Oak Ridge National Laboratory's Strategic Research and Development Insights for Digital Twins*

Anees Al-Najjar[1], Valentine Anantharaj[1], Dilip Asthagiri[1], Vittorio Badalassi[1], Prasanna Balabrakash[1], Tom Beck[1], Wesley Brewer[1], Anthony Cabrera[1], Jiahao Cheng[1], Debraj De[1], Patxi Fernandez-Zelaia[1], Antigoni Georgiadou[1], Swarnava Ghosh[1], Chathika Gunaratne[1], Cory Hauck[1], Prashant Jain[1], Anuj Kapadia[1], Mariam Kiran[1], Olivera Kotevska[1], Dan Lu[1], Massimiliano Lupo-Pasini[1], Jason Mayeur[1], Alexander Melin[1], Marshall McDonnell[1], Narasinga Rao Miniskar[1], Michael Muhlheim[1], Van Ngo[1], Jim Nutaro[1], Femi Omitaomu[1], Sarp Oral[1], Pradeep Ramuhalli[1], Nageswara Rao[1], Diego Rosso[2], Arjun Shankar[1], Arpan Sircar[1], Jeremy Smith[1], Gautam Thakur[1], Rama Vasudevan[1], S.V. Venkatakrishnan[1], Kris Villez[1], Feiyi Wang[1], Patrick Widener[1], Ross Wang[1], Haowen Xu[1], Guannan Zhang[1], Neil Zhang[1], and Amirkoushyar Ziabari[1]

[1]Oak Ridge National Laboratory
[2]University of California, Irvine

Introduction

Oak Ridge National Laboratory (ORNL) is pleased to provide our response to the NITRD RFI on Digital Twins Research and Development. Digital twins are virtual representations of physical systems, leveraging real-time data to simulate and predict behaviors. ORNL is advancing digital twin technology across various disciplines, including neutron scattering, networking, science ecosystems, supercomputing, secure facilities, mobility technologies, materials design and discovery, power systems, fusion reactors, biological sciences, and earth observation. These efforts aim to enhance scientific research, operational efficiency, and decision-making processes. ORNL facilities, such as the High Flux Isotope Reactor (HFIR), Grid-C, Spallation Neutron Source (SNS), and Oak Ridge Leadership Computing Facility (OLCF), provide the infrastructure to develop and demonstrate these digital twin technologies. In this document, we lay out key challenges, research gaps, and future opportunities based on our experience with digital twins that aim to serve as useful contributions towards a National Digital Twins R&D Strategic Plan. In the remaining document, we address nine of the thirteen topic areas specified in the RFI.

1 Artificial Intelligence (AI)

Prior R&D at ORNL has utilized AI/ML to advance digital twin development for various applications. This strategy involves integrating simulations with real-time data, enabling continuous learning and optimization. For nuclear and energy systems, digital twins can enhance design, safety, and operational efficiency. In

*This document is approved for public dissemination. The document contains no business-proprietary or confidential information. Document contents may be reused by the government in the National Digital Twins R&D Strategic Plan and associated documents without attribution.

fusion reactors, they provide insights into plasma behavior and material interactions, informed by experimental data. Neutron scattering experiments benefit from autonomous measurement and experimental steering. Inverse materials design uses AI to generate optimized material structures. Mobility technologies are enhanced through realistic traffic scenario generation and AI-driven control systems. Cellular modeling and earth system simulations require multimodal data integration and advanced computational models. ORNL's supercomputing resources enable real-time data processing, supporting these digital twins' continuous learning and application across various domains. However, developing robust digital twins requires overcoming numerous challenges:

