A Vision for AI in ESS Science: A Report on the AI4ESS Roundtable

Charuleka Varadharajan (LBNL), James Stegen (PNNL), Pamela Weisenhorn (ANL), and Forrest M. Hoffman (ORNL)

Introduction

The U.S. Department of Energy's Environmental System Science (ESS) program advances fundamental understanding of the coupled physical, chemical, and biological processes that govern land, water, and subsurface systems. ESS research integrates models with experimental data collection in an iterative, hypothesis-driven **Model–Experiment (ModEx)** cycle: data are used to generate model predictions, which guide new observations and experiments that deepen process understanding and improve predictive capacity. **This approach positions ESS to connect process-based science with artificial intelligence (AI), including machine learning (ML) to accelerate discovery while preserving mechanistic fidelity.**

There are significant opportunities to leverage AI in ESS to advance the DOE Secretary's priorities for unleashing energy innovation, strengthening reliability of critical infrastructure, and ensuring environmental resilience that depend on accurate prediction of ecosystem processes. Reliable energy use and expansion requires forecasts across multi-year to decadal timescales, research that ESS is uniquely positioned to provide. While system heterogeneity and complexity challenge conventional methods, purely data-driven AI models can fail to generalize beyond training domains. **Integrating process understanding from ESS research into AI models can enable predictions that are robust, generalizable, and trustworthy.**

Motivated by these opportunities, ESS convened the **AI4ESS Roundtable: Opportunities and Advances in the Application of AI in Environmental System Science in August 2025**. The roundtable brought together researchers from across DOE national laboratories to assess the state and potential of AI in ESS. Participants considered how advances in AI can drive new scientific discoveries, accelerate the ModEx cycle, and enable deeper integration of observations, experiments, and models. Participants identified near-term (≤1 year) and mid-term (1–3 years) opportunities to apply AI to ESS science, aligning with DOE's mission. Key themes included:

- Developing AI agents for (semi-)autonomous science, AI-ready data, and harmonized data-model workflows to improve ESS productivity and accelerate ModEx cycles.
- Using AI-guided approaches, advanced instrumentation, and edge-HPC to scale and optimize new measurements.
- Identifying how ESS process understanding and digital–physical testbeds can improve state-of-art AI models, and enable robust, trustworthy predictions.

The tactical and strategic opportunities identified at the roundtable provide ESS opportunities to accelerate DOE's mission and demonstrate leadership in the responsible use of AI. With rapid advances in AI, building momentum across ESS is evident in a range of nascent efforts, beyond those highlighted in the near-term exemplars outlined in the Research Frontiers section.

Cross-cutting AI Grand Challenge: Scaling and Transferability

A central challenge in Earth and environmental science is predicting how water, energy, and elements move and interact across a **wide range of spatial (**molecular to kilometers) **and temporal scales** (seconds to decades). **Transferability** — applying knowledge learned in one system to others with sparse or no observations — was identified as another critical challenge for ESS. Scaling and transferring insights from a limited number of study sites to other regions or the entire nation is difficult because of extreme heterogeneity and complex feedbacks in the Earth system (Figure 1), but is a grand challenge that ESS is well poised to address.

One promising approach to these challenges is to develop **foundation models** for environmental systems: large, flexible base models trained on diverse, widely available datasets that can be fine-tuned for specific applications having sparse data without starting from scratch. Both commercial and academic models (e.g., NASA/IBM Prithvi, Google AlphaEarthFoundation, NVIDIA/Fourcastnet) have shown that such models can perform multiple scientific tasks. For ESS, such models would integrate multimodal data — from in situ field observations and laboratory measurements to remote sensing products and process model outputs — to capture system structure, function, and variability, and learn shared patterns in underlying processes (e.g., fluid flows, heat transfer). Applications include spatiotemporal predictions of water, energy, and biogeochemical fluxes at large scales beyond the capabilities of current models, downscaling coarse observations to decision-relevant resolutions, identifying drivers of response variables, and revealing new process understanding or emergent patterns from the data.

Participants emphasized that such models must be accurate and deliver the **"right answers for the right reasons,"** with interpretable predictions grounded in basic scientific principles. Embedding mechanistic constraints and process-based representations from ESS research are key to ensuring the models generalize across locations, timescales, and environmental regimes. Benchmark datasets for training and testing at ESS field sites through ModEx cycles will be essential to evaluate model generalization and failures. AI-enabled observations can be used to help generate the vast quantities of data needed to train and validate such models. AI-guided autonomous instruments such as drones and robotics, with advanced wireless and edge computing, can shorten ModEx cycles and increase observations by an order of magnitude.

