S A M Y R A D 2026

SMART ALGORITHMS & MODELS: YOUR REACH TO ADVANCED DIALOGUES · 5–6 OCTOBER 2026 · SEVILLE, SPAIN
SAMYRAD OPEN-ACCESS PROCEEDINGS | PUBLISHER: INTELLIGENT NUDGE

PAPER • OPEN ACCESS

A Research-Embedded Implementation Model for Design of AI Applications in Higher Education

Hiba Kassab and Nancy Maynes

Nipissing University, Canada

Article DOI: https://doi.org/10.68286/samyrad2026.309

ABSTRACT

In the context of higher education, particularly in the current era of expansive growth of Artificial Intelligence (AI), the challenges that confront students and faculty are numerous and multifaceted. Potential paths for researching and defining effective implementation of this technology are complex. This article provides an emerging model for further discussion of exploring, utilizing, and expanding the implementation of AI in educational contexts, with a focus on higher education. Meta-synthesis has been applied to 1679 studies to identify themes/recurring knowledge and uncover recurring knowledge to inform processes for adopting AI use in education and optimizing its potential to enhance instruction and learning. ChatGPT GenAI synthesis identified 12 recurring meta-themes, whereas Copilot GenAI synthesis identified 13 recurring meta-themes that describe how GenAI technologies are used and experienced in higher education environments. Both AI tools produced frameworks to illustrate their findings. ChatGPT's framework synthesizes the mixed role of GenAI in higher education, highlighting its pedagogical values and associated risks, and tracing its implications for learning processes, student experience, instructional design, and institutional governance, whereas Copilot's framework illustrates how effective GenAI integration relies on coordinated pedagogical design, institutional governance, operational stability, and equity-focused practices. Four main ideas have been identified as consistently displayed in previous studies on use in higher education, including learning, pedagogy, governance, operations, and inclusion/industry, potentially influencing institutional and instructor readiness to use AI effectively in the higher learning environment.

KEYWORDS Meta-synthesis, Artificial Intelligence, ChatGPT, Higher education, Copilot

How to cite: Kassab, H., & Maynes, N. (2026). A research-embedded implementation model for design of AI applications in higher education. SAMYRAD 2026 Open-Access Proceedings. Intelligent Nudge. https://doi.org/10.68286/samyrad2026.309

Introduction

In higher education, amid the rapid growth of Artificial Intelligence (AI), the challenges facing students and faculty are numerous and multifaceted. AI and immersive technologies such as virtual and augmented reality are reshaping "virtual classrooms" into "modified learning formats and flexible course modules" [1], while simultaneously redefining the global educational landscape. As Elayyan [2] notes, integrating Industry 4.0 technologies into education requires cultivating "liquid or soft skills [such as] critical thinking, coordinating with others, verbal communication, and time management" alongside traditional knowledge-based skills, a shift accelerated by the rise of Massive Open Online Courses (MOOCs) as a democratizing, disruptive force in education [2].

These advancements also bring challenges: redefining "learning outcomes and curricula" and preparing students for a labour market where traditional roles may give way to emerging professions requiring AI literacy [2]. This article explores these implications, emphasizing the necessity for higher education to

adapt and prepare students for an AI-driven future. It is intended to help post-secondary instructors and institutions adopt a proactive, evidence-informed approach to this reality by synthesizing current qualitative research on the instructional use of AI and how best to use it.

Background

While article length limitations prohibit lengthy definitions of AI and claims of its popularity in our culture [3]; [4], it is valuable to define the different types of AI as they are currently presented in the research literature, while simultaneously recognizing that these classifications are in continuous shift with the developments in technology. DiCarlo [5] categorizes AI into three stages: Artificial narrow intelligence (or ANI), "artificial general intelligence (or AGI), and artificial super intelligence (or ASI)". Currently, most AI applications use ANI, designed to execute "specific tasks and to directly follow its digital command", in applications such as recommendation systems, speech recognition, and automated grading tools, generally purposed to enhance task efficiency.

AGI represents AI systems capable of independent reasoning, "problem-solving, learning, and long-term planning abilities, as well as and, eventually, much better than any human" [5]. AGI can learn, adapt, and generate novel solutions similar to human cognition; however, it also raises concerns regarding ethical decision-making, job displacement, and the potential loss of human oversight in critical sectors. Additionally, users will have experienced the cultural nature of the current "learning" of some AGI tasks. For example, depending on where a program was trained in certain content, a lesson on writing structure might default to the five-paragraph essay format (indicating that it was trained in North American curricula) but exclude oral storytelling traditions or argumentative structures common in other cultures or educational systems.

DiCarlo's third stage, ASI, refers to "superintelligence, … intelligence that surpasses the cognitive abilities of any individual human or even the combined intellect of all humanity" [5], with the resultant existential risks it entails. Speculation about ASI risks has implications for education, society, and global stability, highlighting the need for informed, responsible AI development and regulation to ensure that technological advancements align with human values and safety in higher education.

Our current applications of AI draw on several disciplines that have contributed to its current structures. Among these machine learning uses [6], are, analysis of algorithms to identify patterns and optimize performance through exposure to extensive datasets [6]; machine-to-machine interactions (also called IoT or the Internet of Things) that extend beyond human-to-machine interactions [7]; [8]; formal representations of algorithmic thinking and computational models [9]; intelligent tutoring systems and frameworks for educational technology systems [10]; human information processing (e.g., memory, attention, and decision-making) applied to computational frameworks and selective attending [11]; formal models of reasoning and problem-solving [12]; neural-network architectures, brain structure and functions and neural network mimicking [13]; computational linguistics foundations for parsing human language into machine-readable structures [14]; and the simulation of human cognitive processes [4].

