scientists and stakeholders. To help meet this challenge, DOE’s Earth and Environmental System Sciences Division (EESSD) incorporated a novel model-experiment (MODEX) approach, linking interdependent observation and model development into its management philosophy and strategic planning (Figure 1).

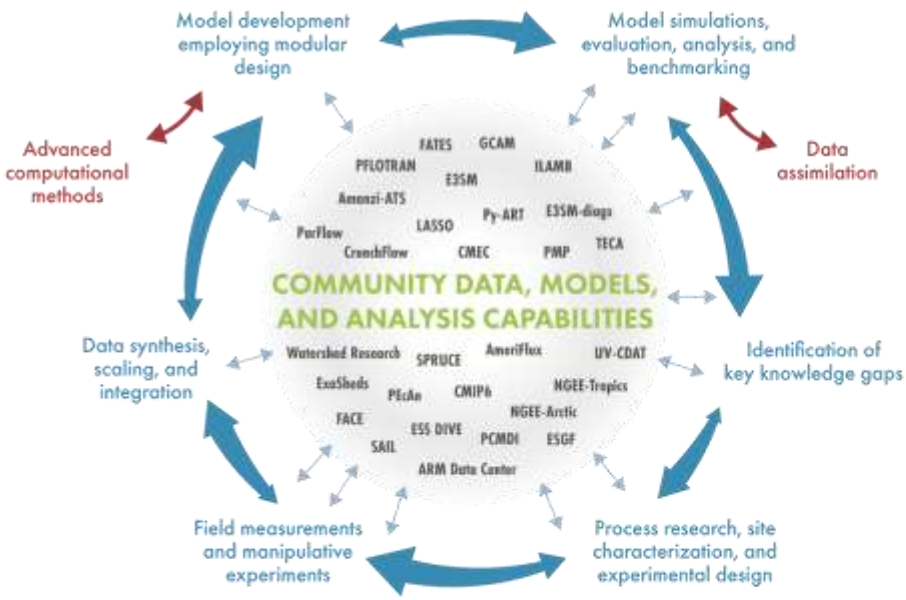

Figure 1: Schematic of the MODEX approach to scientific discovery (outer ring) and various DOE data, models and analysis capabilities that should be linked as community resources based on open science principles (inner sphere).

Despite advances in high-resolution modeling and better observational capabilities, scientific and decision-making communities have increasingly sought predictive capabilities that exceed current knowledge. Some examples include an urgent need for more accurate prediction of extreme events, more complete characterization of uncertainty in models and data necessary to constrain scientific findings, and effective bridges between observational designs and useful predictions.

To meet these pressing needs, a paradigm shift is necessary to build the integrative research framework of the future. Building such a framework will require attention to the design of integrative research approaches that take advantage of recent scientific and technological advances – such as artificial intelligence and exascale computing, etc. – that are not widely incorporated in EESSD research. A successful shift has the potential to result in an entirely new framework that will integrate new observational strategies with capabilities in automated data quality validation; edge computing; new nonlinear and multiscale data assimilation methodologies; model parameter estimation and feature detection using AI; and hybrid prediction models that combine physics with AI.

The hierarchies of models that must be considered in such a paradigm shift include the hybridization of one-dimensional and multi-dimensional ESMs, large eddy simulations, and agent-based models; AI-generated surrogates for such models; and scale-aware, AI-based analytics to complement traditional physical approaches to evaluating uncertainties. With advanced computing and AI as leadership disciplines within DOE, we envision new and heretofore unforeseen possibilities for EESSD to use these new frameworks to discover next generation science, revolutionizing our Nation's predictive capabilities and advancing its scientific agenda.

This exercise to explore a paradigm shift in prediction science follows the July 2020 release of the AI for Science technical report, prepared by a consortium of DOE national laboratories to identify scientific opportunities for AI in the upcoming decade. Two weeks later, the White House issued a joint OSTP-OMB President's S&T memorandum, that highlighted Earth system predictability as a national science priority and identified AI and edge computing as areas that the federal agencies should continue to develop. More recently, DOE collaborated with other agencies to identify science capabilities that the research community could rapidly and aggressively apply to advance prediction science in Earth science research.

The DOE vision is to radically improve predictive capabilities by applying AI methods to build a new integrative system that spans the continuum from observations to predictive modeling. This effort will require the exploration of AI across the MODEX enterprise (Figure 1) to determine the most impactful applications along the observation-modeling continuum.

