

A Digital Twin-Driven Framework for Predictive Maintenance in Complex Engineering Systems

Longfeng Hao

School of Electrical Engineering and Computer Science, Oregon State University, Corvallis,
OR, USA.

hao565@oregonstate.edu

Weday Singht

Department of Computer Science, Binghamton University, Binghamton, NY, USA.

usingh@binghamton.edu

Abstract

The increasing scale and interdependence of complex engineering systems, such as power grids, aerospace fleets, and advanced manufacturing lines, demand a fundamental shift from reactive and scheduled maintenance to predictive strategies capable of anticipating degradation and preventing failures in real time. This paper proposes a comprehensive digital twin-driven framework that integrates multi-layered system architecture, advanced data infrastructure, model governance, and lifecycle sustainability. Departing from narrowly focused technical implementations, the framework is examined through the lens of large-scale systems research, emphasizing structural trade-offs among model fidelity, data fidelity, computational cost, and operational latency. The analysis extends beyond pure engineering to encompass governance mechanisms, algorithmic fairness, policy implications, and the socio-technical resilience of maintenance ecosystems. A conceptual architecture is presented that couples physics-informed and data-driven hybrid twins with a continuous validation layer and a human-in-the-loop decision interface. Particular attention is paid to interoperability across heterogeneous asset fleets, the role of semantic standards, and the edge-cloud continuum that underpins scalable twin deployment. The paper further interrogates the ethical dimensions of predictive maintenance, including liability attribution when autonomous recommendations fail, and the risk of systemic bias against under-instrumented assets. Deployment pathways are assessed from both greenfield and brownfield perspectives, highlighting the tension between rapid digitalization and the preservation of legacy operational knowledge. By synthesizing cross-domain case illustrations from aviation, wind energy, and industrial robotics, the framework's capacity to balance robustness, adaptability, and long-term economic and environmental sustainability is critically evaluated. The article closes with a forward-looking research agenda centered on federated twin ecosystems, generative artificial intelligence for scenario generation, and the establishment of international regulatory benchmarks for digital twin certification in safety-critical maintenance operations.

Keywords

Digital Twin; Predictive Maintenance; Complex Engineered Systems; System Architecture; Cyber-Physical Infrastructure; Sustainability; Governance.

1. Introduction

Modern engineering systems have evolved into highly interconnected socio-technical assemblages where component-level reliability is no longer sufficient to guarantee systemic

resilience. From offshore wind farms spanning hundreds of turbines to continental-scale power transmission networks and commercial aircraft fleets operating under extreme variability, the cost of unscheduled downtime now transcends individual asset economics to threaten energy security, supply chain stability, and public safety. Within this context, predictive maintenance has emerged as a strategic imperative, leveraging data-driven foresight to orchestrate interventions before failures propagate across coupled subsystems. Yet, while machine learning models have demonstrated promising diagnostic and prognostic accuracy in controlled laboratory settings, their translation into operational environments remains fraught with system-level challenges that cannot be resolved by algorithmic sophistication alone. A fundamental gap persists between the theoretical capabilities of predictive analytics and the organizational, infrastructural, and regulatory realities of large-scale asset management. This paper argues that bridging this gap requires a holistic digital twin-driven framework that treats the maintenance lifecycle as a system of systems, wherein physical assets, virtual representations, data infrastructures, human decision-makers, and institutional policies co-evolve under persistent uncertainty.

Digital twins, as high-fidelity virtual replicas that maintain bidirectional information flows with physical counterparts, offer a uniquely integrative substrate for predictive maintenance. Unlike conventional simulation models that operate offline or static data dashboards that merely display sensor streams, digital twins possess the capacity to ingest real-time operational telemetry, update their internal states, and generate forward-looking condition forecasts that are contextually grounded in asset-specific history. However, the contemporary landscape of digital twin research remains fragmented, with industrial case studies often emphasizing sensor instrumentation and visualization, while academic literature tends toward algorithmic optimization in idealized settings. What is missing is a systematic treatment of the architectural, governance, and sustainability dimensions that determine whether a digital twin-driven maintenance strategy can scale across heterogeneous fleets, adapt to evolving asset configurations, and earn the trust of regulatory bodies and frontline personnel. The framework developed in this paper addresses this lacuna by structuring the problem space into five interconnected layers, from physical asset connectivity to institutional governance, and by subjecting each layer to rigorous systems analysis that illuminates trade-offs, failure modes, and pathways toward equitable and robust implementation.