AI is reshaping education by creating interconnected systems to personalize learning experiences. This shift may optimize educational processes and demonstrate the broader impact of IoT, where devices interact autonomously to enhance efficiency and accessibility across various domains [15]; [16]; [17]. It may also make the post-secondary education world uncertain about adoption steps and its own values and principles as AI emerges into instructional and institutional territory. This revolution underscores the urgency of our understanding of current research on AI in our learning environments.

Methodology

A meta-synthesis is needed to identify the themes that emerge from the interaction among faculty who teach, students who learn, administrators who process, and the AI tools available to enhance this collaboration. Systematic literature reviews and meta-level studies can provide "systematic ways of collecting and synthesizing previous research" [18] to create a theoretical framework for AI adoption [4]; [19] and form the basis for "building a conceptual model or theory" [18].

This study began with a broad inclusion of relevant qualitative studies, gradually applying further criteria to prioritize measures of validity, reliability, generalizability, and relevance, and refining the full range of available studies to those most likely to enrich a meta-synthesis theory. By synthesizing diverse perspectives, there is potential to uncover the nuanced ways in which AI is reshaping higher educational practices, through "gathering rich data to discover underlying patterns and construct theories that are grounded in the data" [20].

Adopting the interpretivist/constructivist framework in meta-synthesis and grounded theory has key implications, including how data are collected, analyzed, and how valid the conclusions may be [21]. The meta-synthesis procedure as outlined by Walsh and Downe [22] was adopted for this study. Six steps of that procedure were followed. including:

Framing a meta-synthesis exercise: defining clear research questions, such as "What are the emergent themes in AI applications in higher education?"

Locating relevant papers: conducting a comprehensive search for qualitative studies about AI in higher education across various databases.

Deciding what to include: selecting all relevant qualitative studies, focusing on those related to AI in higher education, teaching, learning, or administration.

Appraisal of studies: evaluating the quality and relevance of the chosen studies, with an emphasis on inclusivity.

Analytical technique: employing a grounded theory approach

for data analysis, extracting relevant data through comparison, contrast, and thematic synthesis.

Synthesis of translation*:* integrating the findings to reveal refined meanings, exploratory theories, and new concepts that contribute to a comprehensive understanding of the topic.

This was an iterative process, using constant comparative approaches to isolate studies from the original 1679 (see Table 1), to focus analysis on those most closely aligned with strong qualitative methodologies and therefore most reliable in their conclusions. After filtering the initial body of related AI studies using various criteria, 154 studies were selected for detailed analysis of themes and findings. Table 1 shows the original body of studies from the full range of AI search engines that were used.

Table 1. Number of articles generated from the different AI tools and other search engines

AI tool/ Search Engine	Number of articles generated
Undermind (April 04-2025)	598
ChatGPT03-Deep Research	58
ChatGPT03-Search	124
Undermind (July 10-2025)	122
Research Rabbit (July 26-2025)	488
Deepseek and Scholar	201
Other articles	88
Total	**1679**

To investigate AI, AI tools were used: ChatGPT (versions 3 and 3-Deepsearch), Undermind, and ResearchRabbit, to identify relevant studies. In addition, online databases for journal articles and chapters were employed in the broadest stages of the search (e.g., Google Scholar). As searches were initiated, inclusion criteria were made specific to ensure adherence to the standards and the post-positivist, constructivist approach. These inclusion criteria included: relevance to AI in higher education; qualitative study design; peer-reviewed; currency (within the last 10 years); empirical studies with rich thematic insights, and a focus on sustainability and educational outcomes. This study excluded previous research that used only quantitative or mixed methods, was conducted in unrelated educational contexts, was out of date (10 years or more), was published in a language other than English, was not peer-reviewed, or had an obsolete technological focus. The application of both inclusion and exclusion criteria to the initial data set of 1679 articles is displayed in Table 2.

Table 2. Number of articles per criterion for inclusion/exclusion

Criterion	Number of Articles
Duplicates	260
Mixed Methods	130
Quantitative	243
Not English, irrelevant, no record found, retracted, restricted access	299
Date parameter	33
Non-Peer-reviewed	15
Qualitative	792

The remaining 154 studies that met the rigid criteria for this purpose were analyzed using the Walsh and Downe [22] method.

Research Design

This study employed a qualitative meta-synthesis to integrate and interpret findings from published qualitative papers on GenAI in higher education. This method permits researchers to move beyond aggregation by re-analyzing and conceptually extending qualitative results across multiple studies [23]. An analytical strategy was applied for the retrieval of the thematic and interpretive synthesis [24], which involves (1) generating codes that remain faithful to participants' or authors' language; (2) grouping conceptually related codes into higher-order themes; and (3) developing integrative propositions that transcend individual studies. This approach is widely used in qualitative synthesis because it enables researchers to integrate qualitative findings methodically while maintaining transparency and traceability to the original data sources [25]. The analysis focuses on identifying recurring patterns in educational practice, as well as the opportunities and challenges associated with GenAI adoption.

Data Sources and Coding Architecture

The synthesis drew on two distinct outputs generated using the two GenAI tools: Copilot and ChatGPT. Copilot was chosen because it is the AI tool approved for use across Universities and Colleges in Canada, whereas ChatGPT was picked for its global reputation as the go-to GenAI tool. The developed coded datasets ("Coding-Copilot" and "Coding-ChatGPT" workbooks) each contain qualitative excerpts extracted from the same peer-reviewed qualitative studies on GenAI in higher education contexts. Each coded entry retained four elements:

1. The source article from which the excerpt originated;
2. The excerpt itself, representing a relevant segment of qualitative evidence;
3. The initial code assigned during the coding phase; and
4. The corresponding thematic interpretation.

Maintaining this structured coding format ensured that each analytical theme could be traced back to its original textual evidence, thereby preserving methodological transparency and an auditable trail of interpretation. Establishing such traceability is essential in qualitative synthesis research to enhance transparency and accuracy [23]. The coded articles were subsequently consolidated into 13 themes (Copilot) and 12 themes (ChatGPT), which served as the structural basis for the two synthesis results.