- Data Quality and Reliability: Ensuring accurate and reliable sensor data is crucial, as digital twins rely on real-time data, which can be noisy or compromised.
- Legacy Data Management Systems: Outdated systems limit the scalability and efficiency of digital twin deployment.
- Data Privacy: In critical infrastructure applications data sharing is restricted due to proprietary concerns.
- Continuous Updating: Digital twins must continuously be updated with new data, which can be challenging given the systems' complexity.
- Security Concerns: Robust measures are necessary to protect against cyber-attacks on digital twin systems.
- Standard Protocols: Establishing standard protocols for digital twin design, deployment, and maintenance is necessary for safety and efficiency.
- Integration of Multi-Physics Models: Developing comprehensive models that integrate different physical domains (e.g., plasmas, materials) remains a significant gap.
- High-Throughput Virtual Representation: Creating robust and high-throughput virtual representations of complex systems is challenging, requiring further AI model development.
- Scalability of Inversion Algorithms: Solving inverse problems with non-linear physics models at scale is an ongoing challenge.
- Trustworthy AI Models: Ensuring AI models are reliable and transparent is essential, especially in critical applications like nuclear safety and weather prediction.
- Simulation and Data Analytics: High-quality simulations and real-time data analytics are expensive, requiring significant computational resources.
- Model Trustworthiness and Manufacturability: Addressing trustworthiness and feasibility in manufacturing remains crucial for materials design.

ORNL's interdisciplinary approach, leveraging its extensive facilities and expertise, aims to address these challenges, advancing the state-of-the-art in digital twin technologies for diverse applications.

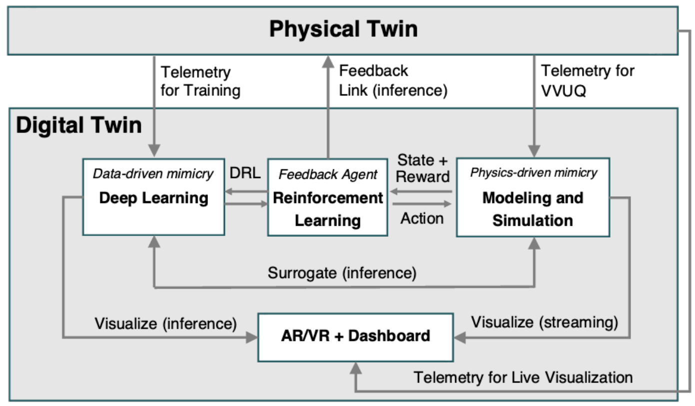

Figure 1: Typical DT component interaction patterns [11]

1.1 Digital Twin Model Integration and Real-time Performance

Topic: AI, Focus area: Integration of Digital Twins with AI

Digital Twins (DT) typically consist of a collection of various models (AI/ML, simulation), control agents, system telemetry data, and AR/VR components, all inter-communicating to create a virtual replica of their physical counterparts. Figure 1 illustrates typical interaction patterns between AI/ML and different components of digital twins. For example, telemetry data may be used to train data-driven AI/ML models to "mimic the structure, context, and behavior of a physical counterpart" [21].

AI/ML models may be deployed in several ways. They can serve as surrogate models to replicate specific aspects of the digital twin. Frameworks such as HPE SmartSim can integrate simulations for either online inference [23] or online training [4]. Reinforcement learning (RL) uses a simulated environment to train an agent to make optimal decisions for the physical twin. Typically, simulations act as the training environment.

Studies have identified at least six different execution motifs or patterns for deploying AI with traditional simulations, along with the middleware that may be used and performance implications [10]. Performance is influenced by various factors and scenarios. For instance, inferencing can be conducted online, offline in batches, or by streaming [8].

Challenges and Gaps: Real-time inference remains a significant challenge in integrating simulations with AI models. Achieving real-time performance may involve several strategies, including reduced precision, scaling across computational resources, or implementing AI surrogates. Numerous studies have investigated scaling up computational resources to achieve real-time performance (e.g., [6, 7, 9]).