A foundation model would enable scaling and transfer of processes across the Earth system's extreme variability. Participants acknowledged the value of light-touch, cross-project coordination through emerging shared datasets, benchmarks, and interoperable workflows as a natural pathway toward the development of an ESS foundation model. Such a model would be a powerful tool for multiple scientific applications, such as those described below.

ESS Research Frontiers for AI Application and Development

Roundtable participants identified five research frontiers with strong potential for rapid progress using AI (Figure 1). Each presents challenges such as heterogeneity, sparse data, and scaling, while offering opportunities to advance both the science and the tools used to study it. Examples of near-term (≤1 year) and longer-term opportunities (1–3+ years) were identified. This would require integrating existing ESS data into AI-ready datasets, benchmarking commercial models versus data-driven, mechanistic and hybrid models with data from AI-enabled ModEx field campaigns.

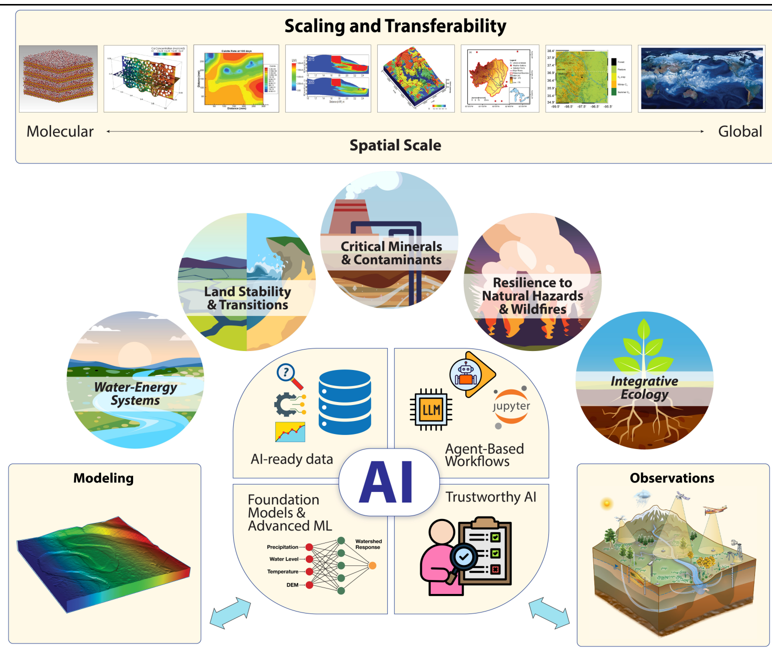

Figure 1. ESS hallmarks of process-based science and ModEx can be combined with AI to address research frontiers and the grand challenge of scaling and transferability. This would rely on core capabilities founded on trustworthy AI supported by AI-ready datasets from existing or new observations, agents, and advanced ML methods, including foundation models.

Water-Energy Systems

Reliable and cost-effective energy systems depend on a steady supply of high-quality water for hydropower, cooling for power plants, industrial facilities, and AI data centers. Water management is also energy intensive, including pumping, treatment, distribution, and storage. The tight coupling between water and energy makes predicting water availability and quality fundamental to DOE's mission. Achieving accurate predictions is challenging because water systems are shaped by both **natural drivers** (e.g., precipitation, snowmelt, vegetation, and soil processes) and **human interventions** (e.g., groundwater pumping, reservoir management, and land use change). For example, snowpack dynamics strongly affect streamflow and water supply for the Western U.S. but remain highly uncertain due to natural variability and land use changes.

AI can offer new capabilities to address these challenges, particularly when combined with ESS's process-based knowledge. State-of-the-art deep learning and hybrid models offer new capabilities such as predicting streamflow, groundwater, and water quality dynamics at regional to continental scales, while capturing effects of human activities. However, the models have been built only for variables and locations with abundant measurements and cannot model coupled processes, limiting their accuracy and utility. Integrating AI with mechanistic representations can address this and enable hypothesis testing — for example, how snowpack and vegetation dynamics jointly control streamflow — thereby advancing both fundamental understanding and applied prediction.

Near-term exemplar: Demonstration using AI to integrate existing data and simulations from ESS testbeds with ModEx to scale streamflow, groundwater, and water quality predictions to other regions. This would support energy-water planning and infrastructure resilience for applications such as hydropower operations, drought or flood mitigation, and agriculture.