Coding and Thematic Consolidation

The coding process followed an iterative thematic synthesis approach, consistent with qualitative meta-synthesis methodologies described by Thomas and Harden [24] Initial coding preserved the verbatim text retrieved from the studies using Copilot and ChatGPT, and the language and meaning of the original excerpts were kept as closely as possible to maintain fidelity to the source material. Two workbooks were created to

conduct the thematic analysis for each GenAI tool.

At the specific-code level, retaining open codes that reflect participants' or authors' language to maintain the semantic integrity of the original language during early coding is a recommended practice because it minimizes interpretive distortion during the initial analytical stage [26].

Following the initial coding phase, codes were grouped into broader conceptual categories through axial coding, a process that organizes related codes according to semantic and conceptual similarities. This stage resulted in the development of 13 thematic categories for Copilot and 12 for ChatGPT, representing key domains of GenAI use and impact in higher education.

To improve thematic precision, the coding framework underwent a two-pass refinement process, including:

Pass 1: Rule-based classification. In the first refinement stage, decisions were applied to ensure consistent thematic assignment. For example:

- Codes referring to hallucinated outputs or inaccurate AI responses were classified under Technical Limitations.
- Codes describing tutoring, explanations, or conceptual clarification were assigned to Learning Support.

Pass 2: Pragmatic thematic adjustment. A second refinement stage adjusted the codes based on the contextual interpretation of the quote. For example:

- Codes referring to time savings or productivity improvements were reassigned to the Efficiency and Productivity category.
- Codes referring to the integration of human instruction and AI systems were grouped under Human-AI Pedagogy.

This iterative consolidation process reduced reliance on overly broad categories such as Other or General while improving the conceptual clarity of the thematic structure.

Thematic Synthesis and Analytical Calibration

The synthesis progressed through three levels of analysis when using the Copilot GenAI tool.

Within-category synthesis: Coded excerpts within each category were examined to identify recurring subthemes. These subthemes were informed by two quantitative indicators, involving:

- Article counts indicating the number of unique studies contributing evidence to a theme; and
- Occurrence counts indicating the frequency with which the theme appeared in the dataset.

Meta-level conceptual synthesis*:* The final stage of analysis involved developing higher-order propositions that connected individual themes into a broader interpretive framework; patterns observed across the dataset were grounded in the coded evidence extracted from the source studies.

Throughout this process, the emphasis in the findings was calibrated to the thematic counts (recorded in Table 3) to reduce the risk of over-representing themes supported by limited evidence, thereby maintaining alignment between analytical claims and the distribution of coded data.

Limitations and trustworthiness

Several strategies were used to enhance the trustworthiness of this synthesis. First, traceability was maintained through structured coding of Excel ensuring an audit trail [26]. Second, unique codes, occurrences, and contributing articles, allowed thematic emphasis to be adjusted according to the distribution of evidence across the dataset and to reduce the tendency to overinterpret themes based on limited sources.

Nevertheless, several limitations should be acknowledged. The coded dataset reflects heterogeneous source coverage, meaning that some thematic categories draw on fewer contributing studies than others. For example, themes related to data privacy and security were supported by fewer articles and should therefore be understood as offering directional insight rather than definitive conclusions. Additionally, the rapidly evolving nature of GenAI technologies introduces the possibility of data deviation, in which technological developments may overtake the evidence available in the literature.

Tool-Specific sensitivity. A further limitation concerns the extent to which the findings reported here are attributable to GenAI-assisted synthesis in general versus the particular behaviour of the two tools used, ChatGPT and Copilot, at the specific versions and points in time this study was conducted. The two tools organized identical coded evidence according to different underlying logics, producing frameworks that diverged in the number of excerpts extracted, the themes, and the structure, despite being subject to the same human validation and adjudication procedure. This suggests that at least part of the study's output reflects each tool's default synthesis behaviour rather than solely the properties of the underlying literature. Because GenAI tools are updated frequently and their synthesis behaviour is not fully documented or controlled by the human using them, this study's specific findings may not replicate identically if repeated with newer versions of these tools, or with other GenAI systems entirely. This limitation does not undermine the validity of the human-adjudicated codes and themes, which were independently verified against source excerpts regardless of which tool proposed them; rather, it constrains the generalizability of the specific frameworks presented in Figs. 1 and 2, which should be read as one exemplification of GenAI-assisted synthesis rather than a stable ground truth. Future work replicating this design across additional GenAI tools, or at multiple points in time with the same tools, would help distinguish which patterns are durable features of the underlying evidence and which are exploratory of a given tool's synthesis at a given point in its development.

Table 3. Thematic findings summary using Copilot GenAI: Limitations and Trustworthiness

Theme	Concise definition	Unique codes	Occurrences	Unique articles
Learning Support & Understanding	On-demand explanations, scaffolding, reassurance	216	286	**20**
Engagement & Participation	AI as conversation starter; critique of outputs	94	127	**5**
Efficiency & Productivity	Time compression for routine tasks	98	112	**7**
Assessment & Feedback	AI for breadth/speed; humans for specificity/affect	63	67	**2**
Accessibility & Inclusion	Language support, captioning, neurodiversity	97	113	**4**
Academic Integrity & Ethics	Concerns: plagiarism, fabricated citations; redesign	123	146	**5**
Data Privacy & Security	Sandboxing, provenance, consent	30	33	**2**
Human–AI Pedagogy	Designed use; role-play; scenario-based learning	88	95	**4**
Technical Limitations & Accuracy	Hallucinations; formulaic style; shallow reasoning	115	122	**6**
Infrastructure & Access	Bandwidth, devices, premium features	43	47	**4**
Training, Policy & AI Literacy	PD, clear policies, norm-setting	107	119	**5**
Accessibility of Data Governance	Privacy, Integrity interface: watermarking, provenance	—	—	**—**
Industry & Workplace Readiness	Task reallocation; AI literacy for work	25	26	2

Methods validation and trustworthiness

Following Lincoln and Guba's [27] framework, the trustworthiness of the synthesis was evaluated through the criteria of credibility, dependability, and confirmability to strengthen transparency, reliability, and interpretive validity [27]; [23]; [28].

