

GenTutor-HE: A Hybrid Human-in-the-Loop Framework for Personalized, Ethically Governed Generative AI Integration in Higher Education

Dr. V. Maria Antoniate Martin¹, Praveen J²

¹Assistant Professor, Department of Information Technology, St. Joseph's College, Trichy, India E-mail:

²PG Student, Department of Information Technology, St. Joseph's College, Trichy, India

Abstract- Generative Artificial Intelligence (GenAI) has rapidly transitioned from an experimental technology into a mainstream instructional resource across higher-education classrooms. Building on five recent systematic and conceptual reviews of AI in education, this paper synthesizes the reported responses, attitudes, behaviors, applications, benefits, and challenges associated with GenAI adoption in university-level teaching and learning. The review of existing literature reveals that while GenAI substantially improves personalization, engagement, and assessment efficiency, its integration is constrained by ethical concerns, over-reliance risks, unequal infrastructure access, and a persistent lack of human-centered oversight. To address these gaps, this paper proposes GenTutor-HE, a hybrid, human-in-the-loop framework that combines a generative AI tutoring engine, an adaptive knowledge-tracing module, and an explicit ethical-governance layer that filters generated content for bias, privacy risk, and factual reliability before it reaches learners. The proposed methodology follows a design-science research approach comprising requirement analysis, architectural design, prototype development, and mixed-method evaluation. A comparative analysis of existing intelligent tutoring systems, adaptive learning platforms, and generative content-based systems is presented, along with the advantages and limitations of the proposed model. The paper concludes that a balanced, ethically governed hybrid approach offers a more sustainable pathway for GenAI adoption than fully automated systems, and it outlines future enhancements including multilingual support, emotion-aware feedback, and federated-learning-based privacy preservation.

Keywords— Generative Artificial Intelligence, Higher Education, Intelligent Tutoring Systems, Adaptive Learning, Human-in-the-Loop, Ethical AI Governance, Personalized Learning

I. INTRODUCTION

Artificial Intelligence (AI), and more specifically Generative AI (GenAI) tools such as ChatGPT and other large language models (LLMs), have become deeply embedded in the everyday practice of higher-education teaching and learning. Reviews of the literature published between 2020 and 2026 consistently report that the number of empirical studies examining GenAI in classrooms has grown sharply, with the overwhelming majority of studies concentrated in 2023 and 2024, corresponding to the rapid global uptake of conversational AI tools following the public release of ChatGPT.

The appeal of GenAI in education stems from its capacity to personalize instructional content, automate routine assessment tasks, generate on-demand explanations, and provide round-the-clock

feedback to learners regardless of geographic location. At the same time, researchers have raised recurring concerns regarding the quality and reliability of generated content, the erosion of critical thinking through over-reliance, data-privacy risks, algorithmic bias, and unequal access to the infrastructure required to benefit from these tools.

This paper makes three contributions. First, it consolidates the findings of five recent reviews of AI and GenAI in education into a single, structured literature review organized around responses, attitudes, behaviors, applications, benefits, and challenges. Second, it compares the dominant families of existing AI-based educational systems - intelligent tutoring systems (ITS), adaptive learning platforms, learning-analytics systems, and generative content engines - highlighting their respective strengths and weaknesses. Third, it proposes GenTutor-HE, a hybrid architecture that

couples a generative tutoring engine with an explicit ethical-governance layer and a human-in-the-loop review mechanism, intended to capture the pedagogical benefits of GenAI while mitigating its most frequently reported risks.

II. LITERATURE REVIEW

1. Responses, Attitudes, and Behaviors toward GenAI

A systematic review synthesizing 99 empirical studies published between 2020 and August 2024 examined how GenAI use affects students' and teachers' responses, attitudes, and behaviors in higher-education classrooms. The majority of studies (73 of 99) reported positive emotional responses such as curiosity, satisfaction, and enjoyment, while a smaller but significant proportion (46 of 99) reported negative emotional responses including distrust, frustration, and fear of academic misconduct. Regarding attitudes, 80 of 99 studies reported that GenAI enabled a new, integrated approach to professional learning, while 44 studies raised concerns about the low quality of generated content and 34 studies flagged the risk of over-reliance. In terms of behavior, GenAI was found to promote cognitive-processing skills such as critical thinking and problem-solving in 52 of 99 studies, although 25 studies found no statistically significant improvement in task performance when GenAI was used, indicating that outcomes are highly context-dependent.

