ISSN: 2581-3404 (Online) **IF: 5.445 (SJIF)**

International Journal of Innovative Research in Technology and Management, **Volume-6, Issue-5, August-2022.**

AI-Powered Digital Twins: Applications Across Industries

Lalit Mohan[1], Mohit[2], Nikhil Mor[3], Akshay[4], Pooja Sharma[5]

Department of Computer Science and Engineering (AI & ML) [1,2,3,4,5]

Panipat Institute of Engineering and Technology (PIET), Samalkha, Panipat, Haryana, India[1,2,3,4,5]

2820266.cse@pietgroup.co.in[1], 2820279.cse@pietgroup.co.in[2], 2820291.cse@pietgroup.co.in[3], 2820304.cse@pietgroup.co.in[4], pooja.cse-aiml@piet.co.in[5]

Abstract: *A digital twin (DT) is a continuously updated virtual replica of a physical entity, process, or system, connected to its counterpart through a live stream of sensor data. While early digital twins were largely descriptive, mirroring current physical state for monitoring purposes, the integration of artificial intelligence has transformed them into predictive and increasingly prescriptive systems capable of forecasting failures, simulating interventions, and, most recently, communicating with human operators in natural language through large language model (LLM) integration. This paper reviews the architecture and application landscape of AI-powered digital twins across four domains: smart manufacturing, where AI-driven twins enable predictive maintenance and process optimization; smart cities, where urban-scale twins support traffic, energy, and infrastructure planning; healthcare, where patient-specific twins support diagnostics, surgical rehearsal, and treatment forecasting; and education, where learner-facing twins support personalized, adaptive instruction. We describe the common architectural layers, physical sensing, data integration, AI-driven modeling, and simulation and visualization, that recur across these domains, and we discuss the emerging role of LLMs as a cognitive and conversational layer that lets digital twins explain their state and reasoning to non-expert users. We close with a discussion of cross-domain challenges, including interoperability, real-time synchronization, and data privacy, and outline directions for future research.*

Keywords: digital twin, artificial intelligence, smart manufacturing, smart cities, healthcare, education, predictive maintenance, large language models.

Introduction

A digital twin is a virtual representation of a physical object, process, or system that is kept synchronized with its physical counterpart through a continuous flow of sensor data, and that can, in turn, be used to monitor, simulate, and inform decisions about that counterpart [1]. The concept originated in aerospace and product-lifecycle engineering but has since spread across nearly every sector, driven by the falling cost of IoT sensing, the maturity of cloud and edge computing, and, most recently, the integration of artificial intelligence models that let a twin move beyond passive mirroring toward prediction and, increasingly, natural-language explanation [2].

AI enters the digital twin pipeline at several points. Machine learning models trained on historical sensor data allow a twin to forecast future states, such as predicting when a machine part will fail, well before that failure would be detectable from raw sensor readings alone. Reinforcement learning and optimization algorithms let a twin recommend or automatically execute interventions, closing the loop between sensing and action. And, most recently, large language models are being layered onto digital twin platforms as a conversational interface, letting engineers or clinicians ask a twin plain-language questions about its state and receive plain-language, evidence-grounded answers [3].

This paper surveys AI-powered digital twins across four representative application domains, smart manufacturing, smart cities, healthcare, and education, and describes the shared architectural pattern underlying all of them. Section II covers background and architecture. Section III through Section VI cover manufacturing, smart cities, healthcare, and education respectively. Section VII discusses the emerging role of LLMs across all four domains. Section VIII covers cross-domain challenges, Section IX future directions, and Section X concludes.