The subsequent sections are organized as follows. Section 2 grounds the discussion in the evolution of the digital twin paradigm, distinguishing its properties from adjacent concepts and establishing the foundational requirements for predictive maintenance in complex systems. Section 3 presents the conceptual architecture, detailing the roles of hybrid modeling, the continuous validation loop, and the human-machine interface. Section 4 examines the data infrastructure and integration challenges that arise when twins must be federated across organizational and semantic boundaries. Section 5 shifts the lens toward governance, policy, and ethical considerations, foregrounding questions of liability, fairness, and algorithmic accountability. Section 6 addresses deployment strategies, sustainability imperatives, and lifecycle management, while Section 7 probes the robustness, fairness, and socio-technical resilience required for long-term viability. Finally, a concluding synthesis outlines cross-cutting insights and a research agenda for the coming decade.

2. The Digital Twin Paradigm in Large-Scale Engineering

The concept of a digital twin has undergone a significant transformation since its early articulation in the context of product lifecycle management, where it denoted a virtual

representation of a manufactured artifact intended to mirror its as-built and as-maintained states [1]. Contemporary interpretations, shaped by the convergence of the Internet of Things, cloud computing, and high-performance simulation, extend this vision to continuous synchronization between physical assets and their virtual counterparts throughout the operational phase. In a large-scale engineering context, a digital twin is not a single monolithic model but a dynamic federation of data, physics-based simulations, statistical surrogates, and domain-specific knowledge graphs that together capture the state, behavior, and degradation trajectory of an asset. Crucially, the twin must also encode the asset's relational context: its position within a larger topology, its interactions with neighboring subsystems, and its exposure to shared environmental and operational stressors. This systemic embedding distinguishes digital twins from mere component-level digital shadows and demands an architectural perspective that explicitly accounts for data provenance, model provenance, and the temporal consistency of distributed states [2].

The promise of digital twins for predictive maintenance lies in their capacity to fuse heterogeneous evidence streams—vibration spectra, thermal images, oil debris counts, operational logs, and historical repair records—into a coherent narrative of asset health that evolves in near real time. Data-driven prognostic models, such as deep recurrent neural networks or gradient-boosted survival trees, can detect subtle precursor patterns that elude threshold-based alarm systems; however, their opacity and susceptibility to spurious correlations present acute risks when applied to safety-critical assets. Physics-informed approaches, grounded in fracture mechanics, tribology, or electro-thermal degradation models, offer higher interpretability and extrapolation capacity but struggle to capture the idiosyncratic wear patterns that emerge under complex, multi-physics loading. The architectural challenge, therefore, is not to choose between data-driven and physics-based paradigms but to design a hybrid twin that dynamically arbitrates between them, leveraging the strengths of each while managing uncertainty propagation across the model chain. This arbitration function must be informed not only by predictive accuracy metrics but also by system-level considerations such as the cost of false positives that trigger unnecessary maintenance dispatches and the cost of false negatives that risk catastrophic failure of a shared infrastructure node.

Related work in the aerospace domain illustrates the interplay of these tensions. Modern engine health management systems already employ a mixture of real-time performance tracking, physics-based gas path analysis, and machine learning classifiers that detect anomalous vibration signatures. Yet field experience reveals that the most consequential failures often occur at the interfaces between subsystems—for instance, when a lubrication system degradation undetected by the primary engine twin propagates into accessory gearbox distress. This observation underscores the need for a framework in which digital twins are not only vertically integrated from sensor to decision support but also horizontally federated so that cross-asset interactions and cascading effects become visible to maintenance planners. The evolution toward such federated ecosystems, however, exposes deep-seated issues of data ownership, intellectual property protection, and semantic interoperability, which are further examined in the subsequent sections.