Thematic Findings: Process and Results

The thematic synthesis followed a structured multi-stage analytic procedure designed to identify recurring conceptual patterns across the coded dataset [29]; [23]; [24]. The same prompts were used in both Copilot and ChatGPT at each level of the analysis.

Meta-Themes Identified in the Qualitative Meta-Synthesis

The ChatGPT GenAI identified 12 meta-themes, whereas the Copilot GenAI distinguished 13 recurring meta-themes that describe how GenAI technologies are used and experienced in higher education environments.

Resulting Conceptual Framework

After completing the data analysis and retrieving the meta-themes for each GenAI tool, each output was uploaded to the corresponding tool to produce a conceptual framework. Although the datasets were extracted from the same qualitative studies, the two tools produced different themes and conceptual frameworks.

ChatGPT:

Fig. 1 illustrates the interaction between ChatGPT, other large language models, and higher education systems. The model identifies two primary dimensions: educational functions and educational risks. Educational functions include learning support, writing and language development, and research assistance, which enable students to access explanations, develop academic writing, and explore research ideas. In parallel, educational risks include cognitive dependence, ethical concerns related to AI misuse, and reliability issues associated with biased or inaccurate AI-generated outputs. These factors collectively influence learning processes and shape the student learning experience, including engagement, productivity, and learner autonomy. The framework further demonstrates how these dynamics contribute to broader pedagogical transformation, such as assessment redesign and AI-integrated teaching practices, and ultimately inform institutional governance responses, including AI literacy initiatives, policy development, and responsible AI use in higher education.

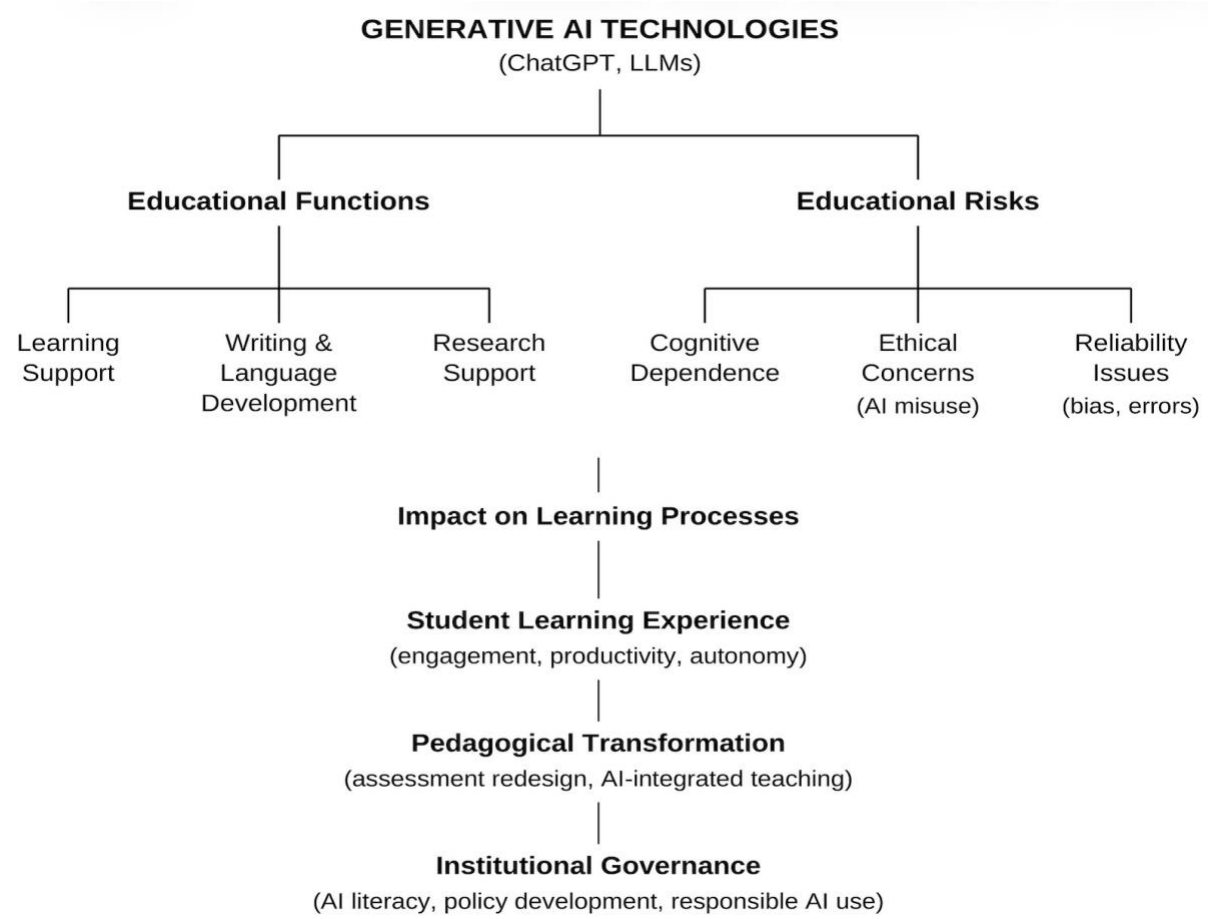


Fig. 1. Unified Conceptual Framework of GenAI Integration in Higher Education: Integrated Socio-Technical Model of Gen AI in Higher Education created by ChatGPT (ChatGPT-generated).