2. The Four Roles of AI in Education

A separate systematic review covering 30 articles indexed in Scopus and Google Scholar between 2014 and 2023 proposed a four-role framework describing how AI functions within educational applications: as an intelligent tutor (the largest category, encompassing intelligent tutoring systems and adaptive/personalized learning systems), as an intelligent student (a comparatively underexplored role in which AI systems learn from human interaction), as an intelligent learning tool or partner (supporting data collection, concept mapping, and knowledge-graph construction), and as a policy-making advisor (assisting institutional decision-makers with macro- and micro-level analysis of

educational trends). This framework underscores that AI's role in education extends well beyond simple tutoring automation.

3. Applications, Benefits, and Challenges

A conceptual review of AI in education identified four principal application areas: content design (using natural-language generation to produce consistent, personalized instructional material), content delivery (enabling flexible, location-independent access to instruction), automated assessment (grading objective item types and, increasingly, essays), and feedback and support (chatbot-based, round-the-clock assistance). Reported benefits include greater flexibility, improved accessibility, reduced institutional costs, and a shift in the teacher's role toward facilitation. Reported challenges include ethical issues such as data privacy and bias, unequal access linked to socio-economic status, dehumanization of the learning experience, inadequate institutional preparedness, and the risk that students fail to develop independent critical-thinking skills.

4. GenAI as a Transformative Force

A qualitative content-analysis study of AI as a transformative agent in modern learning environments similarly concluded that AI supports independent, self-paced learning and increases student motivation and engagement, while cautioning that unregulated use may foster technological dependency, reduce systematic thinking, and increase the incidence of plagiarism. The study emphasized that the teacher's evolving role as facilitator and mediator is central to ensuring that GenAI is used wisely rather than as an unquestioned substitute for human judgment.

5. AI Adoption Factors in Learning Systems

A recent systematic review of 32 peer-reviewed studies published between 2020 and 2025 examined the factors influencing AI adoption in learning systems through the lens of the Technology Acceptance Model (TAM) and socio-technical systems theory. The review found that adoption is shaped jointly by technical factors (ease of use, infrastructure readiness), social and pedagogical factors (AI literacy, institutional support, cultural

adaptation), and data-readiness factors (data quality and model adaptivity). Reported pedagogical benefits included learning-outcome improvements of up to 68.3% in some adaptive-assessment studies and engagement-efficiency gains of up to 20.26% in intelligent-tutoring-system deployments, while ethical and privacy concerns, algorithmic bias, and infrastructure limitations were identified as the principal barriers to sustainable adoption.

6. Summary of Gaps

Across all five reviews, three recurring gaps emerge: (i) most existing systems optimize for either personalization or assessment efficiency but rarely embed an explicit, auditable ethical-governance mechanism; (ii) the teacher's role as a human-in-the-loop safeguard is frequently discussed conceptually but rarely operationalized as a concrete system component; and (iii) comparatively little empirical

attention has been paid to the instructor's perspective relative to the student's. These gaps motivate the hybrid framework proposed in Section 4.

III. EXISTING ALGORITHMS AND SYSTEMS

Before presenting the proposed framework, it is useful to compare the dominant categories of AI-based educational systems reported in the literature. Table 1 summarizes four major categories: rule-based/classical Intelligent Tutoring Systems (ITS), adaptive learning platforms driven by knowledge-tracing algorithms, learning-analytics/prediction systems, and generative-AI-based tutoring or content-generation systems.

Table 1. Comparison of Existing AI-Based Educational System Categories

Category	Representative Technique	Key Strength	Key Limitation
Classical Intelligent Tutoring Systems (ITS)	Rule-based dialogue trees, cognitive tutors (e.g., AutoTutor)	High pedagogical transparency; well-validated learning gains	Limited flexibility; costly to author and maintain content
Adaptive Learning Platforms	Item Response Theory, Computerized Adaptive Testing, Zone-of-Proximal-Development models	Dynamically matches task difficulty to learner ability	Requires large, high-quality item banks and continuous calibration
Knowledge-Tracing / Learning-Analytics Systems	Bayesian Knowledge Tracing, deep-learning models (e.g., self-attentive convolutional models)	Accurate prediction of learner performance and mastery	Prediction accuracy depends heavily on data volume and quality; can encode bias
Generative-AI Tutoring / Content Systems	Large Language Models (e.g., ChatGPT-based tutors), retrieval-augmented generation	Natural-language interaction; rapid content generation; scalable feedback	Hallucination risk, over-reliance, weak built-in ethical safeguards

As shown in Table 1, each category addresses a different facet of the teaching-learning process, but no single category simultaneously provides open-ended natural-language interaction, personalized difficulty adjustment, and an explicit ethical safeguard. Classical ITS and knowledge-tracing

systems are pedagogically well-validated but rigid, whereas generative-AI systems are flexible and scalable but comparatively unconstrained. This complementary gap is the primary motivation for the hybrid architecture proposed in the next section.

