2026 ASEE ANNUAL CONFERENCE June 21–24, 2026 | Charlotte Convention Center, NC
Host Sponsor UNC Charlotte
Paper ID #50945

Narrative-Driven Digital Twin Framework with Generative AI for Enhancing Collaborative Robotics Education

Luis Enrique Ramos Dominguez
Brenda Jamileth Galdamez Rivera
Dr. Zhou Zhang, State University of New York, College of Technology at Farmingdale

I am an Assistant Professor at SUNY Farmingdale State College. My teaching and research interests include robotics and virtual reality in engineering education. I have a Ph.D. and a bachelor's degree in Mechanical Engineering, and my master's degree is in Electrical Engineering. I have over seven years of industrial experience as an electrical and mechanical engineer. I also have extensive teaching and research experience with respect to various interdisciplinary topics involving Mechanical Engineering, Electrical Engineering, and Computer Science.

Narrative-Driven Digital Twin Framework with Generative AI for Enhancing Collaborative Robotics Education

Abstract

The growing adoption of collaborative robots (cobots) in modern industry calls for educational approaches that emphasize intuitive reasoning about robot behavior, safety, and feasibility rather than early dependence on low-level programming. This paper presents a narrative-driven digital twin framework designed to support intent-centric learning in collaborative robotics education. The system integrates Generative AI with an immersive simulation developed in Unreal Engine (a real-time 3D graphics platform used to visualize robot motion, workspace interaction, and collision conditions) allowing students to specify robotic tasks using natural language. Narrative inputs are translated into structured symbolic task representations grounded in robot kinematics and constraint models before execution, ensuring transparency, feasibility, and safety-aware learning. An exploratory pilot study involving two undergraduate students examined interaction feasibility and learning-related behaviors through open-ended task activities, with system logs, observations, and reflective feedback analyzed descriptively. Results indicate that 85–87% of narrative task descriptions were successfully translated into executable motion plans, with iterative refinement driven by constraint feedback. Observed interaction patterns suggest emerging spatial reasoning, constraint awareness, and sustained exploratory engagement despite no prior robot programming experience. Beyond demonstrating system feasibility, this work contributes a structured integration of narrative interaction, constraint-grounded symbolic mediation, and immersive digital twins as an educational robotics paradigm. The exploratory findings establish a preliminary empirical basis for future scaled evaluation of conceptual development, transfer, and learner agency in human-centered automation education.

Keywords: Collaborative Robot, Generative AI, Digital Twins, Simulation-based Learning

1. Introduction

The rapid adoption of collaborative robots (cobots) in advanced manufacturing and service-oriented industries has fundamentally reshaped expectations for engineering graduates[1,2]. Unlike traditional industrial robots that operate in fenced and isolated environments, cobots are designed to work in close physical proximity to humans[3]. This shift introduces new technical and pedagogical demands, requiring engineers to reason simultaneously about kinematics, workspace constraints, safety, and human-centered interaction. As a result, robotics education must evolve beyond code-centric instruction toward learning experiences that emphasize intuitive understanding, safe experimentation, and systems-level reasoning.

Despite their growing industrial relevance, cobots remain challenging to integrate effectively into undergraduate curricula[4]. Conventional robotics education often relies on low-level programming languages, vendor-specific interfaces, and mathematically intensive formulations that present steep learning curves for novice learners[5]. While these approaches are essential for advanced robotics training, they frequently delay meaningful engagement with real robot behavior. Students may spend disproportionate time mastering syntax and tooling, rather than developing intuition about robot motion, reachability, safety constraints, and human–robot interaction. Practical constraints further limit hands-on learning. Industrial cobots are costly, laboratory access is restricted, and institutional safety policies often reduce student autonomy during early-stage instruction. Although cobots are inherently safer than traditional manipulators, improper task specification or misunderstanding of motion constraints can still pose risks. Consequently, introductory laboratory experiences are often highly scripted, limiting experimentation, iteration, and exploratory learning. These limitations hinder the development of spatial reasoning, confidence, and conceptual understanding.

Digital twin technology offers a promising pathway to address these challenges by enabling high-fidelity virtual representations of physical robotic systems[6]. Digital twins provide a risk-free environment for visualizing robotic behavior. Game-engine-based platforms, such as Unreal Engine, further enhance these experiences through realistic physics and immersive visualization. However, many existing digital twin solutions remain programming-intensive, limiting their accessibility for early-stage learners. At the same time, advances in Generative Artificial Intelligence—particularly large language models (LLMs)—have introduced new opportunities for human–machine interaction[7]. Natural language interfaces allow users to express task intent at a conceptual level, abstracting away low-level syntax while preserving logical structure. In an educational context, this capability enables a shift from a code-first paradigm toward intent-centric learning, where students focus on understanding system behavior, feasibility, and safety implications. Motivated by these developments, this paper presents a narrative-driven digital twin framework that integrates Generative AI with an Unreal Engine–based simulation of a Universal Robots UR3e collaborative robot[8,9]. Students interact with the digital twin using narrative-style natural language instructions, which are interpreted by a large language model into structured task representations grounded in formal kinematic and constraint models. Robot behavior is then simulated within an immersive environment, allowing students to visualize motion, evaluate safety constraints, and iteratively refine task strategies without requiring physical hardware.

The proposed framework is designed to complement, rather than replace, traditional robotics instruction. By lowering barriers to entry, it enables earlier and more meaningful engagement with collaborative robotics concepts while preserving opportunities for deeper technical exploration. Through the integration of immersive digital twins and constrained AI-assisted interaction, this work advances a human-centric, accessible, and safety-aware approach to collaborative robotics education aligned with Industry 5.0 principles[10].

The instructional perspective of this work is informed by constructivist and experiential learning principles, which emphasize knowledge development through interaction, reflection, and iterative refinement of understanding. The narrative-driven digital twin environment is designed as a cognitive scaffold that allows learners to formulate intent, observe consequences, and revise strategies in response to constraint-based feedback. This aligns with inquiry-based learning models in which conceptual understanding emerges through exploration rather than procedural instruction. Generative AI interaction serves as a mediation layer that lowers syntactic barriers while preserving agency and critical reasoning, supporting early engagement without replacing analytical thought. This conceptual framing connects the system design, evaluation methodology, and interpretation of learning behaviors investigated through the research questions that follow.

To guide evaluation of the proposed framework's educational value, this study is structured around the following research questions:

RQ1. To what extent can a narrative-driven digital twin framework enable novice learners to formulate, execute, and iteratively refine collaborative robot tasks without prior experience in low-level robot programming?

RQ2. What forms of spatial reasoning, constraint awareness, and task-planning behaviors emerge when learners interact with a constraint-aware digital twin grounded in formal robot kinematics?

RQ3. How does natural-language interaction mediated by Generative AI influence learner engagement, confidence, and exploratory experimentation in early-stage collaborative robotics education?

RQ4 (Exploratory). To what extent do learners demonstrate progression from high-level narrative task specification toward informed reasoning about robot feasibility, reachability, and safety constraints?

These questions frame the pilot study as an exploratory investigation of educational feasibility and learning processes rather than a statistically powered evaluation of learning gains.

