A Framework for Enhancing Self-Generating Digital Twins with Hierarchical Models, Physics-Informed Machine Learning, Gen-AI, and Dynamic Diffusion Models for Real-Time SHM and Predictive Maintenance

SHADY ADIB

ABSTRACT

This paper introduces a conceptual framework to advance Self-Generating Digital Twins (SGDTs) by integrating Hierarchical Digital Twin (HDT) architectures, Physics-Informed Machine Learning (PIML), Generative Artificial Intelligence (Gen-AI), and dynamic diffusion models. Current SGDT implementations face significant challenges in real-time Structural Health Monitoring (SHM) due to the complexity of processing vast, multi-dimensional data streams and the inherent difficulties in achieving accurate predictive insights across large-scale systems. The reliance on a singular SGDT framework often results in inefficiencies when addressing multi-level structural complexities and managing uncertainties, highlighting the necessity for a scalable and layered approach. The proposed framework leverages HDTs to decompose global and local structural behaviours into interconnected layers, simplifying computational demands while enhancing interpretability. By embedding fundamental physical laws and constraints into the learning process, PIML improves the efficiency and accuracy of Gen-AI models. Gen-AI autonomously refines the digital twin models, synthesising actionable insights from real-time data streams, while dynamic diffusion models facilitate precise damage evolution predictions under varying operational and environmental conditions. This proposal addresses critical gaps in real-time SHM by overcoming the limitations of standalone SGDT systems and introducing a methodology that is both scalable and adaptable. It offers a novel pathway for optimised resource allocation, reduced downtime, and improved sustainability in infrastructure management. Although validation through real-world application is a future objective, the framework provides a robust foundation for advancing SHM technologies to meet the growing demands of mobility, autonomy, and sustainability.

INTRODUCTION

The increasing complexity of modern infrastructure systems, along with the growing demand for sustainability, resilience, and operational intelligence, has intensified interest in the development and deployment of Digital Twin (DT) technology. A DT refers to a digital replica of a physical asset that is continuously updated with real-time data and capable of reflecting the asset's current state, performance, and environmental interactions [1, 2]. Originally conceived for aerospace and manufacturing, DTs are now being adopted in the civil engineering domain, particularly for SHM, where they enable condition-based assessment, life-cycle planning, and predictive maintenance of critical assets [3].

Despite this progress, traditional DT models often fall short in high-demand environments due to their static design, limited scalability, and inadequate integration of phys-

Shady Adib, Researcher in Digital Twins, Email: shadyadib93@outlook.com. PhD in Civil Engineering, Department of Civil and Geospatial Engineering, Newcastle University, Newcastle upon Tyne, United Kingdom.

ical laws. The concept of SGDTs has emerged to overcome these limitations, introducing autonomous updating and continuous learning capabilities. These SGDTs aim to respond dynamically to real-time data streams from sensors and monitoring devices without requiring manual intervention [4]. However, most SGDT implementations still face major challenges related to system complexity, model generalisation, and uncertainty management.

A significant limitation arises from treating infrastructure systems as monolithic entities, which restricts the ability of digital twins to represent multi-scale structural behaviour. Moreover, although machine learning models are often employed to improve predictive capabilities, they frequently disregard the governing physics of the system, leading to poor extrapolation and reduced trustworthiness under diverse operational conditions [5]. Furthermore, the influx of high-frequency, high-dimensional sensor data imposes computational burdens that impair real-time analysis and decision-making.

This paper proposes a new conceptual framework for SGDTs that integrates several advanced computational paradigms into a cohesive system. The framework adopts a hierarchical approach to DT development, allowing the decomposition of global infrastructure behaviour into modular and interconnected subsystems. This enhances scalability and provides a clearer understanding of how local changes impact system-level responses [6]. To improve the generalisability and physical fidelity of machine learning models, the framework employs PIML, which embeds physical laws directly into the learning process and reduces reliance on purely data-driven techniques [7].

To further enhance the system's autonomy, Gen-AI is introduced to continuously refine DT models, synthesise missing or corrupted sensor data, and propose new operational insights based on historical and real-time information [8]. Finally, dynamic diffusion models are integrated to enable spatio-temporal reasoning and damage prediction, capturing how structural changes evolve across time and space under variable environmental and loading conditions [9].