IV. PROPOSED SYSTEM: GENTUTOR-HE ARCHITECTURE

This paper proposes GenTutor-HE, a hybrid architecture that integrates a generative AI tutoring engine with an adaptive knowledge-tracing model and an explicit ethical-governance layer, coordinated through a human-in-the-loop review mechanism. Figure 1 presents the overall system architecture.

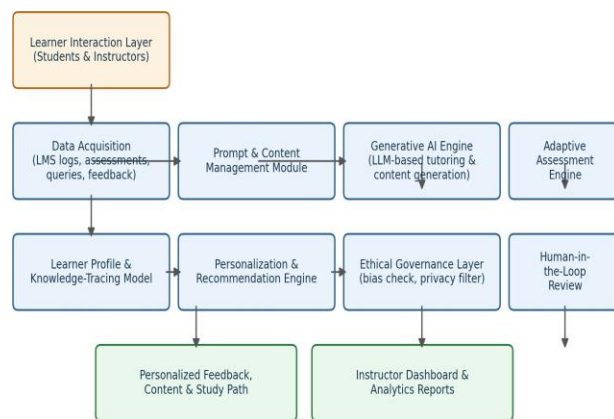


Figure 1. Proposed GenTutor-HE System Architecture

The architecture consists of six interacting layers. The Learner Interaction Layer captures natural-language queries, assignments, and feedback from students and instructors. The Data Acquisition module aggregates learning-management-system logs, assessment records, and interaction histories. The Prompt and Content Management module structures these inputs into task-appropriate prompts before they reach the Generative AI Engine, which produces tutoring dialogue, explanations, and draft feedback using a large language model. In parallel, the Adaptive Assessment Engine applies computerized-adaptive-testing principles to select the next appropriate task difficulty for each learner. The Learner Profile and Knowledge-Tracing Model maintains a continuously updated estimate of each

learner's mastery of specific concepts, feeding this estimate into the Personalization and Recommendation Engine, which tailors the sequencing and format of content. Critically, all generated content and recommendations pass through an Ethical Governance Layer that screens for factual inconsistency, bias, and privacy risk (for example, by detecting personally identifiable information before it is stored or displayed) before reaching the learner. A configurable subset of interactions - particularly high-stakes assessments, flagged content, or first-time deployments in a course - is additionally routed to a Human-in-the-Loop Review step, in which the instructor validates or overrides system output. Finally, the system produces two outputs: personalized feedback, content, and a recommended study path for the learner, and an analytics dashboard for the instructor summarizing class-wide mastery trends and any content flagged during governance review.

1. Design Rationale

The design deliberately keeps the instructor inside the decision loop rather than treating GenAI as a full substitute for the teacher, directly addressing the 'lack of human touch' and 'over-reliance' concerns repeatedly identified in the reviewed literature. The explicit ethical-governance layer operationalizes the ethics, privacy, and bias concerns that prior studies discuss largely at a conceptual level, converting them into a concrete, auditable pipeline stage rather than leaving them as an afterthought.

V. METHODOLOGY

The development and evaluation of GenTutor-HE follows a Design Science Research (DSR) methodology, comprising four phases, summarized in Table 2.

Table 2. Design Science Research Methodology for GenTutor-HE

Phase	Activities	Expected Output
1. Problem Identification & Requirements	Synthesis of reviewed literature; semi-structured interviews with instructors and students to elicit functional and ethical requirements	Requirements specification and use-case models
2. Architectural Design	Definition of system layers (Figure 1); selection of underlying LLM, knowledge-tracing algorithm, and governance rule-set	System architecture and data-flow specification
3. Prototype Development	Implementation of a minimum-viable prototype integrating an LLM API, an adaptive-testing module, and a rule-based/classifier-based content filter	Working prototype deployed on a pilot course
4. Evaluation	Mixed-method pilot study: pre/post learning-outcome tests, System Usability Scale (SUS) survey, instructor review logs, and thematic analysis of flagged-content cases	Quantitative learning-gain results and qualitative usability/ethics findings

Evaluation would combine quantitative measures (pre-/post-test learning gains, task-completion time, System Usability Scale scores) with qualitative measures (thematic analysis of instructor review logs and student interview transcripts), consistent with the mixed-methods approaches most frequently reported as effective in the reviewed literature.

Advantages and Limitations of the Proposed Framework

Table 3 summarizes the anticipated advantages and limitations of GenTutor-HE relative to the existing system categories reviewed in Section 3.