Despite recent progress in coupling simulations with AI models, more complex digital twin workflows, encompassing all components shown in Figure 1, still require robust ecosystems for development and benchmarks to assess performance. These ecosystems may integrate existing frameworks such as SmartSim [23] for AI-Sim coupling and benchmarks such as XRBench [18] for assessing AI-VR coupling and SimAI-Bench [3] for assessing AI-Sim coupling performance.

1.2 Enhancing Cognitive and Generative Digital Twin Interaction

Topic: AI, Focus area: Leverage generative AI for digital twin modeling & simulation with the consideration of the potential impact on a digital twins' physical counterpart

Motivation The advancement of generative AI models, including Large Language Models (LLMs) and physics-based generative models, has revolutionized Digital Twins (DTs) in scientific research. These models enhance DTs' ability to interact with their physical counterparts [20] by analyzing vast amounts of unstructured and structured data. This integration allows DTs to provide detailed responses [30], enhance data synthesis [19], improve decision-making [27], facilitate accurate simulations [29], and support intelligent operations [13]. The synergy between generative AI and DTs promotes citizen science and informed

decision-making through Human-AI partnerships, making scientific research more efficient and impactful.
Challenges Despite their successes, generative AI models face significant challenges. LLMs, for instance, suffer from hallucinations, generating plausible-sounding but incorrect information [16]. This issue is pronounced in specialized domains due to training on static datasets that lack depth in these fields [5]. Additionally, physics-based generative models can struggle with the computational complexity and accuracy required for realistic simulations. The inability to access dynamic, up-to-date, and comprehensive specialized knowledge results in inaccuracies and a failure to provide reliable answers in scientific contexts [12]. Establishing trustworthiness and reliability in these models is crucial, particularly in high-stakes applications. Integrating dynamic, domain-specific data and continuous training is essential to improve precision and relevance.
Vision Addressing these limitations involves integrating Retrieval-Augmented Generation (RAG) into DTs and extending the scope to include physics-based generative models. Cognitive Digital Twins (CDTs) incorporate advanced AI for learning, reasoning, and decision-making [22]. Embedding RAG enables the dynamic retrieval of relevant information, providing accurate insights and accelerating research processes. Physics-based generative models can simulate complex systems with high fidelity, enhancing the realism and applicability of DTs. Establishing robust validation and verification frameworks will be critical for advancing the trustworthiness of these models. This might include incorporating systematic and continuous feedback loops from domain experts and deploying advanced evaluation metrics to assess accuracy and relevance.
Examples Autonomous Decision Optimization uses LLMs for optimizing intermodal freight transportation by gathering requirements, automating decision strategies, and conducting data analytics [27]. Autonomous Scientific Discovery utilizes RAG for process conformance in healthcare, dynamically retrieving up-to-date information to improve process models [17]. There is a wide range of LLM evaluation frameworks, however, there are not specifically focused on digital twins.
Implementation Strategy Create a comprehensive scientific database, evaluation frameworks, integrate it with a generative language model, and fine-tune it with domain-specific data to enhance responses and reliability. Create and refine physics-based generative models tailored to specific applications within DTs. Establish robust validation and verification frameworks involving domain experts and advanced evaluation metrics.
Conclusion Integrating trustworthy LLMs with DTs advances scientific research by enabling sophisticated analysis and interaction with vast unstructured data. Addressing LLM challenges through RAG integration enhances DTs' capabilities, supporting various research scenarios and continuous learning. This integration promises to make scientific research more efficient, comprehensive, and impactful.

2 Is it fit for purpose? Understanding valid use and its relation to the value of information

Topic: Business, Focus: Evaluate value/return on investment
The recent report on Foundational Research and Future Directions for Digital Twins [21] underscores the importance of validation for specific purposes, emphasizing that the success of a digital twin hinges on models that accurately represent the physical counterpart, provide predictions with known confidence, and meet computational constraints. Critical yet under-investigated aspects include the quality and quantity of information the digital twin provides and its value. The model must deliver necessary and accurate information to be useful, and the value of this information must exceed the costs of data, construction, and integration. Balancing value relative to cost and the scope of valid use is essential for successful and economical digital twin efforts. Research is needed to provide a robust framework for considering problems of model validation alongside questions of cost and value. While the problem of model validation has been studied extensively, research seeking to understand and quantify the relationship between validation, cost, and value is relatively new.