Broader opportunities include, but are not limited to:

- **Groundwater mapping and managed aquifer recharge** strategies that optimize timing and location of recharge to stabilize water supplies.
- **Water quality predictions** that link integrated watershed processes (e.g., vegetation shifts, biogeochemistry) with streamflow and reservoir temperature.
- **Transferring insights across basins** using foundation models, transfer learning, site-similarity analysis, guiding efficient data collection, and adaptive monitoring.

Together, these advances can create scalable water-energy prediction systems informed by mechanistic knowledge, and directly relevant to DOE's energy reliability goals.

Resilience to Natural Hazards and Wildfires

Natural hazards such as wildfires, floods, droughts, and severe storms pose persistent risks to U.S. energy systems. These events can damage infrastructure and disrupt transmission and water resources essential for energy production. Hazards can also occur together as **compound events** (e.g., drought and wildfire) — producing far greater impacts than individual events. Predicting such hazards and their impacts on ecosystems and infrastructure is a critical challenge for DOE.

AI, particularly in combination with ESS process knowledge, can transform hazard prediction and response. By integrating satellite imagery, aircraft and drone collected data, ground-based sensors, weather models, and process-based simulations, AI can detect hazards in near real time, track their evolution, and forecast cascading impacts on energy systems. Hybrid AI models are especially promising for capturing interactions among vegetation, human activities, weather, and hydrology that drive hazard dynamics, with predictions grounded in physical processes.

Near-term exemplar: Demonstration of wildfire prediction and response that fuses geospatial data to deliver early warnings and guide adaptive ModEx field campaigns. This would support post-fire ecosystem recovery assessments and predict how burned landscapes alter water quality, sediment transport, and ecosystem services critical for energy infrastructure.

Broader opportunities include, but are not limited to:

- **Compound hazard prediction**, using AI to capture nonlinear interactions between drought, wildfire, storms, and land management practices.
- **Feature tracking and data assimilation**, where high-throughput predictions dynamically direct observational assets (e.g., drones) to maximize information gain during hazard events.
- **Transfer learning** to extend hazard models trained in data-rich regions (e.g., California) to other areas where DOE has energy infrastructure (e.g., Columbia River Basin).

By embedding these insights into AI workflows, ESS can deliver models that not only predict hazards but also explain them — building trust and enabling effective decision-making.

Critical Minerals and Contaminants

Reliable U.S. energy systems require secure supplies of **critical minerals** such as lithium, cobalt, and rare earth elements, which are essential for advanced batteries and energy storage, high-efficiency power electronics, and nuclear energy systems. Current U.S. supplies are limited and heavily reliant on imports, creating vulnerabilities in the energy supply chain. **Legacy contaminants** from DOE's historical operations that continue to affect soil and water quality at multiple sites, posing risks to human and ecosystem health, is another national priority.

ESS is uniquely positioned to advance solutions because its researchers study the coupled physical, chemical, and biological processes that control mineral mobility and contaminant transport across scales. However, these processes are highly complex: **microbial activity**, **redox reactions**, **fluid flow**, and **geochemical transformations** interact from pore to watershed scales, and are affected by environmental conditions. This complexity makes conventional models difficult to calibrate and transfer. AI can improve models to capture this complexity, particularly in linking small-scale biological processes with large-scale hydrology and geochemistry.

Near-term exemplar: Create benchmark datasets around DOE sites, integrating subsurface observations, laboratory experiments, monitoring records and other geospatial data (e.g., geologic maps) to train hybrid AI–reactive transport models that simulate mineral speciation and transport.

Broader opportunities include, but are not limited to:

- **Developing and validating AI-enabled proxies for key subsurface properties and processes** by fusing surface observations, geophysical signatures, and process simulations to augment sparse observations and infer drivers (e.g., permeability, redox state, mineralogy).
- **Designing AI-enabled extraction strategies guided by coupled biotic–abiotic processes** that integrate microbial, mineralogical, geochemical, and hydrologic datasets.
- **Accelerating contaminant remediation design** by using AI to detect patterns in contaminant plume dynamics, optimize monitoring networks, and evaluate treatment scenarios.

By combining AI with ESS's deep expertise in **subsurface biogeochemistry**, DOE can develop predictive tools that both support secure mineral supply chains and accelerate remediation of contaminated environments, directly advancing energy security and environmental stewardship.