Call for White Papers

All interested researchers are asked to read the current version of the Earth and Environmental Systems Sciences Division (EESSD) Strategic Plan. This plan enumerates five grand challenges that frame the Division's investments through 2023: integrated water cycle, biogeochemistry, high-latitude science, drivers and responses, and data-model integration. Since the development of the plan, the following cross-cutting areas of interest have emerged: predictability of extreme events, terrestrial aquatic interfaces, regions of high gradients (e.g., coastal zones and watersheds), and integration of AI and other new technologies into scientific research.

The purpose of this announcement is to solicit white papers from the scientific community that focus on development and application of AI methods in areas relevant to EESSD research with an emphasis on quantifying and improving Earth system predictability, particularly related to the integrative water cycle and associated water cycle extremes. White papers must describe novel and innovative approaches to improving the predictability of the Earth system and explain why these approaches are expected to succeed. We expect that a novel framework, derived from white paper concepts and a series of workshops, will improve capabilities for knowledge capture and distillation that provide future computational constructs across the EESSD research enterprise. Authors should consider novel approaches, needed resources, technologies available and/or in the pipeline, and unforeseen developments that have the potential to transform the Earth system research enterprise out to FY2030 and beyond. We anticipate that exascale computing, edge computing, 5G/6G, and use of quantum computing and quantum sensors will be further developed during this period.

White papers should adhere to the following criteria and constraints:

- White papers should identify a transformational science question that involves the water cycle and associated water cycle extremes as a centerpiece in conjunction with the EESSD's cross-cutting areas described above. This science question may include, for example, extreme precipitation, extreme drought, strong perturbations on surface water or groundwater systems, and/or extreme flooding or inundation, and the impacts of these events on biogeochemistry, terrestrial aquatic interfaces and high latitude/gradient regions.
- Responses should be framed around one or more of the following focal area, ensuring that technology and techniques are incorporated as a critical component of a development pathway:
    1. Data acquisition and assimilation enabled by machine learning, AI, and advanced methods including experimental/network design/optimization, unsupervised learning (including deep learning), and hardware-related efforts involving AI (e.g., edge computing)
    2. Predictive modeling through the use of AI techniques and AI-derived model components; the use of AI and other tools to design a prediction system comprising of a hierarchy of models (e.g., AI driven model/component/parameterization selection)
    3. Insight gleaned from complex data (both observed and simulated) using AI, big data analytics, and other advanced methods, including explainable AI and physics- or knowledge-guided AI
- White papers should identify the primary relevant focal area (1,2, or 3) - the workshops will be arranged around these focal areas. Responses should include how unique DOE capabilities would be brought to bear on the scientific question, including e.g., exascale computing, existing data holdings, community modeling programs and unique observational capabilities. White papers must provide sufficient detail to show how the conceptual idea can more rapidly and fully address the scientific question.
- White papers must address the data-model integration grand challenge presented in the EESSD Strategic Plan. Ideas that incorporate data generated by the Atmospheric Radiation Measurement (ARM) Facility, Environmental Molecular Science Laboratory (EMSL), Next Generation Ecosystem Experiments (NGEEs), Science Focus Area (SFA) observatories, and/or other DOE-generated information are encouraged. Use of data provided by other agencies is also encouraged, as long as it is supplemented by and enhances DOE-supported datasets and software.
- Develop a high-level approach that incorporates one or more of the following as a critical component of a development pathway (if unknown, state any known barriers to understanding the pathway):
    - Topics must employ machine learning or another AI techniques or technologies
    - Topics are encouraged to employ other emerging technologies, e.g., edge computing, nonlinear data assimilation, 5G/6G and advanced wireless
    - Topics should consider a level of difficulty that demands advanced computing capabilities that are likely to be available over the next decade
    - Topics that are more forward looking and explore the application of quantum sensors and/or quantum computing related to AI are also encouraged
- Responses should outline how code and other tools generated in the process of addressing the author-selected science question will be made reusable and findings reproduceable. Such tools can include (but are not limited to) open source code, packaging to allow deployment on a variety of hardware, outreach and workforce development. White papers should outline how FAIR (Findable, Accessible, Interoperable, Reusable) principles will be incorporated into products (e.g., data, networks, models, tools, etc.).