We envision a research and development program with at least three distinct, but nonetheless intertwined, topics. First, what is the return on investment for simulation? Some of the investment challenges faced by organizations that make extensive use of modeling and simulation are described in a 2011 report "Calculating Return on Investment for U.S. Department of Defense Modeling and Simulation" (Defense Acquisition University (DAU), Defense Acquisition Research Journal, April 2011). Second, when is a digital twin good enough? Distinct from how accurate or precise the model is when providing information, we ask the question of how accurate or precise does it need to be for a specific purpose? Moreover, how can we decide when the necessary accuracy and precision have been satisfactorily demonstrated? Answers to these questions will involve issues of risk incurred by the possibility of insufficient or inaccurate information; statistical questions underlaying a quantifiable approach to risk assessment; and other topics concerning the use of information within the intended context. Third, how much should be invested in a digital twin? What should its scope of use be? Questions of cost, value, and risk are expected to motivate the creation of digital twins that have a more or less narrow scope of use. How broad or narrow a scope is economical while offering an acceptable level of risk?

3 Advanced Data Management for Digital Twins

Topic: Data, Focus: Governance methods for data collection, curation, sharing, and usage
Motivation: Integrating digital twin-based development into experimental, design, and manufacturing processes necessitates combining traditional high-performance computing (HPC) (scale-up) modeling and simulation with distributed (scale-out) machine learning/AI analysis. These workflows demand advanced data management beyond the current capabilities of DOE's leadership computing facilities. Data will be shared between HPC systems and local clouds, and its computation and transfer need to be managed for workflow correctness and performance. Ensuring FAIR (Findable, Accessible, Interoperable, Reusable) compliance will enhance scientific discovery by making digital twin processes more transparent and verifiable.
Strategy: Implementing advanced data management in digital twins requires innovations in several areas. Lifecycle metadata must be attached to data throughout its lifecycle, making it discoverable, queryable, and accessible even in archival storage, with repositories functioning as active computational elements rather than static storage. Automated curation is necessary due to the growing size and complexity of data, necessitating AI-based tools to detect anomalies and manage metadata, ensuring data is ready for digital twin integration. Federated operation is essential for interdisciplinary research, requiring cross-organizational data and metadata management to support large-scale digital twins. Additionally, consistent policies and governance across organizations are needed to guide data access, use, and sharing, supporting the other strategy components.
Current Gaps and Challenges: Current tools capture limited metadata, resulting in fragmented and incompatible data management services. Ad hoc approaches and siloed metadata formats make it difficult to query relationships between data artifacts. The volume, velocity, and variety of data exceed the capacity of human curators, necessitating automated tools to adequately appraise and describe datasets. Establishing federated data management mechanisms is challenging due to the involvement of multiple policy and infrastructure organizations. Additionally, defining and agreeing on cross-organizational policies often pose a barrier to implementing technical solutions.

4 A Digital Twin of Science Ecosystem

Topic: Ecosystem, Focus: Integrated Research Infrastructure
Motivation Developing science workflows across distributed instruments and high-performance computing platforms is complex and time-consuming, requiring resource allocation for development, debugging, and testing, especially in the early stages. Digital twins of science ecosystems facilitate the development and testing of workflows with minimal or no need for physical resources. These twins emulate physical components, supporting remote instrument control, measurement transfer, integration of simulation and analytics mod-

els, ecosystem messaging, and AI applications for autonomous workflow orchestration. Examples include microscopy workflows, beam-line instrument control, and ecosystem and network profile estimation [1, 2].