Land Stability and Transitions

The Arctic and U.S. coastal zones are critical to the nation's energy system. The Arctic provides oil, gas, and infrastructure corridors, much of which is built on permafrost. When permafrost thaws, it destabilizes the ground, causing deformation that can damage pipelines, roads, and facilities. At the same time, U.S. coastlines and deltas are zones of rapid change, where erosion, sediment redistribution, storms, and sea level rise threaten communities, ports, energy facilities, and access routes. Predicting these transitions is difficult because hydrology, vegetation, soils, and geomorphology interact in complex, highly variable ways. AI models that combine mechanistic understanding of **cryosphere, hydrology, and coastal processes**, can reveal how these interactions dynamically evolve through space and time to impact land stability.

Near-term exemplar: Development of a permafrost ground-deformation predictor that combines snow and thermal physics simulations with geospatial AI. This tool would allow direct testing of transferability across Arctic study sites.

Broader opportunities include, but are not limited to:

- **Mapping permafrost, ground ice, and ice wedges** using remote sensing, geophysical surveys, and AI methods for spatial extrapolation.
- **Assessing thaw impacts** on vegetation, hydrology, and infrastructure stability through hybrid AI–physics models and remote sensing observations.
- **Modeling coastal change** at high resolution, coupling atmospheric, hydrologic, and geomorphic processes to anticipate erosion and storm impacts on energy infrastructure.

ESS researchers also highlighted the importance of capturing **abrupt transitions and extremes through ModEx campaigns**, such as thaw slumps, storm surges, or compound events like permafrost thaw during anomalously wet years. Such cases are critical stress tests for both AI and process models and will strengthen prediction of infrastructure risk to support DOE's mission to ensure **resilient, secure, and sustainable energy infrastructure** in rapidly changing landscapes.

Integrative Ecology for Reliable Energy Systems

Ecological systems, including soils, plants, and microbial communities, regulate the fluxes of water, carbon, and nutrients and shape land-surface states for energy planning and operations. Long-term forecasts relevant to energy infrastructure require understanding of coupled above- and belowground processes and their responses to environmental variability and management. Ecological systems adapt to hydrology and biogeochemistry; mediate storage, transformation, and transport; and determine ecosystem resilience with nonlinear dynamics. Biotic–biotic and biotic–abiotic interactions span molecular to landscape scales, with feedbacks, thresholds, hysteresis, and time lags that confound simple scaling and transferability. Heterogeneity in soils, lithology, vegetation, and microclimate adds complexity, particularly with sparse observations.

ESS has an opportunity to integrate AI with ModEx campaigns to derive process insights by coupling metagenomic datasets with environmental drivers (e.g., soil moisture, temperature, vegetation indices). This integrated ecological perspective supports the other research frontiers through predictive understanding of how biological processes modulate hydrology, geochemistry, and landscape dynamics, complementing rather than superseding their distinct objectives.

Near-term exemplar: Demonstrate an AI-enabled hybrid biogeochemical model, leveraging BER large-sample datasets (e.g., WHONDRS, MONet, Ameriflux, GROW) and other public datasets (e.g., USGS) to infer microbial drivers of biogeochemical fluxes across field sites.

Broader opportunities include, but are not limited to:

- **Identifying thresholds and tipping points**, such as microbial or soil shifts that trigger large-scale ecosystem state transitions.

- **Linking above- and belowground processes**, integrating plant trait data, soil microbiomes, and hydrology to forecast productivity and ecosystem resilience.
- **Scaling dynamics and fluxes,** using AI to link microbial metabolic models to ecosystem models validated using measurements at ESS field sites

Advancing AI-enabled **integrative ecology** will result in trustworthy ecological prediction at the spatiotemporal resolutions needed for decadal forecasting and robust energy system planning.

Trustworthy and Explainable AI

> **Key Takeaway:**
> Developing new approaches that enable extrapolation beyond training regimes with physical consistency is an active area of scientific AI research. ESS science, data, and models can both challenge and advance the state-of-the-art in this field.

The roundtable emphasized that trustworthy and explainable AI must be central to ESS applications. Trustworthiness was defined broadly to include reliability, accuracy, robustness under uncertainty and non-stationarity, transparency of assumptions, and provenance tracking. Explainability implies understanding why an AI model produces its outputs, ensuring results enhance process understanding rather than obscure it.