By combining these complementary components, the proposed SGDT framework addresses the shortcomings of existing DT implementations in SHM. It supports real-time monitoring, predictive maintenance, and intelligent decision-making, while promoting scalable architecture and interpretability. The result is a robust foundation for advancing infrastructure resilience, minimising unplanned downtime, and aligning infrastructure management with the goals of digital transformation and climate-aware sustainability.

The remainder of the paper is organised as follows: Section 2 presents the methodology, describing the theoretical and computational foundations that support the integration of sensing, physics-informed learning, generative modelling, and diffusion-based reasoning. Section 3 introduces the proposed framework, outlining its layered architecture and the interaction between its key components. Section 4 provides a discussion of the framework's expected outcomes, practical implementation considerations, and potential applications in real-world SHM systems. Finally, Section 5 concludes the paper and outlines directions for future research, including validation strategies and deployment pathways.

METHODOLOGY

The methodology underpinning this study is designed to overcome the limitations of existing SHM approaches by developing a next-generation SGDT framework. This framework combines multiple complementary technologies to achieve real-time monitoring, intelligent anomaly detection, and predictive maintenance of civil infrastructure assets. The methodological approach is grounded in the integration of four pillars: hierarchical system modelling, physics-informed learning, generative modelling, and uncertainty-aware propagation dynamics. Data acquisition begins with sensor instrumentation embedded in the infrastructure. These include accelerometers, strain gauges, thermocouples, and displacement sensors, which are installed across key locations of the physical structure. The data from these sensors is streamed continuously and preprocessed for noise reduction, time synchronisation, and feature extraction. This serves as the primary input into the SGDT system.

The first computational step involves the implementation of a HDT architecture, which allows for the structure to be decomposed into multiscale subsystems. Rather than relying on a monolithic digital replica, the HDT captures system behaviour at both the global and local levels. For instance, in a bridge application, the global model might represent the entire deck response under traffic loads, while local models might focus on critical joints or supports. This structure supports parallel processing, improves scalability, and enhances interpretability by linking specific data patterns to structural elements. Once the HDT is instantiated, the PIML module is activated. Here, physical laws governing the system, such as equilibrium, constitutive relations, and compatibility, are embedded directly into the loss functions of neural networks. This approach guides the learning process so that data-driven predictions adhere to known mechanical behaviours. By using Physics-Informed Neural Networks (PINNs) or other hybrid solvers, the model can produce realistic structural responses even in the absence of dense labelled data. This step reduces overfitting, increases robustness, and ensures compliance with engineering theory.

The Gen-AI layer is then engaged to enhance the self-adaptivity of the SGDT. Using models such as Variational Autoencoders (VAEs) or Generative Adversarial Networks (GANs), the system generates synthetic data streams, fills in missing measurements, and reconstructs unobserved failure scenarios. This generative capability not only strengthens the model's resilience to sensor failures but also enables forward simulation under unseen loading or environmental conditions. Moreover, Gen-AI can be used to simulate structural degradation pathways, training the SGDT to anticipate deterioration before it occurs. Lastly, the dynamic diffusion layer models how damage indicators and anomalies propagate across the structural network. Using graph-based diffusion processes or spatio-temporal partial differential equations, the framework quantifies uncertainty and temporal dynamics of evolving damage states. This enables early-warning mechanisms, load redistribution strategies, and prioritised maintenance recommendations. Overall, the proposed methodology is holistic, modular, and scalable. It bridges physics-based modelling with data-driven intelligence and offers a pathway toward real-time autonomous SHM for future infrastructure systems.

PROPOSED FRAMEWORK

The proposed framework for advancing SGDTs is built upon an integrated, multi-layered architecture that supports real-time SHM and predictive maintenance. It unifies sensing technology, hierarchical modelling, physics-informed learning, generative intelligence, and spatio-temporal forecasting into a cohesive system capable of evolving alongside the physical asset it represents. At its foundation, the framework begins with the physical infrastructure and its associated sensor network. These sensors continuously capture structural responses, environmental conditions, and operational loads. Data collected from the field is transmitted in real-time, forming the primary input stream that drives the behaviour of the DT. The types of data involved include acceleration, strain, displacement, temperature, and other operational indicators relevant to the asset's performance.