Conceptual Framework: GenAI in Higher Education

The conceptual framework presented in Fig. 1 was produced by ChatGPT after uploading the ChatGPT Excel workbook. It synthesizes the mixed role of generative artificial intelligence (AI) in higher education, highlighting its pedagogical values and associated risks, and tracing its implications for learning processes, student experience, instructional design, and institutional governance. The framework reflects a structured abstraction of recurrent empirical patterns and theoretical insights across the studies under review, examining tools such as ChatGPT and other large language models (LLMs).

Mixed-Role: Educational Functions and Educational Risks

At the center of the framework is the divergence between educational functions and educational risks, reflecting the two dominant narratives in AI in higher education. This dual structure reflects a growing consensus that GenAI cannot be understood exclusively as a pedagogical modernization, but rather as a sociotechnical phenomenon characterized by both empowering and limiting effects.

Educational Functions

The first dimension, educational functions, examines how GenAI enriches academic tasks and supports learning. Three subdomains emerged consistently across studies.

First, learning support reflects AI's role as an on-demand cognitive platform. GenAI tools provide explanations, clarify complex concepts, and simulate tutoring interactions, thereby supporting individualized and self-paced learning.

Second, writing and language development are among the most prominent applications of GenAI. The literature repeatedly highlights AI's capacity to assist with drafting, paraphrasing, grammar correction, and language refinement, benefiting predominantly multilingual learners and those developing academic writing proficiency.

Third, research support encompasses the use of AI for idea generation, summarization of scholarly materials, and preliminary literature exploration. Unlike writing support, this function contributes to early-stage knowledge production processes, facilitating the organization, development and synthesis of information.

Educational Risks

In contrast, the second dimension, educational risks, raises doubts about the unintentional effects of AI integration. Three primary categories were detected.

First, cognitive dependence refers to the tendency to over-reliance on AI systems, leading to reduced critical thinking, diminished engagement in problem-solving, and superficial learning. This risk is extremely significant because it directly affects the intellectual abilities underlying learning rather than only external behaviours.

Second, ethical concerns such as academic integrity, plagiarism, authorship ambiguity, and inappropriate reliance on AI-generated content emerged. These concerns are a constant focus in institutional and policy-oriented literature, highlighting tensions between technological capability and academic norms.

Third, reliability issues involve the technical limitations of GenAI, including hallucinations (i.e., the creation of illusory information), factual inaccuracies, and biases. Unlike ethical concerns, which arise from the user's behaviour, reliability issues are intrinsic to the technology itself and raise questions about trustworthiness and validity.

Collectively, these risks embrace cognitive, ethical, and technical dimensions, illustrating the multidimensional challenges of AI adoption in educational contexts.

From Technology to Learning: Impact on Learning Processes

The framework influences learning processes as a first-order outcome of the interaction between educational functions and risks. This level represents a critical analytical connection, highlighting how AI redesigns students' engagement with content, knowledge construction, and learning strategies.

The model underlines that not all risks operate equally at this level. Cognitive dependence is explicitly linked to learning processes because of its direct influence on students' reasoning ability, whereas ethical and reliability concerns have indirect effects.

Student Learning Experience

The framework identifies the student learning experience as an essential domain, encompassing engagement, productivity, and autonomy. These three dimensions represent the lived, experiential consequences of AI integration from the learner's perspective. Engagement refers to students' motivation and active participation in their learning journey. This may be enhanced by using GenAI as interactive support and reducing passive reliance. Productivity emphasizes improvements in efficiency, task completion, and academic output, which are frequently cited as primary benefits of AI tools. Autonomy considers students' ability for self-directed learning, which may be concurrently reinforced (through scaffolding) and weakened (through overdependence) by AI use.

Pedagogical Transformation

Beyond individual learning experiences, the framework illustrates how GenAI accelerates pedagogical transformation, reflecting a shift from progressive tool adoption toward more fundamental changes in teaching practices. First, assessment redesign emerges as a necessary response to the challenges posed by AI-generated content, urging educators to return to traditional evaluation methods. Second, AI-integrated teaching involves purposefully incorporating AI tools into instructional strategies, moving toward more adaptive, technology-enhanced learning environments. This transformation signifies a reconfiguration of the pedagogical model, in which AI is not only a supplementary tool but an active component of instructional design.

Institutional Governance

The framework situates these changes within the institutional governance perspective, emphasizing the general responses required to manage AI integration. Three key integration priorities are identified: 1) AI literacy requires both students and educators to develop the competencies needed to use AI technologies effectively and responsibly. 2) Policy development determines the necessary parameters for academic integrity, acceptable use, and assessment practices. 3) Adopting ethical frameworks and institutional standards demonstrates responsible AI use and confirms that it aligns with educational values, underscoring that the repercussions of GenAI extend beyond classrooms and require coordinated institutional strategies.

Integrative Perspective

This framework presents a multi-level, process-oriented model that details the influence of GenAI from technological capabilities and risks through to individual, pedagogical, and institutional consequences. It conceptualizes AI in higher education as a dynamic sociotechnical system in which benefits and risks coexist and interact, shaping both immediate learning processes and broader educational transformations. By integrating micro (student experience), meso (pedagogy), and macro (governance) perspectives, the framework proposes an inclusive understanding of the evolving role of GenAI in higher education. It also provides the foundation for future empirical research and policy development, as it is grounded in a meta-synthesis. The aim will be to leverage the identified points to maximize educational benefits while mitigating associated risks.