Current gaps and challenges Developing digital twins for science ecosystems requires addressing the ecosystem infrastructure counterparts of instruments, computing, and network elements to match the physical ecosystem representation. In addition, the software environment of workflow and ecosystem modules, including instrument simulators and network emulators, should be available for integration in the emulated ecosystems utilized for science workflow developments. This type of digital twin also requires powerful computing resources to support high-performance distributed computing and GPU computations.

5 A Global Consortium for Supercomputer Digital Twins

Topic: International, Focus: Opportunities for International Collaborations

The ExaDigiT international community is a grassroots effort to develop an open-source framework for developing digital twins for supercomputers. ExaDigiT has multiple modules for modeling energy consumption, cooling dynamics, system workloads, and visual analytics – including a web dashboard for performing "what-if" experiments and an augmented reality module for interacting with the digital twin [11]. Formed by numerous supercomputing centers and universities globally, ExaDigiT also includes industry partners such as HPE and NVIDIA. It features workgroups on various topics, including AI/ML, power and cooling, and visual analytics. The community hosts events and provides resources for modeling data centers, having developed digital twins of several supercomputers already. For more information, visit https://exadigit.github.io.

6 Process Twins for Decision-Support and Dynamic Energy/Cost Prediction in Water Reuse Processes

Topic: Long Term Research, Focus: Research enabling the bidirectional flow between the virtual and the physical assets

Water treatment systems offer significant opportunities for efficiency optimization, which can minimize energy and chemical usage, reduce waste products, and maximize output, thereby lowering the cost of water for potable, industrial, and agricultural uses. These systems are complex and nonlinear, making them ideal testbeds for evaluating the effectiveness of digital twin technology beyond linear approximations. A digital twin must integrate various disciplines, including physics, chemistry, fluid dynamics, economics, and material science, while adapting to data from the physical asset to optimize processes such as energy expenditure for softened cooling water and preventing scaling in heat exchangers.

The evolution of flight simulators from physical twins to advanced digital models highlights the potential of digital twins in other industries. Historically, obtaining experimental data for flight models involved significant safety risks and costs. Modern digital models, like those used in Formula One, allow for rapid design modifications through virtual experiments, reducing the need for physical prototypes. Applying this to water treatment, a small-scale pilot system can safely generate data to train digital twins, allowing for exploration of operational regimes that would be unsafe in full-scale industrial systems. This approach enhances the accuracy of data-driven methods such as AI/ML, statistical regression, and multi-physics models, addressing the current limitations in predicting behavior outside the data's scope.

The project's collaborative nature, involving universities, software companies, and national laboratories, emphasizes significant advancements in data integration, experimental design, data transmission, and techno-economic analysis. This research aligns with the National Science Foundation's focus areas, providing a comprehensive approach to improving water treatment efficiency through advanced digital twin technology and interdisciplinary collaboration.

This project is designed for secure data flow over a wide-area network between the testbed and infrastructure. It is also a highly complex nonlinear system with unpredictable stochastic variation in the physical system that is challenging to model accurately using the most common general digital models. Accuracy requires real-time digital twin calibration using physical data. The project is designed to enable long-term digital twin research by providing a physical system that is not easy to model but easy to perform experiments on and obtain data from. It can also be manipulated with minimal safety concerns compared to an industrial system.

7 The Challenge of Data Privacy

Topic: Responsible, Focus: Intellectual property and privacy

Context In the era of digital transformation, digital twins have become essential in various areas, including manufacturing, healthcare, and urban planning. Digital twins are virtual replicas of physical objects, processes, systems, or environments, facilitating data analysis and system monitoring to enhance decision-making. But how can data be leveraged to save energy or foster synergies in medical research without experimenting on animals or people? Digital twins might be a solution. By accurately replicating real-world entities, they allow us to predict consequences, identify obstacles, and devise strategies to overcome them and maximize benefits. However, despite their vast potential, digital twins raise significant data privacy concerns that research institutions, businesses, and regulators must carefully address.