Participants cautioned against overreliance on black-box approaches, highlighting the need for rigorous validation, benchmarking, and uncertainty quantification. Trustworthy AI requires systematic comparisons between AI and process-based models, as well as evaluation across multiple variables and model construction choices. Roundtable discussions identified explainable AI methods and causal analysis as promising but difficult to apply at scale. Expanding these methods by embedding mechanistic constraints was highlighted as a critical direction for advancing both ESS science and AI itself.

Beyond technical considerations, participants emphasized cultural and institutional dimensions of trust. Trust in models often translates into trust in the modelers, underscoring the importance of transparent communication, clear documentation of assumptions, and visible QA/QC processes. A community-defined **formal risk and ethics framework** with standardized accuracy and uncertainty metrics, and evaluation frameworks for hybrid and AI models can be broadly useful for establishing trust in AI use. Establishing benchmark datasets, systems to identify trusted data and models, and adopting data trust classifications will make model reliability and limitations transparent to users. By addressing both the technical and cultural aspects of trust, ESS can ensure that AI serves as a tool for discovery rather than a barrier to scientific understanding.

Using AI to Accelerate ModEx and ESS Productivity

The **Model–Experiment (ModEx)** framework is a hallmark of ESS research and provides a natural pathway for advancing trustworthy AI. Embedding AI into this iterative process allows the framework to scale more efficiently, adapt in real time, and accelerate discovery. Roundtable discussions highlighted multiple opportunities for AI-enabled ModEx. **Standardized, AI-ready ESS datasets** can be used to train and validate AI models, while **agentic workflows** can integrate model outputs with field observations across facilities, ranging from edge devices to high-performance computing systems. **Digital twins and physical testbeds** can serve as proving grounds where models are stress-tested under diverse conditions, enabling rapid evaluation of transferability and robustness. **Agent-enabled systems** with reinforcement learning can autonomously identify high-value measurements and orchestrate data collection campaigns. **Near-term Exemplar: Using AI agents to enhance ESS productivity**, such as generating model input files, and mining datasets or papers to extract and synthesize relevant information.

Broader opportunities include, but are not limited to:

- **Adaptive, AI-guided sampling**, where AI identifies the most informative observations to reduce uncertainty in predictions, agents and edge computing trigger targeted observations (e.g., drones, in situ sensors) at critical times.
- **Benchmarking testbeds**, to compare predictions across AI-based and process-based models constrained by observations to identify strengths, weaknesses, and failure modes.
- **Cross-scale integration**, using AI (e.g., graph neural networks, transfer learning, generative AI) to bridge observations from pore to watershed scales and beyond within ModEx loops.

To support these capabilities, ESS will require robust cyberinfrastructure (Figure 2) that combines data repositories, workflow orchestration, agentic AI, and model–data integration tools. Such infrastructure should be designed for **transparency, reproducibility, and explainability**, ensuring that AI supports — rather than replaces — scientific reasoning. Workforce development in AI across modelers, field personnel, and data curators was also identified as important, and is being advanced through the ESS AI/ML Cyberinfrastructure Working Group. By combining robust governance with technical innovation, ESS can transition from fragmented, project-specific datasets to a culture of **shared, AI-ready data resources**. This will enable foundation models, hybrid workflows, and ModEx applications to accelerate scientific discovery and DOE mission outcomes.