Human Validation of AI-Assisted Coding

While GenAI tools (ChatGPT and Copilot) were used to generate initial codes, thematic categories, and conceptual frameworks from the corpus of 154 studies, this process was conducted under continuous human supervision rather than as an autonomous analysis by AI tools. This judgment procedure was

applied independently and in parallel to each of the two workbooks; for every source article in both the Coding-ChatGPT and Coding-Copilot datasets, the text was independently read and coded according to the same thematic synthesis procedure [24] before comparing this independent coding against the corresponding AI-generated codes produced by that tool. Where human coding and a given tool's AI-generated coding converged, the code was retained. Where they diverged, the disagreement was adjudicated by returning to the original source excerpt to determine which coding, human, AI, or a revised synthesis of both, most faithfully represented the author's or participants' language and meaning. This process was carried out separately for the ChatGPT and Copilot outputs to preserve the integrity of each tool's independent dataset and to allow the resulting differences between the two conceptual frameworks (Figs. 1 and 2) to reflect genuine variation between the tools rather than cross-contamination introduced during validation. All AI-generated outputs, from the article collection, including evaluating them from inclusion/ exclusion, excerpts retrieval, individual codes, axial groupings, and the final thematic categories and conceptual frameworks for both tools, were reviewed against this standard; outputs found to be inaccurate, unsupported by the source text, or inconsistent were corrected or deleted prior to inclusion in the final synthesis. This dual, tool-specific adjudication process functioned as an additional layer of credibility and confirmability checking, consistent with Lincoln and Guba's [27] criteria, and ensured that the resulting themes and frameworks for each tool reflected a human-verified synthesis of the literature rather than an unexamined AI output.

Copilot:

The conceptual framework (Fig. 2) was produced by Copilot after uploading the Copilot Excel workbook. It positions generative artificial intelligence (GenAI) within four interconnected domains: Learning, Pedagogy/ Governance, Operations, and Inclusion/Industry, reflecting the main areas in which higher education systems are undergoing rapid transformation. These categories reflect the themes discovered across the meta-synthesis, based on the empirical analysis of GenAI in the higher education context using the Copilot tool.

The central node, GenAI in Higher Education, highlights GenAI's role as an adaptive support mechanism that accelerates cognitive, administrative, and creative processes. Solid connections represent primary domain linkages, while dashed cross-links depict secondary but recurrent interactions.

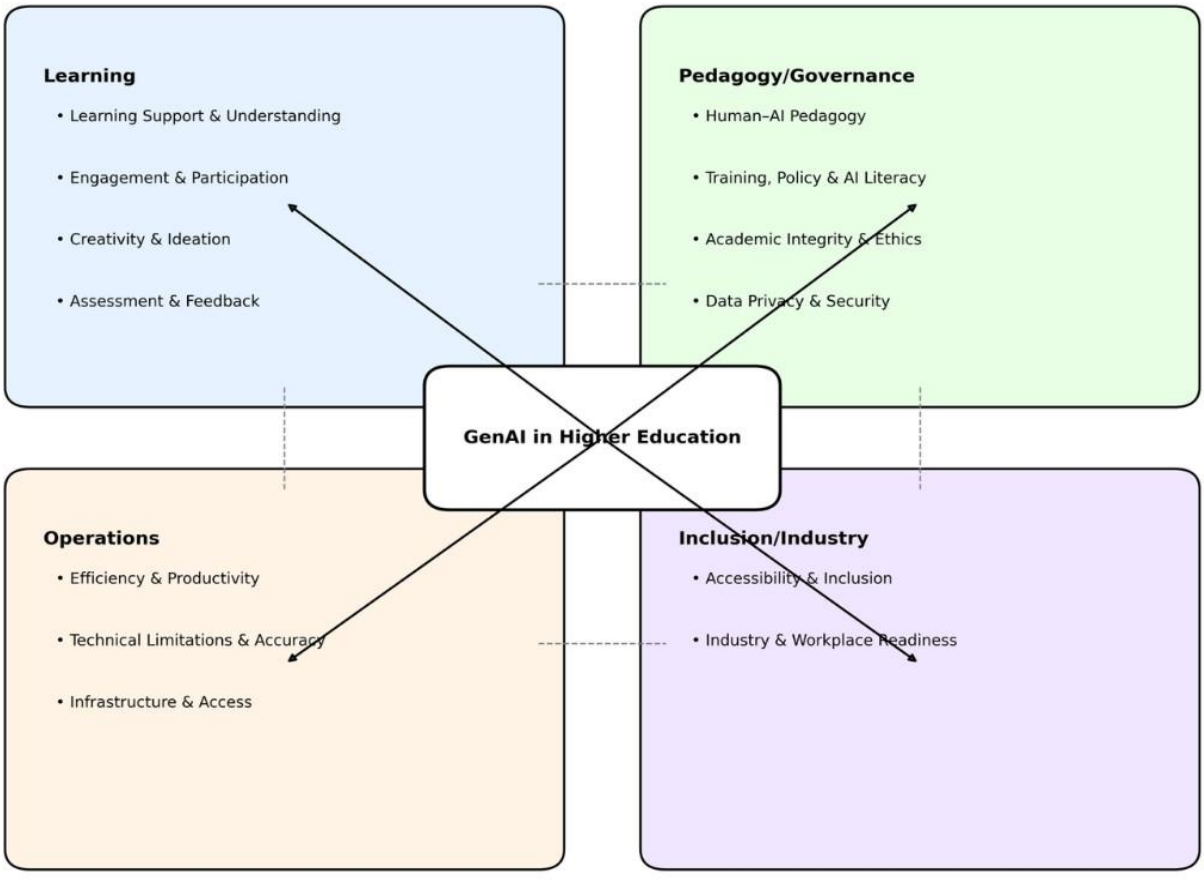


Fig. 2. System-level conceptual framework of GenAI in higher education created by Copilot

The Learning domain portrays how GenAI is reshaping student learning processes, particularly in areas such as personalization, engagement, creativity, and assessment. Multiple reviews indicate that GenAI tools can provide tailored learning opportunities, enhance academic support, and increase student involvement, thereby directly influencing their understanding and participation in meaningful ways. For example, research stresses GenAI's ability to deliver personalized learning experiences and improve student engagement through adaptive capabilities and responsive outputs. GenAI's role in shaping new forms of brainstorming and creative expression supports findings highlighting its flexibility across various educational tasks, while simultaneously raising questions about the integrity of assessments.