3. Conceptual Architecture of the Twin-Driven Framework

The proposed framework organizes the predictive maintenance ecosystem into five logically distinct but operationally interpenetrating layers: the physical asset layer, the data acquisition and preprocessing layer, the digital twin modeling layer, the analytics and decision engine

layer, and the human-in-the-loop interface layer. This stratification is designed not as a rigid waterfall decomposition but as a conceptual skeleton that surfaces the critical points of coupling, contention, and negotiation among stakeholders. At the physical asset layer, a diverse constellation of sensors, embedded controllers, and legacy instrumentation systems produces continuous streams of heterogeneous data whose fidelity, temporal resolution, and semantic clarity vary widely across asset classes and vintages. The data acquisition layer must provide a resilient ingestion fabric that accommodates both brownfield retrofits, where non-intrusive sensor kits are deployed alongside existing supervisory control and data acquisition systems, and greenfield designs where twin-ready instrumentation is built into the asset from the outset [3].

The digital twin modeling layer constitutes the analytical core of the framework. Rather than adopting a single-model strategy, the architecture employs an ensemble of specialized sub-models—physics-based reduced-order models, data-driven surrogate models trained on high-fidelity simulation outputs, and hybrid models that embed domain constraints into neural network topologies—each responsible for capturing a distinct degradation modality. A meta-model orchestration layer selects, weights, and blends sub-model outputs based on real-time estimates of model confidence and operational context. This design reflects a crucial trade-off: increasing the granularity of sub-models improves diagnostic specificity but raises computational cost and risk of overfitting, while overly coarse models provide computational efficiency at the expense of missing rare but critical failure signatures. The framework therefore incorporates an automated model selection policy that adjusts the complexity of the active twin ensemble in response to the asset's proximity to known failure boundaries, enabling a graceful degradation of predictive capability rather than a brittle performance cliff.

Surrounding the modeling layer is a continuous validation and recalibration loop that treats digital twins as living entities subject to drift and decay. As physical assets undergo repairs, retrofits, and gradual wear that alter their baseline signatures, the twin must assimilate new evidence while preserving the memory of rare events that remain essential for accurate tail-risk estimation. This is accomplished through a combination of online Bayesian updating, periodic retraining triggered by cumulative error thresholds, and adversarial stress-testing routines that challenge the twin with synthetic extreme scenarios generated from historical anomaly distributions [4]. The analytics and decision engine consumes twin outputs and translates them into probabilistic maintenance recommendations, resource allocation schedules, and risk-annotated alerts. Critically, these recommendations are not presented as deterministic commands but as decision-support artifacts accompanied by uncertainty quantifications, confidence intervals, and explicit statements of the evidence windows underpinning each forecast. The final layer, the human-in-the-loop interface, is designed to foster appropriate reliance by presenting explainable justifications, enabling what-if counterfactual exploration, and capturing operator overrides as valuable feedback signals that loop back into model refinement and organizational learning.

Across all layers, the architecture enforces modularity and standardized application programming interfaces to prevent vendor lock-in and to allow independent evolution of sensor technologies, model libraries, and visualization dashboards. This modularity is not without cost, as it introduces latency and potential semantic misalignment at each interface boundary. The framework explicitly acknowledges these costs and offers governance mechanisms, including interface versioning policies and automated conformance testing, that are elaborated in the later discussion of deployment and policy.