To clarify the instructional intent of the framework, the learning experience was designed to support the following objectives:

1. Formulate collaborative robot tasks at a conceptual level using natural-language intent description
2. Interpret feasibility feedback related to reachability, constraints, and safety
3. Develop spatial reasoning regarding robot motion and workspace limits
4. Understand safety-aware task planning in human–robot interaction contexts
5. Reflect critically on AI-assisted task interpretation and decision support

These objectives guided activity design and informed interpretation of observed learning behaviors in the pilot evaluation.2. Related Work

2.1 Digital Twins in Robotics and Engineering Education

Digital twin technology has gained significant attention in engineering education as a means of bridging the gap between theoretical instruction and hands-on experimentation[11,12]. A digital twin typically consists of a high-fidelity virtual representation of a physical system that mirrors its structure, behavior, and operational constraints[13,14]. In robotics and mechatronics education, digital twins have been widely adopted to support visualization of kinematics, simulation of control strategies, and exploration of system dynamics without the cost, safety risks, or scheduling constraints associated with physical hardware[15,16,17].

Prior studies have shown that digital twins can enhance conceptual understanding, particularly in areas such as spatial reasoning, coordinate transformations, and motion planning[18,19]. By allowing students to observe cause–effect relationships in real time, digital twins support active learning and iterative experimentation, which are difficult to achieve in traditional laboratory settings[20]. Game-engine-based platforms, including Unreal Engine, Valve Engine and Unity, have further improved the educational value of digital twins by offering realistic physics, immersive visualization, and interactive user interfaces[21,22,23].

Despite these advantages, many educational digital twin implementations remain tightly coupled to traditional programming workflows. Students are often required to write scripts or control logic using vendor-specific languages or robotics middleware, which can limit accessibility for novice learners. As a result, while digital twins reduce hardware-related barriers, they do not fully address the cognitive challenges associated with early-stage robotics education. This limitation motivates the exploration of alternative interaction paradigms that emphasize intent and conceptual reasoning over syntax.

2.2 Generative AI and Natural Language Interfaces in Engineering Education

Recent advances in Generative Artificial Intelligence, particularly large language models (LLMs), have introduced new opportunities for natural language interaction with complex systems[24]. In engineering education, LLMs have been explored as intelligent tutors, automated feedback providers, code generation assistants, and conceptual explainers[25,26]. These tools have demonstrated potential to reduce cognitive load, support self-paced learning, and improve student engagement by allowing learners to interact with technical content in more intuitive ways.

Natural language interfaces are especially relevant in domains characterized by complex syntax and steep learning curves, such as robotics programming. By enabling students to express intent using narrative or task-oriented language, LLMs can abstract low-level implementation details while preserving logical structure. This abstraction aligns with pedagogical goals in early-stage learning, where developing intuition and system-level understanding is often more critical than mastering programming syntax.

However, existing applications of Generative AI in engineering education are largely text-based and disconnected from physical or simulated system behavior. While LLMs have been used to generate code or explain algorithms, their integration with interactive robotics simulations remains limited. Few studies have examined how narrative-driven AI interfaces can be coupled with immersive digital twins to support experiential learning, particularly in the context of collaborative robotics

2.3 Explainable AI in Robotics and Safety-Critical Education

As AI-driven decision-making becomes increasingly embedded in robotics and automation systems, the need for transparency and interpretability has become a central concern, especially in safety-critical and educational contexts[27]. Explainable AI (XAI) seeks to make model decisions understandable to human users, supporting trust, accountability, and learning[28]. Techniques such as SHAP, LIME, and counterfactual explanations have been widely studied in machine learning research and have begun to appear in educational applications[29].

In robotics, XAI has been explored primarily to support operator trust, fault diagnosis, and safety validation[30]. From an educational standpoint, XAI offers additional pedagogical value by enabling students to understand why a system behaves in a particular way, rather than merely observing outcomes[31]. This transparency is particularly important in collaborative robotics, where safety constraints, motion planning decisions, and human–robot interaction policies must be clearly understood.

Despite growing interest in XAI, its adoption in educational robotics platforms remains limited. Most XAI studies focus on static datasets or offline model interpretation, rather than real-time interaction within simulated or cyber-physical environments. Moreover, few educational systems integrate XAI directly into digital twin frameworks in a way that supports student inquiry, reflection, and iterative learning.

2.4 Research Gap

Collectively, prior work demonstrates the individual benefits of digital twins, Generative AI, and Explainable AI in engineering education. Digital twins enhance visualization and experimentation, Generative AI lowers interaction barriers through natural language, and XAI promotes transparency and trust in AI-driven systems. However, these technologies are typically explored in isolation, and their integration into a unified educational framework for collaborative robotics remains underdeveloped.

In particular, there is a lack of educational platforms that combine immersive digital twins with narrative-driven AI interaction to support safe, intuitive, and scalable robotics training. The framework proposed in this paper addresses this gap by unifying digital twins, Generative AI, and XAI within a cohesive, education-centered architecture tailored for collaborative robotics. This integration supports early engagement, conceptual understanding, and future expansion toward cyber-physical and safety-aware learning environments.

While prior studies explore these technologies independently, no reported educational robotics platform integrates immersive digital twin simulation, narrative task specification, symbolic constraint validation, and explainability-oriented feedback into a unified instructional workflow designed specifically for novice collaborative robotics learning. This absence motivates the present work, which contributes both an integrated architecture and exploratory educational evaluation bridging interaction design, robotics modeling, and pedagogical grounding.

3. System Architecture

This work introduces a narrative-driven digital twin architecture that reframes collaborative robotics education around intent rather than code. The core idea is to allow students to express robotic tasks using natural language while ensuring that those intentions are transparently grounded in formal kinematic models, safety constraints, and executable motion plans. In doing so, the architecture lowers barriers to entry without obscuring the physical and mathematical foundations of robot behavior.

The proposed architecture is distinguished by four integrated design principles. First, narrative task specification enables intuitive interaction while remaining physically grounded in robot kinematics and constraints. Second, an intermediate symbolic task representation separates large language model (LLM) interpretation from robot execution, providing a critical layer for validation, explainability, and safety. Third, a constraint-aware digital twin enforces joint limits, collision avoidance, and manipulability considerations within an immersive Unreal Engine environment, allowing students to safely explore both feasible and infeasible behaviors. Finally, the architecture is modular and extensible, supporting future integration with real-time sensing, cyber-physical systems, and advanced learning-based control methods.

Figure 1 conceptually positions the proposed framework within the broader landscape of robotics education, illustrating how narrative interaction, digital twins, and Generative AI–assisted interfaces are integrated into a unified, iterative learning paradigm. Figure 2 then details the operational architecture that enables this integration.