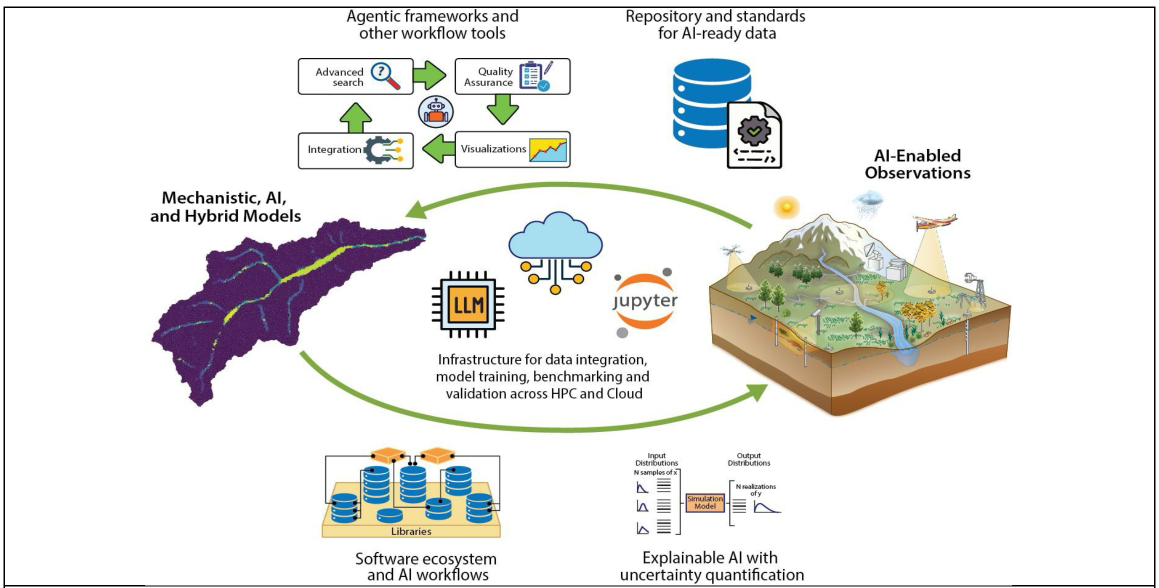

Figure 2. ESS ModEx-based science is ideally suited to leveraging AI to advance understanding and predictive capacity. Unique ESS data generation capabilities and standardized datasets in ESS-DIVE and other BER data systems can be used with AI to inform models via agentic workflows. Resulting models can optimize data generation efforts via explainable AI workflows with uncertainty quantification.

Data for AI: New Data Generation and Benchmark Datasets

High-quality, AI-ready data are essential for advancing AI in ESS, but their availability varies across domains. While remote sensing provides abundant surface observations, subsurface data are often sparse, heterogeneous, and site-specific. Traditional ESS data collection focuses on hypothesis-driven experiments and process model calibration, which do not always align with AI needs for large-sample data. It is essential to generate large, structured datasets for AI workflows.

Roundtable participants emphasized the importance of **proxies** for difficult-to-measure variables. For example, soil moisture can serve as a proxy for microbial and geochemical processes, and vegetation patterns can provide information about subsurface hydrology. Identifying, validating, and standardizing such proxies across sites will enable AI to bridge observational gaps. Participants also noted the value of **event-based indices** (e.g., snowpack thresholds, storm frequency metrics) that allow data to be subset and queried for targeted applications.

Building AI-ready data requires more than access: it requires **standardization, documentation, and machine readability**. ESS-DIVE provides a strong foundation by enabling archival of FAIR data and specifying community formats for AI-readiness, but additional steps are required.

Near-term Exemplar: Creating benchmark datasets by standardization and synthesis of existing ESS datasets (e.g., integration of observations or simulation outputs) that could be used to train and validate AI models and identify critical data gaps.

Broader opportunities include, but are not limited to:

- **Automation of metadata generation and ontology mapping** using AI agents to standardize data and metadata formats to lower the burden on data producers.
- **Integration of datasets** across scales, instruments, and modalities to provide coherent training and validation products.
- **Catalyzing cultural change** with incentives (e.g., recognition, performance credit, badging) and governance that encourage broad publication of high-quality AI-ready data, with mechanisms for provenance tracking and citations for data used in AI models.

AI for Data

AI not only "consumes" data, it also can transform how data are **generated, curated, and used** in ESS. Barriers to AI adoption arise from the time and labor required to restructure, harmonize, and quality-control data for model training and evaluation. Using AI for data workflows can reduce these barriers, accelerate ModEx, and expand the scientific value of ESS datasets.

Near-term Exemplar: Workflow automation, where AI agents can accelerate model setups, standardize data into repository templates, harmonize units, and enrich metadata. These capabilities reduce manual effort and improve reproducibility.

Broader opportunities include, but are not limited to:

- **Adaptive sampling and sensing:** Reinforcement learning and Bayesian optimization can guide optimal measurements such as triggering drone or robotic deployments during events, or siting sensors to maximize information gain. Coupling observing system simulation experiments with AI can quantify the value of new measurements to optimize observations.

- **Synthetic data generation:** Generative AI methods (e.g., variational autoencoders, diffusion models) can produce physically plausible datasets where observations are sparse. Such products must include explicit uncertainty quantification and validation.
- **Data preparation:** Using anomaly detection methods to identify issues in real-time, generative AI to quality control and gap-fill partial time series or spatial coverage. These tools enable AI to make use of heterogeneous data without discarding valuable information.

Participants also emphasized that data prepared using AI must be paired with **guardrails for trustworthiness**. This includes explicit documentation of assumptions, logging of AI-agent actions, and standardized benchmarks for evaluating agent performance. Importantly, AI must complement and not replace human scientific judgment in data collection and interpretation.