4. Data Infrastructure and Integration Challenges

The success of any digital twin-driven maintenance framework is contingent upon a data infrastructure that can reconcile the competing demands of real-time responsiveness, historical depth, and cross-organizational sharing. Complex engineering systems are typically operated by multiple entities—asset owners, maintenance service providers, original equipment manufacturers, and regulatory inspectors—each maintaining their own data silos with idiosyncratic formats, sampling rates, and access restrictions. Achieving a unified view of asset health therefore requires a semantic integration layer that maps these heterogeneous sources onto a common ontology of failure modes, components, and operational contexts. Industrial reference architectures, such as the Reference Architecture Model for Industry 4.0 and the Industrial Internet Reference Architecture, provide conceptual blueprints for this integration, but their instantiation in live operational environments remains a formidable engineering challenge [5]. The proposed framework advocates for the adoption of open, machine-readable asset administration shells that encapsulate not only static design specifications but also dynamic links to twin services, calibration histories, and maintenance records, thereby enabling automated discovery and composition of digital twins across system boundaries.

The data infrastructure must also address the fundamental tension between the localization of analytics at the network edge and the centralization of models in the cloud. Edge computing minimizes latency for time-critical anomaly detection and reduces the burden on bandwidth-constrained communication links, but it restricts the horizon of learning to local asset data, potentially missing fleet-wide degradation patterns that emerge only when thousands of similar assets are compared. A hybrid edge-cloud architecture is thus essential: lightweight twin surrogates operating at the edge perform immediate condition assessment and trigger alerts, while full-fidelity twin instances in the cloud assimilate fleet-wide data, update global models, and push refined parameters back to the edge during periodic synchronization windows [6]. The management of this bidirectional model flow demands robust versioning, rollback capabilities, and security protocols that prevent adversarial injection of corrupted model updates, an issue that gains urgency as predictive maintenance systems become attractive targets for cyber-physical sabotage.

Data quality and provenance are additional pillars of the integration strategy. Sensors in harsh industrial environments are subject to drift, fouling, and intermittent failure, leading to data streams that are riddled with outliers, missing values, and systematic biases. The framework embeds a data health monitoring sub-layer that continuously audits incoming telemetry against physical plausibility constraints and statistical consistency checks, flagging suspect data before it contaminates twin state estimates. Moreover, every data point that informs a maintenance decision carries a provenance record that traces its lineage through the ingestion, transformation, and modeling pipeline, enabling post-hoc forensic analysis when decisions are contested. This emphasis on provenance is not merely a technical nicety but a foundational requirement for regulatory audits and liability adjudication, as discussed in the governance section.

Cross-domain comparisons further illuminate integration imperatives. In offshore wind energy, where turbines from multiple manufacturers are often co-located within a single farm, the lack of standardized data interfaces has historically forced operators to maintain separate condition monitoring systems for each turbine brand, fragmenting the fleet-wide view and obscuring correlated degradation caused by shared wake effects and meteorological

conditions. By contrast, certain aerospace consortia have made progress toward data pooling agreements that allow airlines to contribute anonymized engine performance data to a shared twin repository, yielding collective prognostic models that outperform those trained on any single operator's data. These contrasting trajectories underscore the importance of governance structures that align incentives for data sharing while protecting commercially sensitive information, a theme expanded in the following section.

5. Governance, Policy, and Ethical Dimensions

Predictive maintenance decisions are not purely technical; they are deeply embedded in networks of contractual obligations, safety regulations, and ethical responsibilities. When a digital twin recommends deferring an inspection to a later date, it implicitly accepts a share of the risk, yet current legal frameworks rarely attribute liability to algorithmic agents or their developers. The framework therefore advocates for a governance model in which each predictive recommendation is accompanied by a documented risk statement, confidence bounds, and a clear delineation of residual uncertainties that remain the responsibility of human decision-makers [7]. This model aligns with the emerging principle of meaningful human control in safety-critical autonomous systems and recognizes that maintenance professionals possess contextual knowledge—about impending weather windows, crew availability, or supply chain disruptions—that cannot be fully encoded into a digital model.