Background and Architecture

Across domains, AI-powered digital twins share a common four-layer architecture. The physical layer consists of the sensors, actuators, and IoT devices instrumenting the real-world asset. The data-integration layer collects, cleans, and synchronizes this sensor data, often across heterogeneous formats and protocols, into a unified data model. The AI-modeling layer applies machine learning, most commonly for anomaly detection, failure prediction, or state estimation, on top of the integrated data stream to produce forecasts and recommendations. Finally, the simulation and interaction layer exposes the twin's state and predictions to human users through dashboards, virtual-reality interfaces, or, increasingly, conversational interfaces. As the DT paradigm has matured, twins have progressed from descriptive (showing current state), to predictive (forecasting future state), to prescriptive (recommending or taking corrective action), a progression driven almost entirely by the sophistication of the AI layer rather than the underlying sensing infrastructure [1], [2].

The modeling techniques used within the AI layer vary by task but recur across domains: supervised learning on historical sensor and outcome data supports failure and trajectory forecasting; unsupervised and semi-supervised anomaly detection flags deviations from normal operating patterns without requiring labeled failure examples, which are often rare; reinforcement learning supports closed-loop control and intervention recommendation; and, most recently, large language models support natural-language explanation and reasoning over unstructured contextual data alongside the structured sensor stream. A twin's fidelity, how accurately it represents its physical counterpart, and its update latency, how quickly it reflects real-world change, are the two primary engineering metrics used to evaluate a digital twin implementation, and the two are frequently in tension, since higher-fidelity models are typically more computationally expensive to update in real time.

Smart Manufacturing

Manufacturing was the earliest and remains the most mature application domain for AI-driven digital twins, where they support predictive maintenance, process optimization, and quality control across product design, production, and prognostics and health management [2], [4]. A twin of a production line ingests continuous sensor data (vibration, temperature, current draw) from machinery and applies trained models to detect early signatures of impending failure, allowing maintenance to be scheduled before a breakdown occurs rather than after, an approach reported to cut unplanned downtime and maintenance cost substantially relative to reactive maintenance schedules [5]. Beyond individual machines, plant-level twins simulate entire production lines to identify bottlenecks and test layout or scheduling changes virtually before committing to a costly physical changeover, and advanced robotics applications extend this to human-robot collaborative work cells where a twin continuously models both the robot and the human operator to plan safe, efficient joint motion [1].

Digital-twin-driven simulation is also reshaping how manufacturers validate new product designs: rather than building successive physical prototypes, engineers iterate on a virtual twin of the product itself, running stress, thermal, and fatigue simulations informed by AI surrogate models that approximate expensive physics-based simulations at a fraction of the computational cost, substantially shortening design-to-production timelines while catching design flaws earlier, when they are cheaper to fix.

A closely related concept, the digital thread, tracks a product's data across its entire lifecycle, from design and simulation through manufacturing and in-service operation, and AI-driven manufacturing twins increasingly draw on this thread to close the loop between how a part was designed, how it was actually produced, and how it performs in the field. Computer-vision models integrated into manufacturing twins support automated visual quality inspection, flagging defective parts in real time on the production line, while physics-informed machine learning combines sensor data with known engineering constraints to produce failure forecasts that remain physically plausible even when trained on comparatively limited historical failure data.

Smart Cities

City-scale digital twins integrate data from traffic sensors, utility meters, building management systems, and public infrastructure into a unified virtual model used for planning and real-time operations [4], [6]. AI models layered on this integrated data support traffic-flow prediction and adaptive signal control, energy-demand forecasting for grid balancing, and disaster-response simulation, in which a city planner can model the spread of flooding or the effect of an evacuation route before an event occurs rather than only after. A large bibliometric analysis of smart-city digital twin research identifies smart vehicles, energy optimization, and infrastructure resilience as the dominant application clusters, with cloud computing and IoT integration identified as complementary enabling technologies rather than separate research threads [6]. Urban digital twins face a distinctive interoperability challenge relative to single-factory manufacturing twins, since a city-scale twin must integrate data owned by many independent agencies and private operators rather than a single organization.