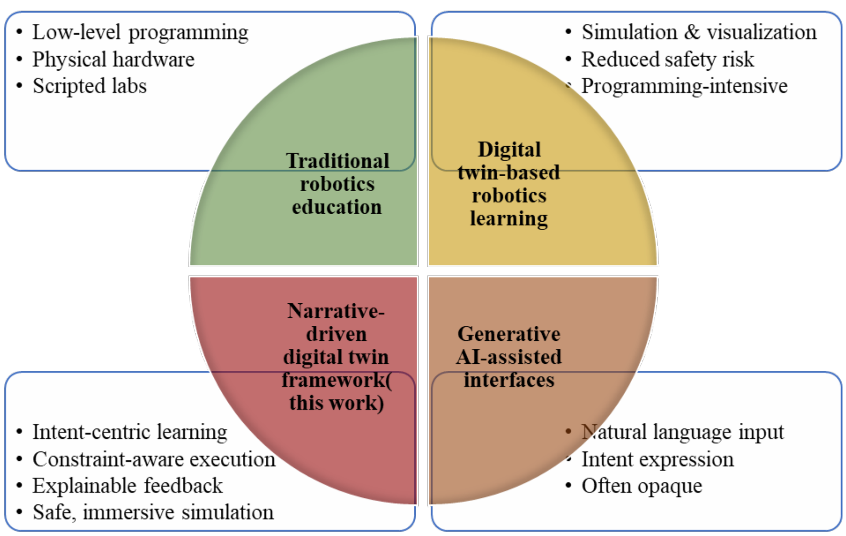


Figure 1: Positioning of Narrative-Driven Digital Twins in Robotics Education.

3.1 Design Philosophy and Theoretical Foundation

The system is designed around a hierarchical intent-to-motion abstraction, in which high-level human intent is progressively transformed into robot motion through increasingly formal representations. This design aligns with classical robot control architectures and human–robot interaction (HRI) frameworks that decouple task specification, motion generation, and execution to improve robustness, safety, and interpretability. From an educational perspective, this hierarchy allows students to engage first with conceptual task formulation before encountering lower-level kinematic and control details. Formally, the system follows the abstraction pipeline:

Natural Language Intent → *Symbolic Task Representation* → *Motion Planning* →*Execution and Feedback*

Rather than relying on Generative AI to directly produce robot control commands, the framework introduces an intermediate symbolic task representation that is explicitly grounded in robot kinematics and safety constraints. This design choice improves transparency, reduces the risk of unvalidated actions, and preserves student agency by separating semantic interpretation from physical execution.

Figure 2 illustrates this hierarchical architecture, showing how semantic interpretation, symbolic task grounding, constraint validation, and motion execution are explicitly separated to ensure that narrative input is translated into physically feasible and safe robot behavior.

From an instructional standpoint, this hierarchical abstraction directly operationalizes constructivist and experiential learning principles introduced in Section 1. By allowing learners to iteratively formulate intent, observe system response, and revise strategies, the architecture transforms theoretical pedagogical positioning into actionable interaction design. The symbolic mediation layer further supports reflective reasoning by exposing constraint-grounded interpretations rather than opaque automation, reinforcing learner agency and conceptual transparency.

3.2 Robot Modeling and Mathematical Foundations

These models are not presented for computational completeness, but to explicitly ground the narrative-driven interaction in formal robotics theory

3.2.1 Kinematic Modeling

The collaborative robot is modeled as a six-degree-of-freedom serial manipulator with joint configuration:

$$q = [q_1, q_2, q_3, q_4, q_5, q_6]^T \tag{1}$$

where $q_i \in R$ denotes the angular position of the $\boldsymbol{i}_{th}$ revolute joint.

The end-effector pose is expressed as a homogeneous transformation matrix $T(q) \in SE(3)$, which encodes both position and orientation. Using standard Denavit–Hartenberg (DH) conventions, the forward kinematics are computed as:

$$T(q) = \prod_{i=1}^{6} T_{i-1}^{i}(q_i) \tag{2}$$

Where $T_{i-1}^{i}(q_i)$ denotes the homogeneous transformation from link $i-1$ to link i, parameterized by joint variable q_i.

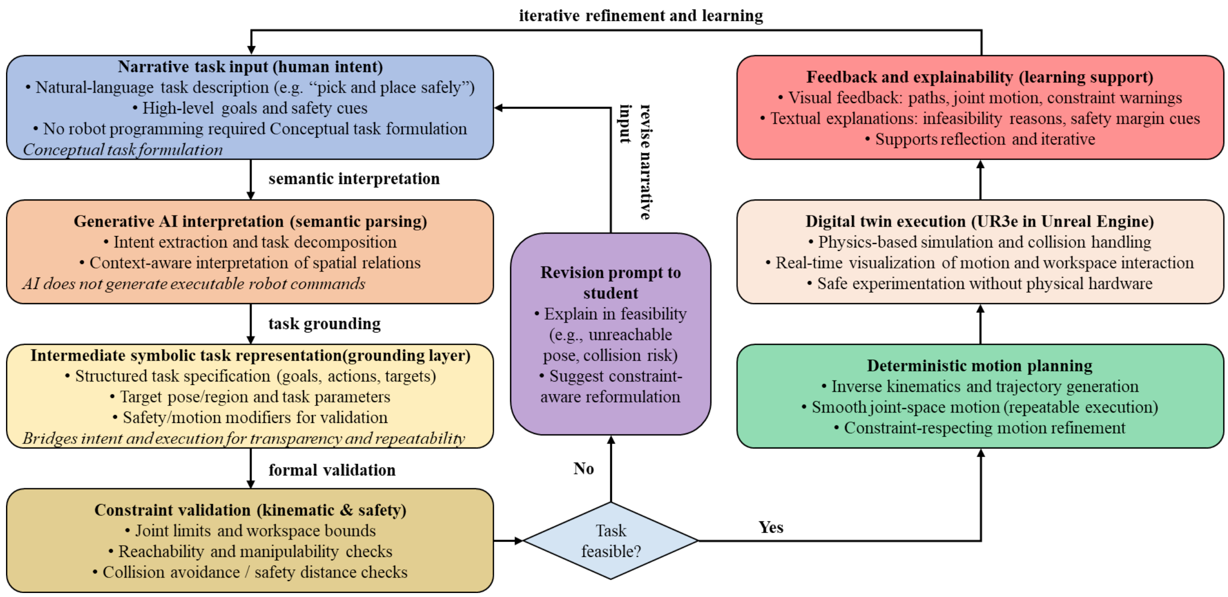


Figure 2: Architecture flowchart of the narrative-driven digital twin framework.

3.2.2 Differential Kinematics and Inverse Kinematics

The differential relationship between joint velocities and end-effector motion is described using the Jacobian matrix $J(q) \in R^{6\times 6}$:

$$v = \begin{bmatrix} \dot{p} \\ \omega \end{bmatrix} = J(q)\dot{q} \tag{3}$$

Where v is the end-effector twist; $\dot{p} \in R^{3}$ is the linear velocity of the end-effector; $\omega \in R^{3}$ is the angular velocity of the end-effector; and $\dot{q} \in R^{6}$ is the joint velocity vector. To compute joint velocities for a desired end-effector twist v_d, inverse kinematics are solved using a damped least-squares (DLS) formulation to improve numerical stability near singular configurations:

$$\dot{q} = J^{T}(q)(J(q) + \lambda^{2} I)^{-1} v_d \tag{4}$$

Where $\lambda > 0$ is a damping coefficient; I is the identity matrix; and $v_d \in R^{6}$ is the desired end-effector twist. This formulation allows smooth and robust motion generation within the digital twin environment while exposing students to core robotics principles.