By embedding AI into the full data lifecycle from acquisition to curation to synthesis, ESS can increase research efficiency and the representativeness of its datasets, reduce uncertainty in models, and accelerate discovery. These capabilities will also enable (semi-)autonomous ModEx workflows, where AI agents improve the cycle of prediction, observation, and understanding.

AI Methods: Model Choice and Development

ESS has long advanced process-based models to test hypotheses and make predictions grounded in well-understood mechanisms. The adoption of AI offers an expanded toolkit for addressing questions that are data-rich, computationally demanding, or poorly captured by existing parameterizations. Participants emphasized that the choice of AI methods must be guided by scientific objectives, available data, and the need for physical consistency.

A diverse portfolio of AI approaches is relevant to and being used in ESS science, including:

- **Classical machine learning (ML):** Methods such as nonlinear regression, random forests, and XGBoost remain valuable for feature identification, clustering, and predictions.
- **Geospatial AI:** Techniques including computer vision and convolutional neural networks (CNNs) are well-suited to multi-modal remote sensing data, enabling high-resolution image recognition, classification, and spatiotemporal prediction.
- **Deep learning:** Recurrent and attention-based models (e.g., LSTMs, transformers) capture complex dependencies in time series, images, and multi-sensor data streams for predictions.
- **Reinforcement learning:** Provides strategies for optimization and control, with applications to adaptive sampling, autonomous field and laboratory systems.
- **Generative models:** Variational autoencoders, diffusion models, and transformer-based foundation models can synthesize new data, fill gaps, and support hypothesis generation.
- **Hybrid modeling:** Physics-informed neural networks, differentiable programming, and emulators (surrogate models) such as neural operators can embed mechanistic understanding into AI architectures. These approaches aim to conserve or learn physical laws, improve extrapolation to new regimes, and accelerate high-cost simulations.

- **Large language models (LLMs):** offer new capabilities for technical language translation, scientific data search and syntheses, guided data and model analytics, and reasoning.

Roundtable participants highlighted that **foundation models** are likely to play a transformative role in enabling AI for ESS. A key question is whether existing commercial or other pre-trained models can suffice as ESS foundation models with fine-tuning, or if additional model development is required. Participants cautioned that foundation models must be stress-tested for **transferability, explainability, and robustness under extremes**.

To ensure trust and transparency, ESS can leverage and extend **benchmarking frameworks** originally developed for process-based models, such as the International Land Model Benchmarking (ILAMB) package and the Program for Climate Model Diagnosis and Intercomparison (PCMDI) Metrics Package (PMP) for AI. These will provide quantitative, standardized performance metrics, reduce overfitting, and identify failure to model key processes.

By advancing and adapting a spectrum of AI methods — from classical ML to foundation models — and embedding them within ESS's ModEx framework, the community can both expand predictive capacity and reveal new process understanding. This dual benefit underscores ESS's role as a leader in developing AI that is both **scientifically trustworthy and mission relevant**.

Collaborations within BER, DOE, and Beyond

ESS can maximize its impact by collaborating across DOE programs, facilities, and external partners to accelerate AI development and deployment. Participants emphasized that ESS's strength lies in coupling process-based knowledge with advanced data generation, which positions it as both a **driver of AI science needs** and a **partner in DOE-wide AI initiatives**.

ESS collaborations with other BER facilities and programs can create shared AI resources:

- **ESS-DIVE, EMSL, JGI, KBase, and NMDC** are already working on providing unified search capabilities, which can enable future integration into an AI-ready data ecosystem.
- Linking ESS data with **other datasets** from across BER will enable cross-domain foundation models, hybrid workflows, and data synthesis at unprecedented scales.

Partnerships with **ASCR** (Advanced Scientific Computing Research) can leverage DOE's computational ecosystem and avoid redundant investments, for example, by:

- Advancing development of data management tools, applied math libraries, edge computing, and advanced AI methods with use of HPC and Integrated Research Infrastructure (IRI).
- Co-developing **scalable tools for streaming, storage, and compression.** This can align computational infrastructure with ESS's needs for ModEx workflows.

ESS can also align with **DOE-wide AI initiatives** by:

- Providing domain expertise to the **AI Manhattan Project II** and **Transformational AI Models consortium (ModCon)** to enhance development of DOE AI models and workflows.
- Actively engaging in the **American Science Cloud**, will enable integration of ESS data into the DOE AI ecosystem, and lower barriers to train, fine-tune, and benchmark AI models.