Algorithmic fairness constitutes another governance dimension that has received insufficient attention in the predictive maintenance literature. In a fleet comprising assets of varying ages, instrumentation levels, and operational histories, data-driven prognostic models naturally exhibit higher accuracy for well-instrumented assets and poorer performance for legacy equipment with sparse sensor coverage. Left unaddressed, this disparity can lead to a systematic bias wherein maintenance resources are preferentially allocated to modern assets, accelerating the deterioration of older assets and widening reliability gaps over time. The proposed framework mitigates this risk through fairness-aware model training that penalizes accuracy disparities across asset cohorts and through transparent reporting of per-cohort performance metrics that enable operators to make equity-conscious resource allocation decisions [8]. Such measures, while imposing a modest cost on aggregate prognostic accuracy, are essential for sustaining the legitimacy of twin-driven maintenance programs across diverse stakeholder groups.

Regulatory policy for digital twins in predictive maintenance is still in its formative stages. Existing safety standards, such as those governing software in airborne systems and industrial process control, provide partial guidance but do not address the unique characteristics of twins that continually evolve through online learning [9]. A forward-looking policy posture would require digital twins used for maintenance credit—the permission to extend inspection intervals based on twin health assessments—to undergo a certification process analogous to the type certification of physical components. This process would verify the twin's architecture, validation protocols, and change management procedures, while accepting that the twin's internal parameters will adapt over time within approved bounds. International harmonization of such certification frameworks is particularly urgent for cross-border infrastructures such as aviation and maritime transport, where inconsistencies in regulatory acceptance can fragment global maintenance ecosystems and undermine safety.

Beyond certification, the framework raises the question of twin data as a semi-public good. Operating data from critical national infrastructure assets, such as grid transformers or water pumping stations, possesses value that extends beyond the immediate asset owner to system

planners, emergency responders, and researchers studying long-term degradation trends. Governance mechanisms that enable data sharing under well-defined privacy, security, and reciprocity agreements can unlock this societal value while protecting commercial interests. The framework envisions a tiered data access model in which immediate operational data remains proprietary, aggregated fleet statistics are shared with regulators, and curated research datasets are released under controlled access protocols, thereby balancing the imperatives of innovation, safety, and competition.

6. Deployment, Sustainability, and Lifecycle Management

Deploying a digital twin-driven predictive maintenance framework in an established industrial environment poses a set of transition risks that extend beyond technical integration. Brownfield deployments must contend with legacy programmable logic controllers, proprietary communication buses, and paper-based maintenance logs that resist straightforward digitization. The framework advocates an incremental deployment strategy that begins with low-intrusion sensor retrofits and open-source edge gateways, progressively building the data foundation while maintaining the existing maintenance workflow as a safety net. During this transitional phase, the twin operates in a shadow mode, generating predictions that are compared against actual maintenance outcomes without directly influencing scheduling decisions, thereby building trust through demonstrated accuracy while revealing gaps in data quality and model coverage. Greenfield deployments, by contrast, offer the opportunity to co-design the physical asset, its instrumentation, and its digital twin from the outset, allowing for optimal sensor placement, embedded model hosting, and native support for the twin's application programming interfaces. The co-design process introduces its own trade-offs, as decisions made during early-stage engineering lock in certain twin architectures for the asset's entire lifecycle, underscoring the need for modularity and upgradeable computing hardware.

Sustainability is an often overlooked dimension of digital twin deployment. The computational footprint of continuously running fleet-wide twins, with their high-frequency data ingestion, model inference, and cloud synchronization, can be substantial. A single large gas turbine twin might generate several terabytes of data per day and require dedicated graphical processing unit clusters for real-time uncertainty quantification. The framework incorporates green twin principles that promote model compression through knowledge distillation, selective data sampling based on information-theoretic criteria, and the use of federated learning to train global models without centralizing raw data, thereby reducing communication energy and storage overhead [10]. Furthermore, the framework aligns with circular economy objectives by enabling life-extension decisions grounded in accurate remaining useful life projections, thus delaying asset replacement and reducing material throughput.