3.2.3 Manipulability Analysis

To provide students with insight into robot dexterity and safety, a manipulability measure is computed using Yoshikawa's metric:

$$w(q) = \sqrt{\det\left(J(q)J^{T}(q)\right)} \tag{5}$$

where $w(q)$ is a scalar value indicating the robot's ability to generate motion in arbitrary directions from configuration q. Low manipulability values indicate proximity to kinematic singularities and are used to generate warnings within the digital twin environment.

3.2.4 Dynamic Model

For completeness and future cyber-physical integration, the robot dynamics are expressed as:

$$M(q)\ddot{q} + C(q, \dot{q}) + g(q) = \tau + \tau_{ext} \tag{6}$$

Where $M(q)$ is the joint-space inertia matrix; $C(q, \dot{q})$ represents Coriolis and centrifugal effects; g(q) is the gravity vector; τ denotes commanded joint torques; and τ_{ext} represents external interaction torques.

Although the mathematical formulations presented in Section 3.2 serve implementation purposes, their inclusion is pedagogically intentional. Explicit grounding of motion execution in forward and inverse kinematics, manipulability analysis, and constraint modeling ensures that narrative interaction remains physically interpretable rather than purely symbolic. Observable feedback derived from these models—such as reachability limits, singularity warnings, and collision violations—directly shapes learner reasoning during interaction. As learners revise task descriptions in response to these constraints, the mathematical structure becomes an implicit instructional mechanism supporting spatial reasoning, feasibility awareness, and safety-conscious planning, thereby connecting system-level modeling with educational outcomes investigated in RQ2. From an instructional perspective, inclusion of these formulations is not intended to present algorithmic novelty but to make explicit the physical grounding of narrative interaction. Their visibility enables learners to connect qualitative task intent with quantitative robot behavior, supporting conceptual bridging between experiential interaction and formal robotics theory.

3.3 Layered System Architecture

The proposed framework is implemented as a four-layer modular architecture, with each layer responsible for a specific stage in the intent-to-motion pipeline.

Narrative Interaction Layer

The Narrative Interaction Layer enables students to specify robot tasks using natural language descriptions. These descriptions encode task objectives, spatial relationships, and safety considerations, encouraging high-level reasoning rather than syntax-driven programming.

Generative AI Interpretation Layer

This layer employs a GPT-based large language model to map natural language input L_{nat} into a structured symbolic task representation S:

$$F_{LLM}: L_{nat} \rightarrow S \tag{7}$$

The symbolic task representation is defined as a finite-state task model:

$$S = (Z, A, \delta, z_0, G) \tag{8}$$

Where Z is the set of task states; A is the set of available actions; δ is the state transition function; z_0 is the initial state; and G is the set of goal states. This intermediate representation enables validation, explainability, and safe grounding prior to execution.

Digital Twin Simulation Layer

The Digital Twin Simulation Layer is implemented using Unreal Engine and hosts a high-fidelity model of the UR3e robot. Joint trajectories are generated using quintic polynomial interpolation:

$$q_i(t) = a_0 + a_1 t + a_2 t^2 + a_3 t^3 + a_4 t^4 + a_5 t^5 \tag{9}$$

where the coefficients a_k are determined from boundary conditions on joint position, velocity, and acceleration. Safety constraints are enforced as: $q_{min} \leq q(t) \leq q_{max}$, and collision avoidance is modeled using a signed distance function: $d(q(t)) \geq d_{safe}$ with d(q) is the minimum distance between robot geometry and obstacles, and d_{safe} is a predefined safety margin.

Feedback and Explainability Layer

This layer provides rule-grounded visual and textual feedback explaining constraint violations and task feasibility; future work will integrate algorithmic XAI methods to further enhance interpretability of AI-assisted task interpretation and planning.

3.4 Learner-Centered Perspective System Workflow

While Section 3.1 describes the system architecture from a formal and implementation-oriented perspective, this section reframes the same process from the learner's interaction and cognitive standpoint. From the learner's perspective, the system operates as an iterative intent-to-feedback learning loop. Tasks that violate kinematic or safety constraints are rejected with explanatory feedback, prompting revision. Feasible tasks are executed within the digital twin, allowing students to observe motion, trajectories, and workspace interaction in real time.

Crucially, the workflow is non-terminal. Whether a task succeeds or fails, students are encouraged to refine their narrative inputs and test alternatives. Through repeated interaction, learners develop intuition about robotic constraints, motion planning trade-offs, and safe human–robot interaction, supporting inquiry-based learning through exploration and reflection.

4. Implementation Details

This section describes the practical realization of the proposed narrative-driven digital twin framework, emphasizing implementation decisions, system integration, and reproducibility in an educational setting. While Section 3 focuses on architectural design and theoretical foundations, this section documents how those concepts were instantiated in a working prototype suitable for undergraduate robotics education.

The implementation prioritizes three objectives: educational clarity, ensuring system behavior is transparent and interpretable; operational robustness, enabling stable execution on standard laboratory computers; and repeatability, supporting consistent learning experiences across users and instructional contexts. Figure 3 illustrates the closed-loop learning paradigm supported by the implemented system, highlighting the iterative refinement of task intent, simulation behavior, and learner understanding.

4.1 Development Environment and Software Stack

The digital twin environment was developed using Unreal Engine, selected for its real-time rendering, physics simulation, and extensibility through C++ and Blueprint scripting. Unreal Engine enables precise control over joint hierarchies, collision handling, and simulation timing, which are essential for modeling collaborative robot behavior in an immersive yet controlled educational environment.

The Universal Robots UR3e model was imported as a hierarchical rigid-body assembly, with each revolute joint mapped to a controllable degree of freedom. Joint axes, offsets, and motion limits were defined according to publicly available UR3e specifications. To maintain real-time performance suitable for classroom use, simplified collision geometries were employed in place of high-resolution meshes. This design choice balances physical realism with computational efficiency and ensures smooth interaction on standard Windows-based laboratory computers.

All robot motion is executed at the kinematic level, consistent with the instructional scope of the current implementation. While the architecture supports dynamic and force-based control, omitting these elements at this stage improves numerical stability and reduces cognitive complexity for novice learners.