Table 3. Advantages and Limitations of the Proposed GenTutor-HE Framework

Aspect	Advantages	Limitations
Personalization	Combines LLM flexibility with knowledge-tracing accuracy for finer-grained adaptation	Requires substantial historical learner data to calibrate effectively
Ethical Safeguards	Explicit governance layer reduces exposure to biased or unreliable generated content	Governance rules must be continuously updated as new failure modes emerge
Instructor Role	Human-in-the-loop review preserves teacher agency and accountability	Adds review workload unless review is intelligently triaged
Scalability	Generative engine allows rapid, low-cost content generation at scale	Increases dependency on third-party LLM providers and compute cost
Engagement	Conversational interface reported to increase motivation and self-paced learning	Risk of over-reliance if governance/review safeguards are relaxed

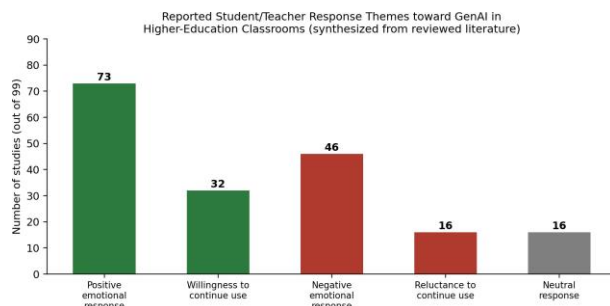


Figure 2. Reported Student/Teacher Response Themes toward GenAI (synthesized from reviewed literature)

Figure 2 illustrates that, across the reviewed studies, positive responses toward GenAI substantially outnumber negative ones, but negative and neutral responses remain non-trivial, reinforcing the argument that ethical governance and human oversight - rather than unmoderated deployment - are necessary for sustainable adoption.

VI. CONCLUSION

This paper synthesized findings from five recent reviews of Artificial Intelligence and Generative AI in higher education, covering student and teacher responses, attitudes, behaviors, applications, benefits, and challenges. The synthesis confirms that GenAI offers substantial pedagogical value, including personalized instruction, improved engagement, and reduced instructor workload, but that this value is consistently offset by ethical concerns, risks of over-reliance, and uneven infrastructure access. In response, this paper proposed GenTutor-HE, a hybrid, human-in-the-loop framework that integrates generative tutoring, adaptive assessment, and an explicit ethical-governance layer. Compared with existing classical ITS, adaptive-learning, and unmoderated generative-AI systems, the proposed framework is designed to combine the flexibility of large language models with the pedagogical rigor of knowledge-tracing methods and the accountability of continued instructor oversight. The proposed Design Science Research methodology provides a concrete pathway for implementing and empirically evaluating the framework in a real classroom setting.

Future Enhancements

- Multilingual generative tutoring to extend accessibility to non-English-speaking learner populations, addressing the language-bias gap identified in prior reviews.
- Emotion-aware feedback generation, building on emerging evidence that emotionally enriched AI feedback can improve well-being without compromising learning outcomes.
- Federated-learning-based model training to strengthen data-privacy protection while still allowing the knowledge-tracing model to learn from institution-wide data.
- Explainable-AI (XAI) dashboards that allow instructors to inspect why the ethical-governance layer flagged or approved specific generated content.
- Longitudinal, cross-institutional evaluation studies to assess the durability of learning gains and instructor-trust levels over multiple academic terms.
- Integration with existing Learning Management Systems (LMS) through open standards (e.g., LTI) to reduce institutional adoption cost and technical barriers.

REFERENCES

1. Wu, F., Dang, Y., & Li, M. (2025). A systematic review of responses, attitudes, and utilization behaviors on generative AI for teaching and learning in higher education. *Behavioral Sciences*, 15(4), 467.
2. Bakti, I. K., Zulkarnain, Yarun, A., Rusdi, Syaifudin, M., & Syafaq, H. (2023). The role of artificial intelligence in education: A systematic literature review. *Jurnal Iqra': Kajian Ilmu Pendidikan*, 8(2), 182-197.
3. Tahir, M., Hassan, F. D., & Shagoo, M. R. (2024). Role of artificial intelligence in education: A conceptual review. *World Journal of Advanced Research and Reviews*, 22(1), 1469-1475.
4. Gafarurrozi, M., Rohman, M. F., & Fathurrohman, R. (2024). The role of artificial intelligence as a transformation of learning in the modern era. *International Journal of Science and Applied Science: Conference Series*, 8(1), 32-40.