The Pedagogy/Governance domain acknowledges that instructional innovation must be anchored in responsible and ethical structures. Governance concerns, including privacy, data protection, fairness, and transparency for the responsible use of GenAI in education are highlighted. Analyses of GenAI governance models in educational research underline the importance of establishing adaptable, multi-level governance frameworks capable of addressing emerging pedagogical and regulatory considerations at macro-, meso-, and micro-levels. These outcomes reinforce the need for well-defined policies, faculty training, and AI literacy in ethically grounded and pedagogically rigorous structures.

The Operations domain exposes the organizational and infrastructural capacity required for effective GenAI integration. It has significant potential to improve efficiency and productivity in higher education by reorganizing administrative processes and supporting decision-making. For instance, GenAI is frequently discussed for its potential to reduce costs, enhance administrative workflows, and optimize service delivery when properly implemented. However, there is a tendency to understate its technical challenges, including limitations in accuracy, system reliability, data security, and the need for rigorous technological infrastructure to support safe use.

Finally, the Inclusion/Industry domain positions GenAI within the broader social and economic landscape, recognizing its implications for equity, accessibility, and workforce preparation. GenAI's predicted societal impact accentuates its potential to support more inclusive educational environments, but it can also aggravate existing inequalities if not implemented equitably. The rapid transformation of workforce expectations requires higher education institutions to prepare students for emerging skills involving AI ethics, AI literacy, and human-AI partnership.

These four domains explain the complex, interdependent ways GenAI is reshaping higher education. The framework highlights how changes in governance organizations shape pedagogical approaches; how operational infrastructures influence possibilities for equitable learning; how industry demands affect curriculum and workforce readiness; and how student learning needs force institutional adaptation. It offers a unified, multidimensional perception for understanding and guiding the responsible, transformative integration of GenAI within higher education.

Divergence Between the ChatGPT and Copilot Frameworks

Although both frameworks were derived from the same 154 article corpus and validated using an identical human-adjudication procedure, the ChatGPT and Copilot syntheses converged on different numbers of themes (12 and 13, respectively) and organized them according to different structural logics. This divergence is itself an informative finding rather than a limitation to be minimized. The ChatGPT framework is organized causally, tracing a pathway from educational functions and risks, through learning processes and student experience, to pedagogical and institutional consequences. The Copilot framework, by contrast, is organized categorically, grouping themes into four equal, interconnected domains (Learning, Pedagogy/Governance, Operations, and Inclusion/Industry) without positioning a directional sequence between them. This suggests that the two tools did not simply disagree about which themes were present in the excerpts but applied different underlying organizing principles to conceptually adjacent codes: one prioritizing process and causality, the other prioritizing domain classification. For example, codes related to academic integrity and data privacy were grouped by ChatGPT under a higher-ranking "Educational Risks" category linked causally to learning outcomes, whereas Copilot situated the same underlying codes within a "Pedagogy/Governance" domain alongside training and policy codes, emphasizing institutional responsibility rather than potential risk. This pattern indicates that GenAI-assisted thematic synthesis is sensitive not only to the evidence provided but to each tool's default synthesis exploration, underscoring the value of the human validation and adjudication procedure (Section 10.b) after independently verifying against source excerpts. Thus, relying on a single GenAI tool's output could mistake one tool's organizing logic for an objective feature of the underlying evidence. Future replications using additional GenAI or more advanced or paid tools would help clarify whether this divergence reflects a unique behaviour of these two specific systems or a more general property of GenAI-assisted qualitative synthesis.

Conclusions

After applying various criteria to strengthen conceptions, 154 articles were identified for detailed conceptual analysis. The figures in this paper are tentative and emergent, subject to further refinement through additional studies. It is, however, important to note the high frequency of occurrence of related themes that have thus far emerged from this study. Taken together, the five domains identified across both frameworks: learning, pedagogy, governance, operations, and inclusion/industry, offer higher education instructors and institutions a starting point for structuring their own AI adoption efforts, rather than approaching GenAI integration as a single, undifferentiated decision. For instructors, the findings suggest that instructional value is concentrated in well-scaffolded uses of GenAI (e.g., explanation, drafting support, research assistance) paired explicitly with strategies

that counter cognitive dependence, such as requiring students to critique or verify AI-generated content rather than accept it uncritically. For administrators and institutional leaders, the consistent appearance of governance-related themes (policy development, data privacy, academic integrity) across both frameworks indicates that institutional policy cannot be an afterthought to classroom-level adoption; clear, proactively communicated norms for acceptable AI use appear to be a precondition for, rather than a response to, effective pedagogical integration. For those responsible for operations and infrastructure, the recurrence of themes related to efficiency, access, and technical limitations suggests that equitable release depends as much on infrastructure investment and reliability testing as on pedagogical design. Finally, the persistence of inclusion and industry-readiness themes across both syntheses points to a need for institutions to evaluate GenAI adoption not only by classroom outcomes but by its effects on equitable access and workforce preparedness. These implications should be read as directional, given the conceptual and evolving nature of the AI in higher education; they are intended to guide further empirical validation and institutional discussion rather than to serve as a finalized implementation checklist.

Conflict-of-interest

The authors declare no financial conflicts of interest related to the research, authorship, or publication of this paper.