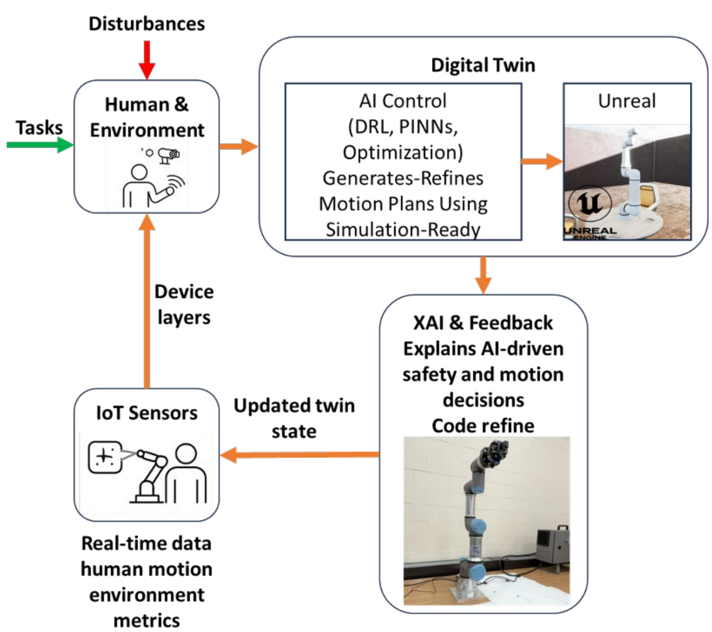


Figure 3: Closed-loop learning cycle illustrating how student input, constraint-aware simulation, and feedback iteratively refine task understanding and robot behavior.

4.2 Task Representation and Generative AI Integration

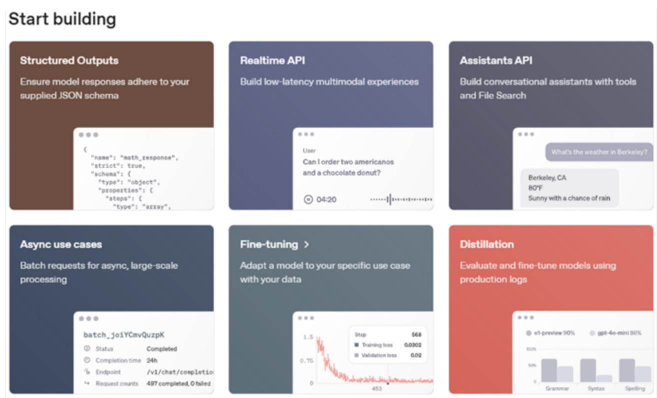
A central implementation decision was to strictly constrain the role of Generative AI. Consistent with the architectural constraints described in Section 3, the LLM is restricted to semantic interpretation rather than executable control generation. Each task is internally represented using a lightweight, deterministic data structure consisting of:

- A target end-effector pose or target region

- An action category (e.g., approach, move, align)
- Optional motion or safety modifiers (e.g., reduced speed, obstacle awareness)

This representation is human-readable and serializable, enabling tasks to be logged, replayed, and inspected by instructors. By decoupling semantic interpretation from execution, the system ensures that identical task specifications yield identical robot behavior, independent of narrative phrasing. This design substantially improves repeatability and fairness in educational use.

Prompt engineering is applied to encourage the LLM to generate consistent, constraint-aware outputs. However, all generated task specifications are validated before execution, ensuring that the LLM functions as an assistive semantic interface rather than an autonomous controller. Figure 4 illustrates the integration of the OpenAI API within the Unreal Engine plugin architecture.




```
FHttpRequestPtr Request = FHttpModule::Get().CreateRequest();
Request->SetURL("https://api.openai.com/v1/completions");
Request->SetVerb("POST");
Request->SetHeader("Authorization", "Bearer YOUR_API_KEY");
Request->SetContentAsString(PayloadJsonString);
Request->OnProcessRequestComplete().BindUObject(this,
&YourClass::OnResponseReceived);
Request->ProcessRequest();
```

Figure 4: Integration of the OpenAI API into the Unreal Engine–based plugin, enabling narrative task interpretation while preserving deterministic and constraint-aware execution.

4.3 Motion Generation and Constraint Enforcement

Once a task specification is validated, motion generation proceeds through a deterministic pipeline. Desired end-effector targets are converted into joint-space trajectories using numerical inverse kinematics implemented via a damped least-squares formulation. The damping coefficient is externally configurable, allowing instructors to demonstrate the effects of kinematic singularities and numerical conditioning without destabilizing the simulation.

Joint trajectories are generated using quintic polynomial interpolation, producing smooth motion with continuous position, velocity, and acceleration profiles. This formulation was selected for its robustness and pedagogical transparency, enabling students to directly relate trajectory smoothness to observable robot behavior.

Safety constraints are enforced at every simulation timestep, including joint limits, workspace boundaries, and collision avoidance conditions derived from signed distance checks. When a constraint violation is detected, execution is halted and a descriptive explanation is generated, reinforcing safety-first reasoning.

4.4 Digital Twin Execution and Visualization

Validated trajectories are executed within the Unreal Engine simulation environment. Joint values are updated in real time, and the resulting robot motion is rendered with synchronized kinematic and collision states. The digital twin provides several visualization layers designed to support learning without overwhelming the user, including joint limit indicators, end-effector trajectory traces, collision proximity highlighting, and workspace boundary visualization.

These visual elements are intentionally minimalistic and context-aware, appearing only when relevant to the current task or when constraints are violated. This design reduces cognitive load while preserving instructional value.

4.5 Feedback, Instructor Control, and Reproducibility

Feedback is provided through a combination of visual cues and concise textual messages derived directly from kinematic analysis and constraint evaluation. When a task is infeasible, the system reports the specific reason (e.g., unreachable pose, joint limit violation, collision risk), supporting transparent reasoning rather than opaque AI explanations.

To support instructional use and reproducibility, key system parameters—including joint limits, safety margins, inverse kinematics damping, and motion speed scaling—are externally configurable. All narrative inputs, task specifications, and execution outcomes can be logged, enabling instructors to review student progress, replay experiments, and compare performance across cohorts.

Importantly, the platform requires no specialized hardware. All experiments were conducted on standard laboratory computers, making the system accessible to institutions with limited resources.

While full algorithmic XAI techniques such as SHAP or LIME are not yet integrated, the current implementation delivers rule-grounded transparency through deterministic constraint reporting and symbolic task inspection. This serves as an instructional precursor to formal XAI deployment planned for subsequent system iterations.

4.6 Implementation Scope and Practical Limitations

The current implementation emphasizes educational robustness and interpretability rather than full physical fidelity. Real-time sensing, force interaction, and impedance control are not enforced in the present system, although the architecture supports their integration in future iterations.

This scoped approach ensures that students first develop a solid understanding of kinematics, constraints, task planning, and safety before progressing to advanced control strategies. Such staged exposure aligns with best practices in engineering education and supports incremental curriculum adoption.

5. Educational Evaluation and Pilot Study

5.1 Evaluation Purpose and Learning Context

This pilot study evaluates the educational feasibility and instructional value of the proposed narrative-driven digital twin framework for collaborative robotics education. Because the pilot involved a small number of participants, the goal was not to claim statistically significant learning gains, but to examine interaction feasibility, characterize learning-related behaviors, and gather exploratory evidence aligned with the research questions guiding this work.