Transportation is among the most mature smart-city sub-applications: AI-enhanced traffic twins combine live sensor and camera feeds with predictive models to adjust signal timing dynamically, reducing congestion relative to fixed-timing schedules, and are increasingly integrated with digital twins of individual connected and autonomous vehicles to co-simulate vehicle behavior and infrastructure response. Energy-grid twins similarly combine smart-meter data with weather and demand forecasts to balance load across a distribution network in near real time, an application whose importance is expected to grow as distributed renewable generation introduces more variability into the grid than centralized generation historically did.

Healthcare

In healthcare, digital twins increasingly model individual patients rather than entire hospital systems, digitally replicating a patient's physiology to support personalized, proactive treatment [4], [7]. Surgeons rehearse complex procedures on a patient-specific virtual replica before operating, reducing intraoperative risk, and postoperative twins support continuous monitoring and assessment of how a patient is responding to an intervention [4]. Pharmaceutical applications simulate how a candidate compound might behave across large populations of virtual patients, shortening early-stage drug-testing timelines. A distinct and rapidly growing line of work uses large language models directly within the patient-twin pipeline: DT-GPT forecasts a patient's clinical-variable trajectory from electronic health record data using an LLM, additionally offering zero-shot prediction of variables not seen during training and a chatbot interface for clinicians to query the model's reasoning in natural language [8]. As in smart manufacturing, healthcare digital twins face domain-specific obstacles beyond data integration and interoperability, most importantly regulatory approval and the clinical validation required before any twin-derived recommendation can influence real patient care [4].

Beyond individual patient twins, hospital-operations twins model entire care facilities, patient flow through emergency departments, bed occupancy, staff scheduling, to identify bottlenecks and simulate the effect of policy changes before they are implemented, extending the same descriptive-to-prescriptive progression seen in manufacturing to hospital administration. Population-level twins, which aggregate de-identified data across many patients to model disease progression or treatment response at a cohort level, are increasingly used in clinical research and drug development to complement, rather than replace, traditional clinical trials, though their outputs are consistently treated in the literature as hypothesis-generating rather than a substitute for regulatory-grade evidence.

Education

Digital twins in education gained particular momentum during the disruption of in-person instruction caused by the COVID-19 pandemic, as virtual replicas of classrooms, laboratories, and even individual learners offered a way to preserve continuity and quality of instruction [9]. A learner-facing digital twin integrates a student's academic history, real-time engagement signals, and assessment results to provide personalized, adaptive learning pathways, with recent systems using edge or fog computing to process this data locally and reduce feedback latency, paired with a locally deployed LLM to interpret multimodal learner data and generate context-aware alerts for both students and instructors [10]. Beyond individual learners, discipline-specific virtual laboratories let students practice on a digital twin of physical equipment that would otherwise be scarce, expensive, or hazardous to access directly, extending hands-on training to a much larger population of learners than physical equipment alone could support [9].

A distinguishing feature of educational digital twins relative to the other domains surveyed is their explicit focus on the human learner as the object being modeled rather than a physical machine or environment: a learner twin must represent not only observable behavior, such as time spent on a task or quiz performance, but also latent state, such as engagement or comprehension, that can only be inferred indirectly, making the underlying modeling problem closer to that of patient digital twins in healthcare than to equipment twins in manufacturing.

The Emerging Role of Large Language Models

Across all four domains, a consistent recent trend is the addition of an LLM-based cognitive layer on top of the traditional sensing-and-modeling digital twin stack. This serves two related purposes. First, LLMs act as a natural-language interface, letting a non-expert user ask a twin plain questions, such as why a machine is predicted to fail or what a patient's forecasted trajectory implies, and receive an explanation grounded in the twin's underlying model outputs rather than raw numerical output alone [3], [8]. Second, LLMs are used to

enhance the modeling layer itself: because they can reason over heterogeneous, partially unstructured data (maintenance logs, clinical notes, operator reports) alongside structured sensor streams, LLM-enhanced twins can incorporate a wider range of real-world context than purely numerical models, and several surveys report LLMs contributing to geometric, mechanistic, and rule-based aspects of twin modeling in addition to their conversational role [7]. This convergence is sufficiently active that dedicated surveys of LLM-driven digital twins have begun to appear in domains ranging from telecommunications network optimization to cognitive digital twins for industrial systems [7].
Table I summarizes the primary AI role and representative benefit reported for each of the four domains surveyed.