External partnerships provide additional opportunities:

- Collaborations with **other agencies** (e.g., NASA, NOAA, and USGS**)** can expand ESS's access to geospatial, monitoring data and support interagency testbeds.
- **Industry partnerships** (e.g., Google, Microsoft, NVIDIA, IBM) can provide complementary capabilities in large-scale foundation models and data, cloud computing, and AI tools.

By leveraging open-source and public-private collaborations, ESS can accelerate domain-specific AI method development while contributing to broader AI innovation, ensuring DOE science priorities shape AI infrastructure that supports discovery and mission-driven applications.

Next Steps

Participants at the AI4ESS Roundtable identified both **tactical near-term steps** and **strategic mid-term investments** to build momentum, while ensuring sustained progress. These actions should demonstrate tangible value to DOE's missions, strengthen trust in AI methods, and lay the groundwork for long-term capabilities such as ESS-specific foundation models.

Possible near-term priorities (<1 year) could include:

- **Launching near-term exemplars** to provide visible demonstrations of ESS-enabled AI.
- **Developing benchmark datasets and evaluation suites** for AI/hybrid models with metrics for comparing model performance and explicit uncertainty reporting.
- **ESS AI-focused Cyberinfrastructure Working Group (CIWG)** to coordinate activities across ESS, including training, seminars, hackathons, and workforce development.
- **Conducting a workshop or roundtable** to develop an ESS foundation model "minimum viable product" (MVP), including targeted use cases aligned with DOE AI initiatives.

Possible mid-term priorities (1–3 years) to consider are:

- **Developing an ESS foundation model** centered around identified use cases and datasets
- **Forming a multidisciplinary AI Center for ESS**, uniting labs and disciplines to advance the foundation-model roadmap and associated cyberinfrastructure.
- **Establishing field testbeds** linked to ModEx campaigns, where AI-enabled digital twins and autonomous systems can be stress-tested in real-world conditions.
- **Demonstrating semi-autonomous data collection**, e.g., leveraging edge computing, in situ sensors, drones, and distributed networks guided by AI agents.
- **Developing a formal risk and ethics framework** for AI in ESS, including metrics, standards for uncertainty quantification, provenance tracking, and responsible use.
- **Integration with DOE-wide AI initiatives** through partnership with ASCR

Together, these activities create a **stepwise path** toward ESS foundation models and trustworthy AI workflows that are both scientifically rigorous and mission relevant. By delivering early exemplars while investing in testbeds, data systems, and governance, ESS can demonstrate immediate impact while building sustainable capacity for the future.

Summary: Advancing ESS Science and DOE's AI Capabilities

Environmental System Science (ESS) brings a unique and powerful capability to DOE: the ability to couple deep process-based understanding of land, water, and subsurface systems with advanced data generation and AI-enabled modeling. By embedding AI within the ModEx framework, ESS ensures that AI accelerates discovery while remaining grounded in physical, chemical, and biological principles. This alignment is critical for producing predictions that are not only accurate, but also trustworthy, transferable, and explainable.

Through the AI4ESS Roundtable, participants identified opportunities for ESS to lead in advancing DOE's mission. Research frontiers in water–energy systems, resilience to natural hazards and wildfires, critical minerals and contaminants, land stability and transitions, and integrative ecology illustrate the breadth of applications where AI can provide scientific breakthroughs and mission-relevant capabilities. The community also highlighted the need to address foundational challenges — scaling across heterogeneity, building trustworthy AI, and preparing AI-ready data — to ensure sustained progress.

To keep pace with AI development, near-term actions such as developing benchmark datasets, launching targeted exemplars, and piloting AI-ready data workflows, will deliver immediate results. Mid-term investments in multidisciplinary AI Centers, testbeds, and semi-autonomous field deployments, paired with governance frameworks and integration with DOE-wide AI initiatives, will establish durable infrastructure for the future.

ESS is poised to deliver transformative capabilities for DOE **by advancing mission-relevant science** using **trustworthy and explainable AI** and development of **AI-ready data and AI-enabled cyberinfrastructure** for ModEx. These efforts will enhance U.S. energy-water security, resilience to natural disasters, critical minerals extraction, and ecosystem management for reliable energy. ESS will hence position DOE at the forefront of applying AI to Earth and environmental sciences, while remaining true to its core mission of fundamental, process-based research.