Specifically, evaluation activities were structured to generate evidence addressing the following areas:

RQ1: Feasibility of narrative-driven task formulation and execution

RQ2: Emergence of spatial reasoning and constraint awareness

RQ3: Effects of natural-language interaction on engagement and experimentation

RQ4 (Exploratory): Indications of progression toward model-informed reasoning about feasibility and safety

The study was conducted during a summer undergraduate research experience with two engineering students with limited prior exposure to industrial collaborative robots and no formal training in vendor-specific robot programming languages. Students primarily interacted with the Unreal Engine–based UR3e digital twin described in Sections 3 and 4, with limited physical robot exposure reserved for late-stage comparison. This structure allowed unrestricted experimentation without early safety or access constraints. Figure 5 documents the pilot implementation and the correspondence between the simulation environment and the physical UR3e used for later validation.

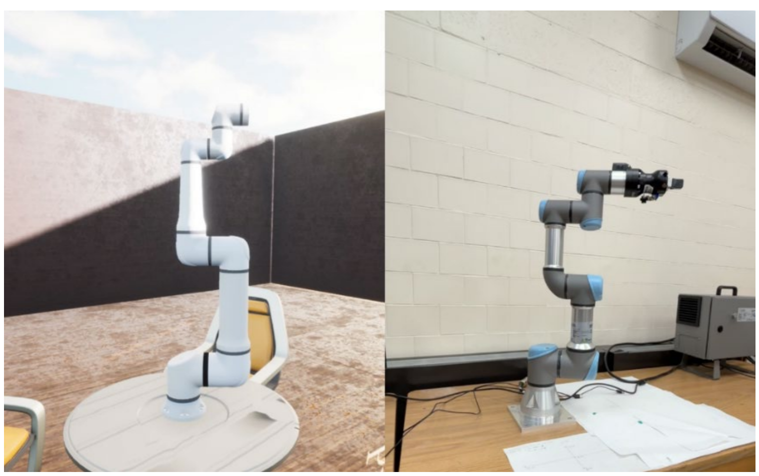


Figure 5: Pilot implementation of the narrative-driven digital twin framework, showing the Unreal Engine–based UR3e simulation (left) and the corresponding physical UR3e robot (right).

5.2 Learning Activities, Data Sources, and Collection Methods

Students completed a sequence of open-ended task activities (e.g., reaching, alignment, and safe motion near workspace boundaries) using narrative task descriptions. They were not given scripted procedures. Instead, they were encouraged to iteratively refine narrative inputs based on system feedback, consistent with the intent-to-feedback workflow in Section 3.4.

Each participant completed approximately 8–10 task interaction sessions lasting 30–45 minutes each over a two-week period. Tasks were not scripted but followed thematic categories (workspace reaching, obstacle-aware motion, and boundary-sensitive positioning) to ensure comparable exposure across participants while preserving exploratory autonomy. This structure supports reproducibility of instructional conditions while maintaining open-ended interaction design.

Multiple evidence sources were collected to evaluate learning processes and feasibility: (1) system interaction logs (narrative inputs, task interpretations, constraint checks, and execution outcomes), (2) researcher observation notes during student interaction and discussion, and (3) student feedback describing usability, perceived learning, and confidence. This multi-source design supports triangulation between quantitative interaction behavior and qualitative learning evidence.

5.3 Interaction Metrics and Exploratory Results

To complement qualitative observations, the system logged a set of interaction metrics that reflect learning-relevant behavior rather than performance on traditional examinations. These indicators quantify iterative refinement, constraint-awareness, and task efficiency.

Table 1 summarizes representative interaction metrics used in the pilot evaluation. The metrics are aligned with the framework's instructional intent: encouraging students to reason about feasibility and safety, revise task formulation, and reduce constraint violations over time through reflective iteration.

The pilot results indicate that 85–87% of narrative task descriptions were successfully translated into valid joint-space motion within the Unreal Engine simulation. Most unsuccessful cases were attributable to ambiguous task descriptions or target requests outside the robot's reachable workspace. Safety checks were frequently activated: approximately 35% of commands were flagged due to collision risks, unrealistic movements, or boundary violations. Importantly, these interventions did not halt learning; instead, they generated interpretable feedback that supported iterative refinement.

Table 1: Student interaction metrics and learning indicators.

Metric Category	Indicator	Student A	Student B	Interpretation
Task Success	Valid task execution rate (%)	87%	85%	Indicates effective grounding of narrative intent
Iteration Behavior	Avg. narrative revisions per task	4.0	4.4	Reflects iterative refinement toward feasibility
Constraint Awareness	Avg. constraint violations per task	2.0	2.4	Captures learning through controlled failure
Safety Reasoning	Tasks flagged by safety filters (%)	34%	36%	Demonstrates active safety screening and awareness
Efficiency	Avg. time to feasible execution (min)	12.0	13.4	Indicates improved efficiency in task formulation
Dominant Error Type	Most frequent infeasibility cause	Reachability	Reachability	Highlights spatial reasoning and workspace limits
Learning Outcome	Independent correction without hints	Yes	Yes	Suggests internalization of constraints

Figure 6 visualizes key interaction distributions. The histograms highlight trends in narrative revision behavior, safety-related constraint encounters, and task execution efficiency, providing insight into how learners adapt their task formulation over time. Consistent with the pilot scope of this study, these plots are descriptive in nature and are not intended to establish statistical significance.

Interaction trends demonstrate alignment with research objectives:

1. Evidence supporting RQ1 is reflected in successful task execution and iterative refinement behaviors
2. Indicators relevant to RQ2 appear in reachability-related errors and constraint awareness patterns
3. Observations supporting RQ3 emerge from sustained experimentation and narrative exploration
4. Behaviors associated with RQ4 include independent correction and reduced infeasibility over time

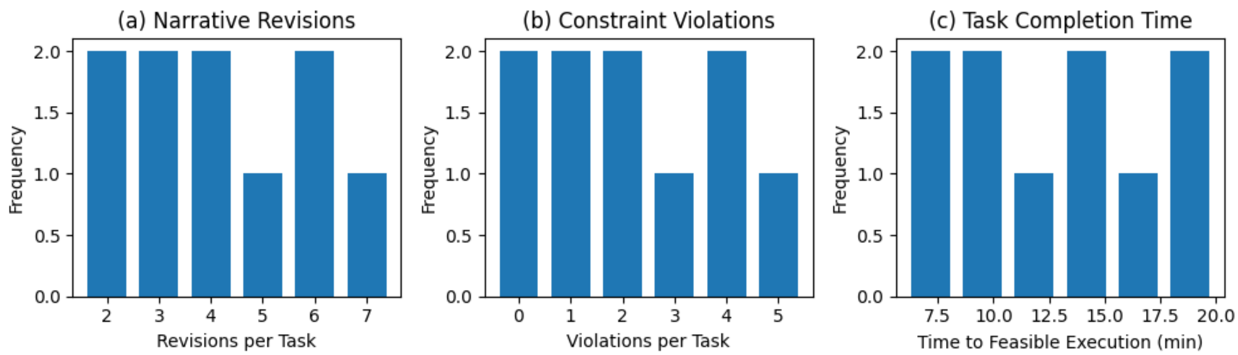


Figure 6: Exploratory interaction metrics from the pilot study: (a) Distribution of narrative revisions per task, (b) distribution of constraint violations per task, and (c) distribution of time required to achieve feasible task execution.