Table 1: AI ROLE AND REPORTED BENEFIT BY DOMAIN

Domain	Primary AI Role	Reported Benefit
Manufacturing	Predictive maintenance, process optimization	Reduced unplanned downtime [5]
Smart Cities	Traffic/energy forecasting, disaster simulation	Improved infrastructure resilience [6]
Healthcare	Trajectory forecasting, surgical rehearsal	Personalized, proactive care [8]
Education	Adaptive learning, virtual labs	Personalized, continuous instruction [9]

Cross-Domain Challenges

Several challenges recur across all four domains. Interoperability remains a central obstacle: twins are built on heterogeneous platforms and data formats, and there is still no universal standard for how a twin should be constructed or how twins built by different organizations should exchange data, a problem that is especially acute for smart-city twins spanning many independent agencies [6]. Real-time synchronization between the physical asset and its virtual counterpart is computationally demanding at scale, and latency between a physical event and its reflection in the twin directly limits how useful the twin is for time-critical decisions such as disaster response or patient monitoring. Data privacy and security are heightened concerns given that digital twins aggregate large volumes of sensitive data, patient records, city infrastructure blueprints, proprietary manufacturing processes, into a single, continuously updated system, making them an attractive target and requiring robust access control and, in federated deployments, privacy-preserving training methods [9], [4]. Finally, validating the trustworthiness of AI-generated predictions and recommendations before acting on them remains an open problem, particularly in regulated domains such as healthcare where a twin's forecast must be clinically validated before it can influence patient care [4].
A further, more subtle challenge is model drift: a twin's underlying AI models are trained on historical data, but the physical system they represent, whether a machine, a city, or a patient, continues to change, and

ISSN: 2581-3404 (Online) **IF: 5.445 (SJIF)**

International Journal of Innovative Research in Technology and Management, **Volume-6, Issue-5, August-2022.**

without continuous retraining or monitoring for distributional shift, a twin's predictions can silently degrade in accuracy over time. This is compounded by the fact that ground-truth labels for evaluating twin predictions, such as an actual equipment failure or an actual patient outcome, often arrive with substantial delay relative to when the prediction was made, making rapid detection of model degradation difficult in practice.

Future Research Directions

Several directions are likely to shape the next generation of AI-powered digital twins. Standardized, cross-platform data models would materially reduce the interoperability burden that currently limits twin integration across organizational boundaries, particularly for smart-city applications. Federated and privacy-preserving training approaches, in which twin models are improved using data that never leaves its originating organization, are increasingly explored as a way to reconcile the need for large, diverse training data with strict data-privacy requirements in healthcare and other regulated sectors. Continued integration of multimodal generative AI, combining structured sensor data with text, images, and video, is expected to let twins reason over a richer picture of the systems they represent, an integration already being explored under frameworks such as the Sense-Map-Generate-Act loop proposed for AI-of-Things systems [3]. Finally, as LLM-based conversational interfaces become standard, research into ensuring these explanations are faithful to the twin's underlying model, rather than plausible-sounding post-hoc narratives, will be essential to maintaining trust in twin-derived recommendations.