Building upon this data foundation is the HDT architecture. Unlike traditional monolithic DTs that attempt to model the entire system as a single entity, HDTs adopt a modular and scalable representation. Each subsystem or critical component within the infrastructure is modelled as an independent but interconnected DT. For example, in the case of a bridge, global models simulate deck-level responses, while local twins focus on bearings, joints, or support elements. This decomposition enables higher computational efficiency, supports parallel model development, and facilitates more targeted diagnostics. Interactions between global and local twins are maintained through clearly defined interface conditions, ensuring coherent system-level behaviour. Above the HDT layer resides the PIML module. This layer addresses a key limitation of standard data-driven models, their tendency to overfit or behave unreliably when extrapolated beyond the training data. By embedding the governing physical laws directly into the architecture of neural networks or into the optimisation process, PIML ensures that the learned models comply with the known principles of structural mechanics. This could include equilibrium equations, material constitutive laws, or energy conservation constraints. The result is a learning system that generalises better, requires less data, and offers interpretable outputs such as damage indicators, stress fields, or modal parameter updates.

To enhance the autonomy of the framework, a Gen-AI module is incorporated. This layer leverages advanced generative models such as VAEs, GANs, or diffusion-based transformers to synthesise missing data, augment sensor readings, and explore unobserved system behaviours. Gen-AI allows the SGDT to evolve continuously, adapt to unseen conditions, and maintain operational awareness even under sparse or corrupted sensor scenarios. Moreover, generative models are useful for simulating failure scenarios or extreme events, supporting proactive planning and resilience assessment. The final component of the framework is the Dynamic Diffusion layer. This layer introduces spatio-temporal forecasting capabilities by modelling how damage or stress concentrations spread across the infrastructure. It accounts for the interconnectivity of structural components and simulates damage propagation using graph-based diffusion equations or dynamic stochastic processes. By integrating uncertainty quantification and time-dependent degradation modelling, this layer supports early warning, prioritisation of inspections, and risk-based maintenance scheduling.

Each module in this framework operates in a continuous feedback loop. Real-time sensor data drives the DT; the twin refines itself through physics-informed learning and generative updates; and the diffusion layer predicts how local events may escalate into broader structural concerns. The result is an intelligent, adaptive, and resilient DT system capable of supporting infrastructure operators in making timely and evidence-based decisions. This proposed framework offers a path forward in overcoming the limitations of conventional SHM systems. The integration of physical principles with AI-driven automation within a hierarchical and dynamic architecture forms a solid foundation for scalable, self-adaptive monitoring systems designed to meet the future demands of infrastructure.

DISCUSSION AND EXPECTED OUTCOMES

The proposed SGDT framework introduces a novel synthesis of hierarchical system modelling, physics-informed learning, generative intelligence, and diffusion-based forecasting for real-time SHM. Its conceptual strength lies in unifying these diverse methodologies into a modular and scalable architecture capable of adapting autonomously to changing structural and environmental conditions. One of the primary advantages of this framework is its hierarchical structure, which enables both local precision and global awareness. This multiscale modelling approach is essential for large civil infrastructure, where local failures can have system-wide implications. By decomposing the asset into interconnected digital subsystems, the framework promotes computational efficiency and facilitates targeted analysis, which would be infeasible using traditional monolithic models.

The integration of PIML further distinguishes this approach. Unlike conventional black-box models, the inclusion of governing physical laws as constraints during training ensures that predictions remain interpretable and physically realistic. This greatly enhances the model's generalisability and reliability, especially in safety-critical applications where interpretability is essential for decision support. Gen-AI adds another layer of resilience and intelligence to the system. Its ability to synthesise plausible sensor data, reconstruct missing information, and simulate unobserved deterioration pathways extends the capabilities of the DT beyond reactive monitoring. This paves the way for predictive and proactive strategies that can reduce maintenance costs, optimise resource allocation, and improve asset longevity.