Challenges Scientists and engineers need real-world data to create accurate systems, which often involves gathering personal information. Data sensitivity and accessibility are significant concerns, as the data used in digital twins can be highly sensitive, containing personal identifiers that trace back to individuals. For instance, in healthcare, patient-specific data used for personalized treatments can reveal deeply private information. Ensuring that only authorized entities can access sensitive data is crucial but challenging due to the multiple vectors of potential exposure in digital ecosystems. Data integrity is another major concern, as there is potential for data tampering or misuse. Inaccurate or manipulated data could lead to faulty predictions about critical infrastructure, endangering public safety or causing financial losses. Additionally, there is a risk of data being used for purposes other than originally intended. For example, data collected to optimize building energy efficiency might be used to infer the behavior patterns of residents without their consent, leading to privacy violations.

Mitigating the Privacy Risks To mitigate privacy risks, incorporate privacy by design into the development and deployment of digital twins, addressing privacy concerns proactively by integrating privacy and data protection principles from the outset of technology design. Enhanced data protection measures should be utilized, including state-of-the-art cybersecurity technologies such as encryption and blockchain to secure data transmission and storage. Implement strong access control and identity verification systems to ensure that only authorized personnel can access sensitive data.

8 Surrogate and generative models for trustworthy digital twins of complex physical and engineering systems

Topic: Trustworthy, Focus: Trustworthiness of digital twins

Digital twins (DTs) for complex systems require advanced AI to handle multi-source, multi-fidelity data, ensuring precision across diverse scenarios. High-dimensional spaces pose challenges, stressing experimental and HPC facilities. AI accelerates both forward and inverse problem-solving in various applications, including materials modeling, fluid dynamics, and engineering.

Challenges Surrogate deep learning (DL) models must be (1) generalizable to different physical scenarios, (2) transferable across system sizes for scale bridging, and (3) confident in predictions to maintain DT quality. These properties define trustworthy surrogate models. However, stable training with multi-source, multi-fidelity, imbalanced data remains challenging. Reinforcement learning (RL) and generative models (GM) algorithms must propose realistic scenarios for DL models to assess DT component responses. Ensuring feasibility constraints in these scenarios is critical but challenging, especially for complex systems without clear mathematical models.

Proposed solutions Trustworthy DL surrogate models should integrate physics to maintain self-consistency in target properties, either through physical correlations or additional physics laws. Multi-modal learning in GMs can uncover hidden correlations, enhancing realistic scenario generation and scientific discovery. These correlations help efficiently sample new parameters, improving modeling accuracy.

Expected outcome Integrating DL surrogate models with RL and GM in a unified AI workflow will create DTs that offer trustworthy virtual descriptions of complex engineering/physical systems.

9 High Performance Probabilistic Ensemble Digital Twins for Time Dependent Problems

Topic: VVUQ, Focus: Foundational and cross-cutting methods

Understanding and harnessing non-equilibrium physical and engineering systems rely heavily on computing. Mathematically rigorous methods that integrate physics, observations, AI/ML parameterizations, and expert knowledge offer the greatest promise. Probabilistic estimation methods, particularly those leveraging Bayesian techniques, can handle non-Gaussian/nonlinear processes and provide a framework for digital twins. These methods, established since the 1960s in stochastic control and data assimilation, enable interaction between the physical and in-silico worlds, producing results with bounded uncertainty and interpretability [14].

Current Challenges Framework adaptability involves creating a general framework that can balance constructs like meta-stable, Markovian, Bayesian, and frequentist approaches to various problems [25]. Computational challenges include developing stable and efficient probabilistic digital twins that manage evolving probabilities and epistemic errors while ensuring statistically convergent results [15]. Sampling techniques need to define sampling stability and develop efficient methods for high-dimensional problems to generate explainable results [26]. Ensuring interpretability in complex, non-equilibrium problems requires addressing multi-modal and multi-scale statistical challenges [28]. Tackling the curse of dimensionality involves employing dimension reduction techniques for models, data, and observations. Additionally, managing pervasive biases and unknowns necessitates achieving convergence for fidelity exchange [24].