References

[1] H. Turillo. (Sep, 19). The Future of AI and Metaverse in Education. Available: https://businessabc.net/the-future-of-ai-and-metaverse-ineducation.
[2] S. Elayyan, "The future of education according to the fourth industrial revolution," Journal of Educational Technology and Online Learning, vol. 4, (1), 2021. . DOI: https://doi.org/10.31681/jetol.737193.
[3] H. Abbass, "Editorial: What is Artificial Intelligence?" TAI, vol. 2, (2), pp. 94–95, 2021. . DOI: https://doi.og/10.1109/TAI.2021.3096243.
[4] L. Chen, P. Chen and Z. Lin, "Artificial intelligence in education: A review," Ieee Access, vol. 8, pp. 75264–75278, 2020. Available: https://ieeexplore.ieee.org/stamp/stamp.jsp?arnumber=9069875. DOI: https://doi.org/10.1109/ACCESS.2020.2988510.
[5] C. DiCarlo, Building a God: The Ethics of Artificial Intelligence and the Race to Control It. Essex, CT: Prometheus Books, 2025.
[6] B. Mahesh, "Machine learning algorithms: A review," International Journal of Science and Research (IJSR).[Internet], vol. 9, (1), pp. 381–386, 2020. . DOI: https://doi.org/10.21275/ART20203995.
[7] W. Dong, D. Pan and S. Kim, "Exploring the integration of IoT and Generative AI in English language education: Smart tools for personalized learning experiences," Journal of Computational Science, vol. 82, (1), pp. 102397, 2024. DOI: https://doi.org/10.1016/j.jocs.2024.102397.
[8] A. Kumar et al, "Transforming Education Through Iot And AI Opportunities And Challenges," Educational Administration: Theory and Practice, vol. 30, (5), pp. 11610–11622, 2024. . DOI: https://doi.org/10.53555/kuey.v30i5.4982.
[9] A. Barr, Artificial Intelligence: Cognition as Computation. Stanford University, 1982.
[10] R. Luckin and W. Holmes, "Intelligence unleashed: An argument for AI in education," UCL Knowledge Lab, 2016Available: https://discovery.ucl.ac.uk/id/eprint/1475756/.
[11] J. Zhao et al, "Cognitive psychology-based artificial intelligence review," Frontiers in Neuroscience, vol. 16, pp. 1–9, 2022. . DOI: https://doi.org/10.3389/fnins.2022.1024316.
[12] D. N. Cassenti, V. D. Veksler and F. E. Ritter, "Editor's Review and Introduction: Cognition-Inspired Artificial Intelligence," Topics in Cognitive Science, vol. 14, (4), pp. 652–664, 2022. DOI: https://doi.org/10.1111/tops.1262.
[13] D. Hassabis et al, "Neuroscience-inspired artificial intelligence," Neuron, vol. 95, (2), pp. 245–258, 2017. DOI: https://doiorg.roxy.nipissingu.ca/10.1016/j.neuron.2017.06.011.
[14] R. S. Rosenberg, "Artificial intelligence and linguistics: A brief history of a one-way relationship," in Annual Meeting of the Berkeley Linguistics Society, 1975, pp. 379–392.
[15] A. Gocen and F. Aydemir, "Artificial Intelligence in Education and Schools," REM: Research on Education and Media, vol. 12, (1), pp. 13–21, 2020. . DOI: https://doi.org/10.2478/rem-2020-0003.
[16] X. Jia, L. Ren and J. Cai, "Clinical implementation of AI technologies will require interpretable AI models," Med. Phys., vol. 47, (1), pp. 1–4, 2020. . DOI: https://doi.org/10.1002/mp.13891.
[17] W. Wang and K. Siau, "Artificial intelligence, machine learning, automation, robotics, future of work and future of humanity: A review and research agenda," Journal of Database Management (JDM), vol. 30, (1), pp. 61–79, 2019. . DOI: https://doi.org/10.4018/JDM.2019010104.
[18] H. Snyder, "Literature review as a research methodology: An overview and guidelines," Journal of Business Research, vol. 104, pp. 333–339, 2019. . DOI: https://doi.org/10.1016/j.jbusres.2019.07.039.
[19] I. Celik et al, "The Promises and Challenges of Artificial Intelligence for Teachers: A Systematic Review of Research," TechTrends, vol. 66, (4), pp. 616–630, 2022. . DOI: https://doi.org/10.1007/s11528-022-00715-y.
[20] K. Charmaz, Constructing Grounded Theory. (2nd ed.) London: Sage Publications, 2014.
[21] E. G. Guba and Y. S. Lincoln, "Competing paradigms in qualitative research," in Handbook of Qualitative Research Anonymous Thousand Oaks: Sage: 1994, pp. 105–117.
[22] D. Walsh and S. Downe, "Meta-synthesis method for qualitative research: A literature review," J. Adv. Nurs., vol. 50, (2), pp. 204–211, 2005. DOI: https://doi.org/10.1111/j.1365-2648.2005.03380.x.
[23] M. Sandelowski, J. Barroso and C. I. Voils, "Using qualitative metasummary to synthesize qualitative and quantitative descriptive findings," Res. Nurs. Health, vol. 30, (1), pp. 99–111, 2007. . DOI: https://doi.org/10.1002/nur.20176.
[24] J. Thomas and A. Harden, "Methods for the thematic synthesis of qualitative research in systematic reviews," BMC Med. Res. Methodol., vol. 8, (1), pp. 45, 2008. . DOI: https://doi.org/10.1186/1471-2288-8-45.
[25] B. L. Paterson et al, Meta-Study of Qualitative Health Research: A Practical Guide to Meta-Analysis and Meta-Synthesis. 2001.
[26] A. Tong et al, "Enhancing transparency in reporting the synthesis of qualitative research: ENTREQ," BMC Med. Res. Methodol., vol. 12, (1), pp. 181, 2012. . DOI: https://doi.org/10.1186/1471-2288-12-181.
[27] Y. S. Lincoln and E. G. Guba, Naturalistic Inquiry. 1985.
[28] S. P. S. Senanayake and D. N. N. J. N. Dean, "Factors Influencing Trustworthiness of Qualitative Research: A Systematic Literature Review," International Journal of Research and Scientific Innovation, vol. 12, (12), pp. 297–319, 2025. DOI: https://doi.org/10.51244/IJRSI.2025.12120028.
[29] J. Hewitt-Taylor, "Use of constant comparative analysis in qualitative research," Nursing Standard, vol. 15, (42), pp. 39–42, 2001. . DOI: https://doi.org/10.7748/ns2001.07.15.42.39.c3052.