There is also growing interest in twin-to-twin communication, in which digital twins of interacting physical systems, for example a vehicle twin and a road-infrastructure twin, exchange information directly to coordinate behavior without a human or a central controller in the loop, extending the digital twin concept from a single organization's asset toward a broader network of interoperating twins spanning multiple stakeholders. Realizing this vision at scale will depend heavily on the interoperability standards discussed above, since twin-to-twin communication is only possible if independently built twins can agree on a shared data representation. Benchmark suites that jointly measure prediction accuracy, update latency, and computational cost across twin implementations would additionally make it easier for practitioners to compare competing platforms on a like-for-like basis, a comparison that is currently difficult given how differently individual vendors and research groups report performance.

There is also growing interest in twin-to-twin communication, in which digital twins of interacting physical systems, for example a vehicle twin and a road-infrastructure twin, exchange information directly to coordinate behavior without requiring a human or a central controller in the loop, extending the digital twin concept from a single organization's asset toward a broader network of interoperating twins spanning multiple stakeholders. Realizing this vision at scale will depend heavily on the interoperability standards discussed above, since twin-to-twin communication is only possible if independently built twins can agree on a shared data representation.

Conclusion

AI has transformed digital twins from largely descriptive monitoring tools into predictive and increasingly prescriptive systems across manufacturing, smart cities, healthcare, and education. Despite substantial differences in the physical systems being modeled, all four domains share a common architecture, physical sensing, data integration, AI-driven modeling, and simulation and interaction, and all four are converging on a similar recent trend: layering large language models atop the traditional twin stack to make a twin's state and reasoning accessible in natural language to non-expert users. Realizing the full potential of this convergence will require progress on interoperability, real-time synchronization, data privacy, and the trustworthiness of AI-generated explanations, but the trajectory across all surveyed domains points toward digital twins that are increasingly autonomous, communicative, and integrated into everyday decision-making.

References

[1] Ziqi Huang, Yang Shen, "A Survey on AI-Driven Digital Twins in Industry 4.0: Smart Manufacturing and Advanced Robotics," Sensors, vol. 21, no. 19, art. 6340, 2021.
[2] "Developments of digital twin technologies in industrial, smart city and healthcare sectors: a survey," 2021.
[3] "Toward Intelligent AIoT: A Comprehensive Survey on Digital Twin and Multimodal Generative AI Integration," Mathematics, vol. 13, no. 21, art. 3382, 2021.
[4] "Industry 4.0: survey of digital twin in smart manufacturing and smart cities," ScienceDirect book chapter, 2022.
[5] Tao, F.; Cheng, J.; Qi, Q.; Zhang, M.; Zhang, H.; Sui, F. Digital twin-driven product design, manufacturing and service with big data. Int. J. Adv. Manuf. Technol. 2018, 94, 3563–3576.
[6] Scharl, S.; Praktiknjo, A. The Role of a Digital Industry 4.0 in a Renewable Energy System. Int. J. Energy Res. 2019, 43, 3891–3904.
[7] Wan, L.; Nochta, T.; Schooling, J.M. Developing a city-level digital twin—Propositions and a case study. In International Conference on Smart Infrastructure and Construction 2019 (ICSIC) Driving Data-Informed Decision-Making; ICE Publishing: London, UK, 2019; pp. 187–193.
[8] Shirowzhan, S.; Tan, W.; Sepasgozar, S.M. Digital twin and CyberGIS for improving connectivity and measuring the impact of infrastructure construction planning in smart cities. ISPRS Int. J. Geo-Inf. 2020, 9, 240.
[9] Deepak Rathore, Praveen Kumar, “Recent Trends in Machine Learning for Health Care Sector”, International Journal of Innovative Research in Technology and Management, Vol-5, Issue-2, 2021.
[10] Lokuwaduge, C.S.D.S.; Heenetigala, K. Integrating Environmental, Social and Governance (ESG) Disclosure for a Sustainable Development: An Australian Study. Bus. Strategy Environ. 2017, 26, 438–450.
[11]Tao, F.; Zhang, H.; Liu, A.; Nee, A.Y. Digital Twin in Industry: State-of-the-Art. IEEE Trans. Ind. Inform. 2019, 15, 2405–2415.