10 Uncertainty Quantification and High-performance Computing

Topic: VVUQ, Focus: Foundational and cross-cutting methods

Motivation Physical systems inherently exhibit randomness and uncertainty, complicating simulation-based or data-driven predictions. Robust and scalable Uncertainty Quantification (UQ) techniques are essential, particularly in HPC environments, to bridge the "predictive/validated HPC" gap. This is crucial for national priorities like fusion energy and drug discovery in digital twins. Automated UQ tools can connect the needs of engineers and scientists using supercomputing platforms, creating reliable digital twins that incorporate simulation and real-time data with reduced risk.

Digital Twins and UQ Digital twins use existing tools to create future models that evolve the domain. Success in digital twins varies across applications, but reliability and robustness depend on UQ methods. Coupling digital twins with UQ builds confidence through multiscale, multi-physics simulations, improving VVUQ performance and prediction fidelity. HPC is vital for these calculations, balancing edge and supercomputer computations for precision.

Challenges and Requirements Massive computing power, mathematical modeling expertise, and scientific domain knowledge are needed to develop problem-solving digital twins. An interdisciplinary approach ensures successful uncertainty assessment and informed decision-making. Long-term investment in reproducible codes is crucial.

Specific HPC Challenges Current HPC practices lack clear paths for coupling diverse CPU/GPU usage across platforms. Opportunities lie in developing interfaces for physics and engineering code coupling, such as in-memory coupling and standardized file formats. Implementing UQ methods builds confidence in error estimation and design under uncertainty, identifying impactful variables and using influential computations strategically. Addressing experiment variability and minimizing downtime, such as in fusion modeling, are key goals.

Acknowledgment

This research was sponsored by and used resources of the Oak Ridge Leadership Computing Facility (OLCF), which is a DOE Office of Science User Facility at the Oak Ridge National Laboratory (ORNL) supported by the U.S. Department of Energy under Contract No. DE-AC05-00OR22725.

References

[1] A. Al-Najjar et al. Enabling autonomous electron microscopy for networked computation and steering. In *2022 IEEE 18th International Conference on e-Science (e-Science)*, pages 267–277. IEEE, 2022.

[2] A. Al-Najjar and N. S. Rao. Virtual infrastructure twin for computing-instrument ecosystems: software and measurements. *IEEE Access*, 11:20254–20266, 2023.

[3] R. Balin and S. Barwey. SimAI-Bench: ALCF benchmarks for coupled simulation and AI workflows, 2024. Available at `https://github.com/argonne-lcf/SimAI-Bench`.

[4] R. Balin et al. In situ framework for coupling simulation and machine learning with application to CFD. *arXiv preprint arXiv:2306.12900*, 2023.

[5] E. M. Bender et al. On the dangers of stochastic parrots: Can language models be too big? In *Proceedings of the 2021 ACM Conference on Fairness, Accountability, and Transparency*, pages 610–623, 2021.

[6] M. Boyer et al. Scalable integration of computational physics simulations with machine learning. In *2022 IEEE/ACM International Workshop on Artificial Intelligence and Machine Learning for Scientific Applications (AI4S)*, pages 44–49. IEEE, 2022.

[7] W. Brewer et al. Production deployment of machine-learned rotorcraft surrogate models on HPC. In *2021 IEEE/ACM Workshop on Machine Learning in High Performance Computing Environments (MLHPC)*, pages 21–32. IEEE, 2021.

[8] W. Brewer et al. Streaming detection and classification performance of a POWER9 edge supercomputer. In *2021 IEEE High Performance Extreme Computing Conference (HPEC)*, pages 1–7. IEEE, 2021.

[9] W. Brewer et al. Scalable interactive autonomous navigation simulations on HPC. In *2022 IEEE High Performance Extreme Computing Conference (HPEC)*, pages 1–7. IEEE, 2022.