5.4 Observed Learning Processes and Educational Interpretation

Consistent with the workflow in Section 3.4, analysis of interaction cycles revealed several recurring learning processes consistent with the research questions.

First, students developed spatial reasoning and kinematic intuition through immersive visualization. Observing joint motion, end-effector trajectories, and workspace boundaries helped students connect narrative intent to physical robot behavior. Over time, students increasingly anticipated reachability issues and revised task descriptions proactively (supporting interpretations aligned with RQ2).

Second, narrative interaction reduced early cognitive barriers by shifting student attention from syntax and tooling toward task intent, feasibility, and safety. Students reported that the ability to describe goals in natural language improved confidence and encouraged experimentation, particularly in early stages when low-level programming often discourages novice learners (supporting interpretations aligned with RQ3).

Third, constraint-aware feedback supported learning through productive failure. When tasks were infeasible, explicit explanations (e.g., unreachable pose, joint limit violation, collision risk) prompted hypothesis testing and revision rather than abandonment. This behavior is consistent with inquiry-based and constructivist learning models, in which understanding is built through experimentation and reflection (contributing evidence toward RQ1 and exploratory indications associated with RQ4).

Across interaction sessions, behavioral patterns suggested movement beyond interface use toward reasoning informed by system constraints. While not sufficient to claim transfer, these observations indicate emerging internalization of feasibility considerations consistent with the exploratory scope of this study.

6. Discussion, Ethical Considerations, and Conclusions

From an evaluation standpoint, internal validity is supported through system logging and triangulation across observation and reflection data. Construct validity is strengthened by alignment between interaction metrics and research question constructs. External validity remains limited due to sample size and context specificity, reinforcing the exploratory scope of conclusions and motivating planned multi-cohort studies.

6.1 Discussion in Relation to Research Questions

Findings from the pilot study indicate that narrative-driven interaction can support early engagement with collaborative robotics concepts while preserving transparency and constraint awareness. Evidence associated with RQ1 demonstrates that novice learners can successfully formulate executable robot tasks through iterative refinement without reliance on low-level programming. Observations relevant to RQ2 show that constraint visibility and kinematic grounding promote spatial reasoning and feasibility awareness during task development. With respect to RQ3, natural-language interaction appears to reduce cognitive barriers and encourage exploratory experimentation, supporting learner confidence and engagement. Exploratory indicators related to RQ4 suggest emerging progression toward model-informed reasoning about reachability and safety constraints, although larger-scale studies are required for confirmation.

These findings demonstrate that reframing collaborative robotics education around intent-centric interaction, supported by narrative-driven digital twins, can meaningfully lower barriers to engagement while preserving conceptual rigor. Rather than simplifying robotics content, the proposed framework reorders the learning process, allowing students to reason about task intent, feasibility, and safety before confronting low-level implementation details. The pilot study suggests that this reordering supports earlier engagement, improved spatial reasoning, and productive exploration, even for students with no prior robot programming experience.

6.2 Educational Insights and Design Implications

A central insight emerging from this study is the educational value of explicit constraint visibility. By making joint limits, reachability, and collision conditions observable and actionable, the system encourages students to understand robot behavior as a constrained physical process rather than a black-box response to commands. This design supports learning through controlled failure, where infeasible actions prompt reflection and revision rather than silent error. Equally important is the deliberate limitation of Generative AI's role. Treating large language models as semantic interpreters rather than autonomous controllers preserves student agency, encourages critical evaluation of AI-assisted outputs, and mitigates risks associated with overreliance on opaque automation. In this way, Generative AI functions as a scaffold for reasoning rather than a substitute for learning.

From a practical and ethical perspective, the framework demonstrates that immersive, high-quality robotics learning experiences can be delivered without physical hardware. This capability has important implications for accessibility, scalability, and instructional flexibility, particularly in programs constrained by cost, safety policies, or laboratory availability. By embedding safety constraints directly into the digital twin and providing transparent feedback, the system reinforces responsible human–robot interaction from the earliest stages of learning. Narrative-based interaction and visual feedback further support inclusion for students with diverse technical and programming backgrounds, promoting equitable participation without compromising instructional rigor.

These observations further suggest alignment with threshold concept theory in engineering education, where conceptual bottlenecks such as spatial reasoning and constraint awareness are overcome through iterative experiential exposure. The narrative-driven digital twin may function as a transitional scaffold supporting passage through these conceptual thresholds, warranting targeted investigation in future longitudinal studies.

6.3 Limitations and Future Directions

This study is exploratory in scope and involves a small participant cohort, limiting generalizability and precluding statistical claims regarding learning gains or transfer. The current implementation focuses primarily on kinematic simulation, and real-time sensing, force interaction, and physical robot feedback were not fully integrated into the instructional workflow.

Future work will expand participant cohorts and incorporate structured assessment instruments, including validated surveys, performance rubrics, and pre/post measures aligned with ABET student outcomes. Planned system extensions include integration of real-time sensing for hybrid digital–physical learning, incorporation of advanced motion optimization techniques, and application of Explainable AI methods to enhance transparency in AI-assisted decision-making. Additional deployments will explore collaborative and remote learning scenarios involving shared digital twin environments.

6.4 Conclusion

This work contributes three primary outcomes to engineering education research: (1) an instructional architecture demonstrating how narrative-driven interaction can scaffold novice engagement with collaborative robotics concepts; (2) a deployable implementation showing that immersive, constraint-aware learning environments can be realized without specialized hardware barriers; and (3) exploratory empirical evidence indicating feasibility for supporting spatial reasoning, constraint awareness, and sustained engagement among early-stage learners. Together, these contributions position intent-centric interaction as a pedagogically viable pathway aligned with emerging Industry 5.0 human-centered automation paradigms.

The pilot cohort size was intentionally limited because the primary goal was interaction feasibility validation and instrumentation testing rather than outcome generalization. Such exploratory pilots are consistent with early-stage educational system evaluation where methodological refinement precedes scaled deployment. Results therefore inform design validity rather than population inference.

Acknowledgments

The authors gratefully acknowledge the support of the PIP grant, Summer Research Awards Program, and the Research Aligned Mentorship Program at SUNY Farmingdale State College, which enabled undergraduate student participation in the research activities reported in this paper. This support provided essential opportunities for students to engage in hands-on development, experimentation, and evaluation, substantially contributing to the outcomes of this work.

References

[1] Dhanda, M., Rogers, B.A., Hall, S., Dekoninck, E. and Dhokia, V., 2025. Reviewing human-robot collaboration in manufacturing: Opportunities and challenges in the context of industry 5.0. Robotics and Computer-Integrated Manufacturing, 93, p.102937.

[2] Zhang, M., Duan, P., Zhang, Z. and Esche, S., 2018, November. Development of telepresence teaching robots with social capabilities. In ASME International Mechanical Engineering Congress and Exposition (Vol. 52064, p. V005T07A017). American Society of Mechanical Engineers.

[3] Pons, P., Soler-Dominguez, J.L. and Navas-Medrano, S., 2025. Designing Gesture-Based Interfaces for Cobots: Insights from a Mixed Reality Elicitation Study. ACM Transactions on Human-Robot Interaction.