6. Sama, H., Nurkayla, A., & Prasetyo, S. E. (2026). Artificial intelligence adoption in learning systems: Systematic literature review. *JISA (Jurnal Informatika dan Sains)*, 9(1), 10-19.
7. Hwang, G. J., Xie, H., Wah, B. W., & Gašević, D. (2020). Vision, challenges, roles and research issues of artificial intelligence in education. *Computers and Education: Artificial Intelligence*, 1, 100001.
8. Ahmad, S. F., Rahmat, M. K., Mubarik, M. S., Alam, M. M., & Hyder, S. I. (2021). Artificial intelligence and its role in education. *Sustainability*, 13(22), 12902.
9. Ouyang, F., & Jiao, P. (2021). Artificial intelligence in education: The three paradigms. *Computers and Education: Artificial Intelligence*, 2, 100020.
10. Zhang, K., & Aslan, A. B. (2021). AI technologies for education: Recent research and future directions. *Computers and Education: Artificial Intelligence*, 2, 100025.
11. Khosravi, H., Shum, S. B., Chen, G., Conati, C., Tsai, Y. S., Kay, J., Knight, S., Martinez-Maldonado, R., Sadiq, S., & Gašević, D. (2022). Explainable artificial intelligence in education. *Computers and Education: Artificial Intelligence*, 3, 100074.
12. Long, D., & Magerko, B. (2020). What is AI literacy? Competencies and design considerations. *Proceedings of the 2020 CHI Conference on Human Factors in Computing Systems*, 1-16.
13. Holmes, W., & Tuomi, I. (2022). State of the art and practice in AI in education. *European Journal of Education*, 57(4), 542-570.
14. Kabudi, T., Pappas, I., & Olsen, D. H. (2021). AI-enabled adaptive learning systems: A systematic mapping of the literature. *Computers and Education: Artificial Intelligence*, 2, 100017.
15. Chiu, T. K., Xia, Q., Zhou, X., Chai, C. S., & Cheng, M. (2023). Systematic literature review on opportunities, challenges, and future research recommendations of artificial intelligence in education. *Computers and Education: Artificial Intelligence*, 4, 100118.
16. Baidoo-Anu, D., & Ansah, L. O. (2023). Education in the era of generative artificial intelligence (AI): Understanding the potential benefits of ChatGPT in promoting teaching and learning. *Journal of AI*, 7(1), 52-62.
17. Yang, A. C. M., Flanagan, B., & Ogata, H. (2022). Adaptive formative assessment system based on computerized adaptive testing and the learning memory cycle for personalized learning. *Computers and Education: Artificial Intelligence*, 3, 100104.
18. Kia, M. A., & Koliouris, A. (2024). ACKT: Self-attentive convolutional model for knowledge tracing. *Procedia Computer Science*, 246, 2723-2732.
19. Alier, M., Pereira, J., García-Peñalvo, F. J., Casañ, M. J., & Cabré, J. (2025). LAMB: An open-source software framework to create artificial intelligence assistants deployed and integrated into learning management systems. *Computer Standards & Interfaces*, 92, 103940.
20. Yang, A. C. M., Lin, J. Y., Lin, C. Y., & Ogata, H. (2024). Enhancing python learning with PyTutor: Efficacy of a ChatGPT-based intelligent tutoring system in programming education. *Computers and Education: Artificial Intelligence*, 7, 100309.
21. Ferguson, C., van den Broek, E. L., & van Oostendorp, H. (2022). AI-induced guidance: Preserving the optimal zone of proximal development. *Computers and Education: Artificial Intelligence*, 3, 100089.
22. Liu, J., Sun, D., Sun, J., Wang, J., & Yu, P. L. H. (2025). Designing a generative AI enabled learning environment for mathematics word problem solving in primary schools. *Computers and Education: Artificial Intelligence*, 9, 100438.
23. Moorhouse, B. L., Wan, Y., Wu, C., Kohnke, L., Ho, T. Y., & Kwong, T. (2024). Developing language teachers' professional generative AI competence: An intervention study in an initial language teacher education course. *System*, 125, 103399.
24. Akgun, S., & Greenhow, C. (2022). Artificial intelligence in education: Addressing ethical challenges in K-12 settings. *AI and Ethics*, 2(3), 431-440.
25. Pedro, F., Subosa, M., Rivas, A., & Valverde, P. (2019). Artificial intelligence in education: Challenges and opportunities for sustainable development. *UNESCO*.

26. Alsaieri, O., Baghaei, N., Lahza, H., Lodge, J. M., Boden, M., & Khosravi, H. (2025). Emotionally enriched AI-generated feedback: Supporting student well-being without compromising learning. *Computers & Education*, 239, 105363.