Lifecycle management of the digital twin itself is equally critical. As assets age, their failure modes evolve, rendering earlier twin configurations obsolete. The framework mandates a structured twin retirement and replacement governance process, wherein each model component is tagged with a validity horizon and subjected to periodic review. When a twin component is retired, its learned parameters and historical prediction records are archived to preserve institutional memory for future forensic analysis and for training next-generation models [11]. This approach treats the digital twin not as a disposable software artifact but as a persistent knowledge asset whose value accumulates over the asset's operational life.

Economic viability remains the ultimate arbiter of framework adoption. The framework supports a layered business case analysis that separates the fixed costs of infrastructure establishment from the variable costs of twin maintenance and the uncertain benefits of avoided failures. By providing transparent uncertainty quantification, the framework enables operators to conduct value-of-information assessments that compare the expected cost of additional sensors or model fidelity improvements against the expected reduction in maintenance expenditure and downtime risk. In industries where downtime costs dwarf maintenance savings, such as semiconductor fabrication or data center operations, these assessments often justify substantial investment in high-fidelity twins, whereas in lower-margin sectors the framework can be scaled down to lightweight anomaly detection proxies that retain core architectural principles while curtailing computational expense.

7. Robustness, Fairness, and Socio-Technical Resilience

Robustness in the context of predictive maintenance twins extends beyond model accuracy under stationary conditions to encompass graceful performance degradation under distributional shifts, sensor attrition, and adversarial perturbations. Complex engineering systems routinely encounter novel operating regimes—extreme weather events, emergent grid configurations, new fuel blends—that lie outside the training distribution of data-driven prognostic models. The framework addresses this challenge through a layered defense-in-depth strategy. First, physics-informed constraints and domain invariants serve as guardrails that detect and reject predictions that violate fundamental conservation laws or material limits. Second, an ensemble of models trained on diverse subsets of historical data and employing distinct inductive biases provides a measure of redundancy, allowing a majority-vote or uncertainty-weighted consensus to suppress outlier predictions [12]. Third, the continuous validation loop triggers a heightened monitoring state when the twin’s self-assessed confidence drops below a threshold, automatically involving human analysts and reducing reliance on automated recommendations until confidence is restored. This strategy reflects a system-level acknowledgment that in safety-critical contexts, a timely acknowledgment of ignorance is more valuable than a confident but erroneous prediction.

Fairness, already introduced in the governance discussion, also manifests as a robustness concern. When maintenance optimization algorithms allocate inspection crews and spare parts across geographically dispersed assets, they must balance efficiency with equitable coverage. An algorithm that systematically defers maintenance on assets in economically disadvantaged regions because those assets are older, less instrumented, or associated with higher model uncertainty may reinforce existing infrastructure inequalities. The framework therefore includes an equity-aware scheduling module that allows operators to visualize the distributional consequences of alternative maintenance policies and to impose fairness constraints that guarantee minimum service levels for under-resourced asset cohorts [13]. This module does not prescribe a single fairness metric but rather surfaces the tension between aggregate cost minimization and distributional justice, leaving the final trade-off to human policymakers accountable to their constituencies.

Socio-technical resilience encompasses the capacity of the maintenance ecosystem to absorb disruptions not only in physical assets but also in the human and organizational fabric that surrounds them. Turnover of experienced maintenance personnel, mergers and acquisitions, and shifts in regulatory posture can abruptly alter the context in which digital twins operate. A twin framework that is tightly coupled to a specific vendor’s software stack or to a particular team’s tacit knowledge may become brittle in the face of such changes. The proposed

framework mitigates this brittleness through open interfaces, comprehensive documentation of model assumptions, and the integration of operator knowledge elicitation routines that continuously translate experiential insights into structured rules and Bayesian priors [14]. Moreover, the framework emphasizes the role of digital twins as boundary objects that mediate communication among reliability engineers, data scientists, business managers, and regulators, fostering a shared understanding of asset health that transcends disciplinary jargon. This communicative function becomes particularly salient when unexpected failures occur and cross-functional teams must collectively reconstruct the decision chain that led to the event, a process that is greatly facilitated by the provenance trails embedded in the twin's decision support outputs.