[4] Wendt, T.M., Blöchle, J.J., Stiglmeier, L. and Hangst, N., 2025, April. From University to Industry: Holistic Robotics Education for Tomorrow's Challenges. In International Conference on Robotics in Education (RiE) (pp. 19-30). Cham: Springer Nature Switzerland.

[5] Aschenbrenner, D., 2017. Human Robot Interaction Concepts for Human Supervisory Control and Telemaintenance Applications in an Industry 4.0 Environment. Bayerische Julius-Maximilians-Universitaet Wuerzburg (Germany).

[6] Al Zami, M.B., Shaon, S., Quy, V.K. and Nguyen, D.C., 2025. Digital twin in industries: A comprehensive survey. IEEE Access.

[7] Akhtar, Z.B., 2025. Generative artificial intelligence (GAI): From large language models (LLMs) to multimodal applications towards fine tuning of models, implications, investigations. Computing and Artificial Intelligence, pp.1498-1498.

[8] https://www.unrealengine.com/en-US, accessed in January, 2026.

[9] https://vention.io/parts/universal-robots-ur3e-collaborative-robot-arm-2446, accessed in January, 2026.

[10] Yanytska, L., 2025. The rise of human-centric manufacturing in the industry 5.0 era. The International Journal of Advanced Manufacturing Technology, 139(9), pp.5067-5077.

[11] Karanam, S.A.K. and Hartman, N.W., 2025. A systematic review of digital twin (DT) and virtual learning environments (VLE) for smart manufacturing education. Manufacturing letters, 44, pp.1597-1608.

[12] Darejeh, A., Chilcott, G., Oromiehie, E. and Mashayekh, S., 2025. Virtual Reality and Digital Twins for Mechanical Engineering Lab Education: Applications in Composite Manufacturing. Education Sciences, 15(11), p.1519.

[13] Wang, Y., Wang, X., Liu, A., Zhang, J. and Zhang, J., 2025. Ontology of 3D virtual modeling in digital twin: A review, analysis and thinking. Journal of Intelligent Manufacturing, 36(1), pp.95-145.

[14] Zhang, Z., Zhang, M., Chang, Y., Esche, S.K. and Chassapis, C., 2014, November. An efficient method for creating virtual spaces for virtual reality. In ASME International Mechanical Engineering Congress and Exposition (Vol. 46507, p. V005T05A016). American Society of Mechanical Engineers.

[15] Wagg, D.J., Worden, K., Barthorpe, R.J. and Gardner, P., 2020. Digital twins: state-of-the-art and future directions for modeling and simulation in engineering dynamics applications. ASCE-ASME Journal of Risk and Uncertainty in Engineering Systems, Part B: Mechanical Engineering, 6(3), p.030901.

[16] Nezzi, C., Gufler, V., Vidoni, R. and Rauch, E., 2025. Kinematic and dynamic modeling of mechanical systems towards Digital Twins. Results in Engineering, p.104874.

[17] Zhang, Z., Zhang, S., Zhang, M. and Esche, S.K., 2018, November. Immersive educational systems with procedure-oriented combinations of real and virtual environments. In ASME International Mechanical Engineering Congress and Exposition (Vol. 52064, p. V005T07A016). American Society of Mechanical Engineers.

[18] Nica, E., Popescu, G.H., Poliak, M., Kliestik, T. and Sabie, O.M., 2023. Digital twin simulation tools, spatial cognition algorithms, and multi-sensor fusion technology in sustainable urban governance networks. Mathematics, 11(9), p.1981.

[19] Ding, H., Seenivasan, L., Killeen, B.D., Cho, S.M. and Unberath, M., 2024. Digital twins as a unifying framework for surgical data science: the enabling role of geometric scene understanding. Artificial Intelligence Surgery, 4(3).

[20] Reethika, A. and Kanaga Priya, P., 2024. Personalize Learning Experience in Education Using Digital Twins with Human-Centered Design and Pedagogy. In Transforming Industry using Digital Twin Technology (pp. 165-185). Cham: Springer Nature Switzerland.

[21] Sharif, K.H. and Ameen, S.Y., 2021, September. Game engines evaluation for serious game development in education. In 2021 International Conference on Software, Telecommunications and Computer Networks (SoftCOM) (pp. 1-6). IEEE.

[22] Jungherr, A. and Schlarb, D.B., 2022. The extended reach of game engine companies: How companies like epic games and Unity technologies provide platforms for extended reality applications and the metaverse. Social Media+ Society, 8(2), p.20563051221107641.

[23] Chang, Y., Aziz, E.S.S., Zhang, Z., Zhang, M. and Esche, S.K., 2015, June. Usability evaluation of a virtual educational laboratory platform. In 2015 ASEE Annual Conference & Exposition (pp. 26-1637).

[24] Hagos, D.H., Battle, R. and Rawat, D.B., 2024. Recent advances in generative ai and large language models: Current status, challenges, and perspectives. IEEE transactions on artificial intelligence.

[25] García-Méndez, S., de Arriba-Pérez, F. and Somoza-López, M.D.C., 2025. A review on the use of large language models as virtual tutors. Science & Education, 34(2), pp.877-892.

[26] Li, W., & Hung, Y., & Altuger-Genc, G., & Zhang, S., & Tatoglu, A., & Zhang, Z. (2025, June), Integrating AI/ML Learning in Senior Projects for Mechanical Engineering Technology Students Paper presented at 2025 ASEE Annual Conference & Exposition , Montreal, Quebec, Canada . 10.18260/1-2--56818

[27] Pandy, G., Pugazhenthi, V.J., Murugan, A. and Jeyarajan, B., 2025. AI-Powered Robotics and Automation: Innovations, challenges, and pathways to the future. European Journal of Computer Science and Information Technology, 13(1), pp.33-44.

[28] Nikiforidis, K., Kyrtsoglou, A., Vafeiadis, T., Kotsiopoulos, T., Nizamis, A., Ioannidis, D., Votis, K., Tzovaras, D. and Sarigiannidis, P., 2025. Enhancing transparency and trust in AI-powered manufacturing: A survey of explainable AI (XAI) applications in smart manufacturing in the era of industry 4.0/5.0. ICT Express, 11(1), pp.135-148.

[29] Salih, A.M., Raisi‐Estabragh, Z., Galazzo, I.B., Radeva, P., Petersen, S.E., Lekadir, K. and Menegaz, G., 2025. A perspective on explainable artificial intelligence methods: SHAP and LIME. Advanced Intelligent Systems, 7(1), p.2400304.

[30] Yeong, D.J., Panduru, K. and Walsh, J., 2025. Exploring the unseen: A survey of multi-sensor fusion and the role of explainable ai (xai) in autonomous vehicles. Sensors, 25(3), p.856.

[31] Zeng, E., Long, Y., Wang, X., Xiao, Y. and Feng, Y., 2025. Literature Review: Personalized Learning Recommendation System in Educational Scenarios: XAI-Driven Student Behavior Understanding and Teacher Collaboration Mechanism. Frontiers in Interdisciplinary Applied Science, 2(1), pp.78-92.