Meanwhile, dynamic diffusion models contribute to situational awareness by forecasting the spatial and temporal evolution of damage or anomalies. Their role in capturing uncertainty and propagating risk makes them particularly suitable for post-disaster assessments or ageing infrastructure operating under varying climate conditions. The expected outcome of implementing this framework is a highly autonomous SHM system that operates with minimal human intervention while maintaining high levels of accuracy, adaptability, and interpretability. In practice, the system would continuously learn from live sensor data, update its internal models, simulate future failure scenarios, and inform asset managers of potential risks in real time.

Additionally, the proposed SGDT framework opens opportunities for:

- Improved maintenance scheduling through accurate prediction of damage progression.
- Integration with digital asset management platforms and Building Information Modelling (BIM).
- Deployment on edge devices to enable decentralised intelligence for smart infrastructure networks.
- Increased safety and resilience of ageing infrastructure assets.

Despite its promising capabilities, real-world implementation of this framework poses challenges, including data privacy concerns, hardware constraints for edge deployment, and the need for high-quality sensor data over long periods. These will be addressed in future research through targeted validation studies and prototype development in collaboration with infrastructure owners and stakeholders.

CONCLUDING REMARKS

This paper presents a comprehensive and forward-looking framework for SGDTs tailored for real-time SHM and predictive maintenance of critical infrastructure. By integrating HDT, PIML, Gen-AI, and Dynamic Diffusion Models, the framework addresses core challenges in scalability, physical interpretability, data resilience, and predictive capability. The hierarchical architecture ensures that both system-wide and component-level behaviours are accurately captured, while physics-informed learning embeds engineering knowledge into data-driven models to enhance trustworthiness. The inclusion of Gen-AI enables autonomous model updating and data enhancement, and the diffusion layer provides dynamic risk assessment across space and time.

The resulting SGDT system is not merely a static replica of the physical asset but a living, adaptive intelligence capable of evolving with the structure it represents. This approach promises to redefine the landscape of SHM by enabling smarter, faster, and more informed decision-making. Future work will focus on the real-world deployment and validation of the proposed framework, integration with edge and cloud platforms, and the development of open-access tools to facilitate adoption by infrastructure operators. Ultimately, this work contributes to the vision of sustainable, intelligent, and autonomous infrastructure systems aligned with the goals of Industry 5.0 and smart cities.

REFERENCES

1. Grieves, M. and J. Vickers. 2014. “Digital Twin: Manufacturing excellence through virtual factory replication,” *White paper*, 1(1):1–7.
2. Tao, F., H. Zhang, A. Liu, and A. Y. C. Nee. 2018. “Digital twin and big data towards smart manufacturing and Industry 4.0: 360 degree comparison,” *Journal of Manufacturing Systems*, 48:92–106.
3. Bosché, F., Y. Turkan, and C. Haas. 2020. “The critical role of the digital twin in future infrastructure,” *Automation in Construction*, 120:103385.
4. Fuller, A., Z. Fan, C. Day, and C. Barlow. 2020. “Digital twin: Enabling technologies, challenges and open research,” *IEEE Access*, 8:108952–108971.
5. Karniadakis, G. E., I. G. Kevrekidis, L. Lu, P. Perdikaris, S. Wang, and L. Yang. 2021. “Physics-informed machine learning,” *Nature Reviews Physics*, 3(6):422–440.
6. Taylor, C., C. Barlow, A. Cleland, and C. Day. 2020. “A hierarchical digital twin framework for infrastructure systems,” *Smart Infrastructure and Construction*, 1(1):1–12.
7. Raissi, M., P. Perdikaris, and G. E. Karniadakis. 2019. “Physics-informed neural networks: A deep learning framework for solving forward and inverse problems involving nonlinear partial differential equations,” *Journal of Computational Physics*, 378:686–707.
8. Bond, R., D. Ramos, and C. King. 2021. “Deep generative modeling for digital twin data augmentation,” *Applied Intelligence*, 51(5):3024–3036.
9. Baraldi, P., F. Mangili, and E. Zio. 2023. “Diffusion models for structural health forecasting under spatio-temporal uncertainty,” in *Proceedings of the European Safety and Reliability Conference (ESREL)*, ESRA.