Manufacturing environments provide instructive illustrations of these socio-technical dynamics. In smart factories, digital twin frameworks have been prototyped for predictive maintenance of production equipment, integrating real-time data with simulation models to anticipate spindle degradation and robot joint wear [15]. Field studies in such settings have revealed that the most successful deployments are those in which shop-floor operators are actively involved in configuring alert thresholds and in providing feedback on false alarms, rather than being passive recipients of algorithmic directives. This participatory approach not only improves model performance by injecting tacit knowledge but also builds the organizational trust necessary for the twin's recommendations to be acted upon under time pressure. The framework generalizes this insight by mandating that every human-in-the-loop interface includes a feedback channel through which overrides, comments, and contextual annotations are captured and used for continuous model refinement and workforce development.

8. Conclusion

This paper has presented a digital twin-driven framework for predictive maintenance that transcends the algorithmic focus of much existing research to engage with the systemic, architectural, governance, and human dimensions that ultimately determine real-world impact. The five-layer conceptual architecture—spanning physical assets, data infrastructure, hybrid modeling, decision analytics, and human interaction—embeds within it a set of design principles that prioritize modularity, uncertainty transparency, fairness, and lifecycle sustainment. The analysis has illuminated several key structural trade-offs: between the fidelity of physics-based models and the adaptability of data-driven surrogates; between the latency advantages of edge processing and the learning depth of centralized fleet twins; between the aggregate cost efficiency of optimized maintenance scheduling and the equity imperative of uniform service levels across asset cohorts; and between the regulatory desire for frozen, certifiable models and the operational necessity of continual learning in the face of evolving asset conditions.

By grounding these trade-offs in cross-domain illustrations drawn from aviation, wind energy, and industrial robotics, the paper has demonstrated that the proposed framework is not an abstract ideal but a practical scaffolding that can be scaled, tailored, and progressively adopted across diverse industrial contexts. The governance analysis further highlighted that digital twin-driven maintenance is not merely a technical upgrade but a socio-institutional transformation that demands new forms of liability attribution, new certification regimes, and new data-sharing compacts between public and private stakeholders.

Looking forward, the research agenda must tackle the integration of generative artificial intelligence for scenario generation and synthetic failure mode discovery, which promises to

extend the robustness of prognostic twins beyond historically observed failures. Equally urgent is the development of interoperable standards for federated twin ecosystems that span organizational boundaries, enabling the collective intelligence of entire industrial sectors to be harnessed while preserving competitive sensitivities. The framework's emphasis on sustainability also calls for quantitative lifecycle assessments that compare the environmental footprint of twin infrastructure against the footprint of the failures it prevents, ensuring that the digital transformation of maintenance contributes positively to global sustainability goals. Realizing this vision will require sustained interdisciplinary collaboration among systems engineers, computer scientists, social scientists, and policymakers, united by the recognition that predictive maintenance in complex engineering systems is a grand challenge for which no single discipline holds all the answers.