[10] W. Brewer et al. AI-coupled HPC workflow applications, middleware and performance. *arXiv preprint arXiv:2406.14315*, 2024.

[11] W. Brewer et al. A digital twin framework for liquid-cooled supercomputers. In *Proceedings of the International Conference for High Performance Computing, Networking, Storage and Analysis*, 2024.

[12] T. Brown et al. Language models are few-shot learners. *Advances in Neural Information Processing Systems*, 33:1877–1901, 2020.

[13] S. L. Choi et al. eGridGPT: Trustworthy AI in the control room. Technical Report NREL/TP-5D00-87440, National Renewable Energy Laboratory, 2024. `https://www.nrel.gov/docs/fy24osti/87740.pdf`.

[14] G. L. Eyink et al. A statistical-mechanical approach to data assimilation for nonlinear dynamics. *Journal of Statistical Physics*, 2003.

[15] D. Foster and J. M. Restrepo. An improved framework for the dynamic likelihood filtering approach to data assimilation. *Chaos*, 32:053118, 2022.

[16] R. Hu et al. Evaluation of hallucination and robustness for large language models. In *2023 IEEE 23rd International Conference on Software Quality, Reliability, and Security Companion (QRS-C)*, pages 374–382. IEEE, 2023.

[17] H. B. Klasky et al. Comparative assessment of data-driven process models in health information technology. Technical Report ORNL/TM-2021/2284, Oak Ridge National Laboratory (ORNL), Oak Ridge, TN (USA), 2021.

[18] Hyoukjun Kwon et al. XRBench: An extended reality (XR) machine learning benchmark suite for the metaverse. *Proceedings of Machine Learning and Systems*, 5:1–20, 2023.

[19] M. Li et al. ChatTwin: toward automated digital twin generation for data center via large language models. In *Proceedings of the 10th ACM International Conference on Systems for Energy-Efficient Buildings, Cities, and Transportation*, pages 208–211, 2023.

[20] S. Mandvikar. Augmenting intelligent document processing (IDP) workflows with contemporary large language models (LLMs). *International Journal of Computer Trends and Technology*, 71(10):80–91, 2023.

[21] National Academies of Sciences, Engineering, and Medicine. *Foundational research gaps and future directions for digital twins*. 2023.

[22] P. Niloofar et al. A general framework for human-in-the-loop cognitive digital twins. In *2023 Winter Simulation Conference (WSC)*, pages 3202–3213. IEEE, 2023.

[23] S. Partee et al. Using machine learning at scale in numerical simulations with SmartSim: An application to ocean climate modeling. *Journal of Computational Science*, 62:101707, 2022.

[24] J. M. Ramirez et al. Probabilistic measures for biological adaptation and resilience. *Physical Review E*, 109:024413, 2024.

[25] J. M. Restrepo. A dynamic likelihood approach to filtering. *Quarterly Journal of the Royal Meteorological Society*, 2017.

[26] J. M. Restrepo and J. M. Ramirez. Calculating probability densities with homotopy and applications to particle filters. *International Journal of Uncertainty Quantification*, 12:71–89, 2021.

[27] J. Tupayachi et al. Towards next-generation urban decision support systems through AI-powered generation of scientific ontology using large language models – a case in optimizing intermodal freight transportation. *arXiv preprint*, 2024.

[28] S. Venkataramani et al. Dimension reduction in slow relaxation processes. *Journal of Statistical Physics*, 2017.

[29] H. Xu et al. Leveraging generative AI for smart city digital twins: A survey on the autonomous generation of data, scenarios, 3D city models, and urban designs. *arXiv preprint*, 2024.

[30] J. Šturm et al. Enhancing cognitive digital twin interaction using an LLM agent. In *2024 47th MIPRO ICT and Electronics Convention (MIPRO)*, pages 103–107. IEEE, 2024.