References

1. Grieves, M., & Vickers, J. (2017). Digital twin: Mitigating unpredictable, undesirable emergent behavior in complex systems. In F.-J. Kahlen, S. Flumerfelt, & A. Alves (Eds.), *Transdisciplinary perspectives on complex systems* (pp. 85–113). Springer.
2. Lee, J., Bagheri, B., & Kao, H. A. (2015). A cyber-physical systems architecture for industry 4.0-based manufacturing systems. *Manufacturing Letters*, 3, 18–23.
3. Tao, F., Zhang, M., & Nee, A. Y. C. (2019). *Digital twin driven smart manufacturing*. Academic Press.
4. Renganathan, K., Duraimutharasan, N., & Venkataraman, R. (2020). Online model validation and recalibration for digital twins in industrial applications. *Journal of Manufacturing Systems*, 56, 217–228.
5. International Electrotechnical Commission. (2016). *Industrial-process measurement, control and automation – Reference architecture model for Industry 4.0 (RAMI 4.0) (IEC/PAS 63088:2017)*. IEC.
6. Shi, W., Cao, J., Zhang, Q., Li, Y., & Xu, L. (2016). Edge computing: Vision and challenges. *IEEE Internet of Things Journal*, 3(5), 637–646.
7. Winfield, A. F. T., & Jirotko, M. (2017). The case for an ethical black box. In M. I. Aldinhas Ferreira, J. Silva Sequeira, & R. Ventura (Eds.), *Cognitive architectures* (pp. 265–279). Springer.
8. Mehrabi, N., Morstatter, F., Saxena, N., Lerman, K., & Galstyan, A. (2021). A survey on bias and fairness in machine learning. *ACM Computing Surveys*, 54(6), Article 115.
9. RTCA, Inc. (2012). *Software considerations in airborne systems and equipment certification (DO-178C)*. RTCA.
10. Li, T., Sahu, A. K., Talwalkar, A., & Smith, V. (2020). Federated learning: Challenges, methods, and future directions. *IEEE Signal Processing Magazine*, 37(3), 50–60.
11. Boschert, S., & Rosen, R. (2016). Digital twin—The simulation aspect. In P. Hehenberger & D. Bradley (Eds.), *Mechatronic futures* (pp. 59–74). Springer.
12. Lakshminarayanan, B., Pritzel, A., & Blundell, C. (2017). Simple and scalable predictive uncertainty estimation using deep ensembles. In *Advances in Neural Information Processing Systems 30* (pp. 6402–6413). Curran Associates, Inc.

13. Barocas, S., Hardt, M., & Narayanan, A. (2019). *Fairness and machine learning: Limitations and opportunities*. MIT Press.
14. Klein, G., Phillips, J. K., Rall, E. L., & Peluso, D. A. (2007). A data-frame theory of sensemaking. In R. R. Hoffman (Ed.), *Expertise out of context* (pp. 113–155). Lawrence Erlbaum Associates.
15. Tao, F., Zhang, H., Liu, A., & Nee, A. Y. C. (2019). Digital twin in industry: State-of-the-art. *IEEE Transactions on Industrial Informatics*, 15(4), 2405–2415.
16. Tuegel, E. J., Ingraffea, A. R., Eason, T. G., & Spottswood, S. M. (2011). Reengineering aircraft structural life prediction using a digital twin. *International Journal of Aerospace Engineering*, 2011, Article 154798.
17. Glaessgen, E. H., & Stargel, D. S. (2012). The digital twin paradigm for future NASA and U.S. Air Force vehicles. In *53rd AIAA/ASME/ASCE/AHS/ASC Structures, Structural Dynamics and Materials Conference (AIAA 2012-1818)*.
18. Susto, G. A., Schirru, A., Pampuri, S., McLoone, S., & Beghi, A. (2015). Machine learning for predictive maintenance: A multiple classifier approach. *IEEE Transactions on Industrial Informatics*, 11(3), 812–820.
19. Wang, J., Ye, L., Gao, R. X., Li, C., & Zhang, L. (2019). Digital twin for rotating machinery fault diagnosis in smart manufacturing. *International Journal of Production Research*, 57(12), 3920–3934.
20. Zio, E. (2009). Reliability engineering: Old problems and new challenges. *Reliability Engineering & System Safety*, 94(2), 125–141.
21. Jardine, A. K. S., Lin, D., & Banjevic, D. (2006). A review on machinery diagnostics and prognostics implementing condition-based maintenance. *Mechanical Systems and Signal Processing*, 20(7), 1483–1510.
22. European Union Agency for Cybersecurity. (2023). *Digital twins in critical infrastructures: Security and resilience perspectives*. ENISA.
23. Siddiqui, F., Kahandawa, G., & Hewawasam, H. (2021). Digital twin-enabled predictive maintenance for industrial assets: A systematic literature review. *IEEE Access*, 9, 127230–127250.