International Journal of Education, Leadership, Artificial Intelligence, Computing, Business, Life Sciences, and Society - https://journalviral.org
DOI: https://doi.org/10.65222/VIRAL.2026.7.41.61

From ChatGPT to Claude: A Literature Review of Large Language Models in Education

Bogdan Costache
Bucharest University of Economic Studies
Costin Petcu
"Carol Davila"University of Medicine and Pharmacy, Bucharest
Vladimir Aurelian Enachescu
Bucharest University of Economic Studies

ABSTRACT: *The rapid diffusion of large language models has fundamentally altered discussions surrounding artificial intelligence in education and generated significant scholarly interest in their implications for teaching, learning, assessment, and institutional governance. Systems such as ChatGPT and Claude have moved generative artificial intelligence from experimental applications to mainstream educational practice, creating new opportunities and challenges for universities, educators, and policymakers. Despite the growing volume of research, the literature remains fragmented across technological, pedagogical, and organizational perspectives, limiting the development of integrated theoretical understandings of how large language models reshape educational systems.*

This study addresses this fragmentation through a structured integrative literature review based on systematic searches of major academic databases and a theory-driven synthesis of existing research. The review examines the evolution of scholarly work on large language models in education, identifies dominant themes, and explores the theoretical foundations that underpin current debates. The findings reveal a rapid transition from studies centered on technological capabilities toward broader concerns involving human-AI collaboration, educational leadership, institutional transformation, and AI governance. At the same time, significant gaps remain regarding long-term organizational implications, regulatory frameworks, and the development of sustainable models of human-machine collaboration in educational environments.

The paper contributes to the literature by providing an integrative conceptual framework that connects fragmented research streams and proposes a future research agenda for the study of large language models in education. The analysis also offers practical insights for educational institutions seeking to navigate the opportunities and risks associated with generative artificial intelligence and highlights directions for future empirical and theoretical investigations.

KEYWORDS: Large Language Models; ChatGPT; Claude; Generative Artificial Intelligence; Artificial Intelligence in Education; Human-AI Collaboration; Educational Leadership; AI Governance; Digital Transformation; Higher Education.

1. Introduction

Artificial intelligence has become one of the most influential technological developments shaping contemporary education. During the last decade, advances in machine learning, natural language processing, and computational power have progressively expanded the capabilities of intelligent systems, enabling them to participate in increasingly sophisticated cognitive tasks that were previously considered uniquely human. The emergence of generative artificial intelligence, and particularly large language models (LLMs), has accelerated this transformation and introduced new possibilities for teaching, learning,

DOI: https://doi.org/10.65222/VIRAL.2026.7.41.61

assessment, and educational management. Unlike earlier educational technologies that primarily supported information delivery or automated administrative tasks, large language models can generate text, summarize complex information, engage in dialogue, explain concepts, and assist with problem-solving in ways that closely resemble human communication.

Educational institutions have responded rapidly to these developments. Universities across the world are experimenting with generative artificial intelligence for curriculum design, student support, content creation, and administrative decision-making. Faculty members increasingly rely on conversational AI systems to prepare educational materials, while students use these tools for tutoring, writing assistance, and information retrieval. Such developments have triggered both enthusiasm and concern. Advocates argue that large language models can increase educational accessibility, support personalized learning, and reduce administrative burdens, whereas critics point to risks related to academic integrity, algorithmic bias, misinformation, and excessive dependency on artificial intelligence (Kasneci et al., 2023; Tlili et al., 2023).

The speed of adoption has also created significant theoretical and institutional challenges. Educational systems traditionally evolve gradually, often requiring years to integrate new technologies into pedagogical practices and governance arrangements. By contrast, generative artificial intelligence has diffused at an unprecedented pace, creating pressures on universities to respond to technologies whose long-term consequences remain uncertain. This rapid transformation raises fundamental questions regarding the future relationship between humans and intelligent systems in educational environments and the role of educational institutions in shaping responsible and sustainable forms of human-AI collaboration.

1.1 The Emergence of Large Language Models

The emergence of large language models represents the latest stage in the broader evolution of artificial intelligence. Earlier generations of educational technologies relied heavily on rule-based systems and relatively narrow applications of machine learning. Computer-assisted instruction systems and intelligent tutoring environments were capable of supporting specific learning activities, but their ability to understand language, generate content, and adapt to complex educational contexts remained limited (Luckin et al., 2016; Zawacki-Richter et al., 2019).

The introduction of transformer architectures fundamentally altered this landscape. The publication of the transformer model by Vaswani et al. (2017) demonstrated that attention mechanisms could dramatically improve the ability of artificial intelligence systems to process and generate natural language. Subsequent developments, including large-scale pretraining and foundation models, created systems capable of performing a wide variety of tasks with remarkable fluency and contextual awareness (Bommasani et al., 2021).

The release of ChatGPT by OpenAI in November 2022 represented a turning point in public awareness and institutional adoption of generative artificial intelligence. Within a matter of months, ChatGPT became one of the fastest-growing digital platforms in history and

DOI: https://doi.org/10.65222/VIRAL.2026.7.41.61

generated an extraordinary wave of interest across educational communities (Dwivedi et al., 2023). Researchers immediately began examining its implications for teaching, assessment, and student learning, leading to an unprecedented increase in scholarly publications concerning artificial intelligence in education (Lo, 2023; Baidoo-Anu & Ansah, 2023).

The subsequent emergence of Claude, developed by Anthropic, further expanded the landscape of educational large language models. Claude distinguished itself through larger context windows, strong document analysis capabilities, and a development philosophy centered on constitutional AI and safety mechanisms (Bai et al., 2022). These characteristics made it particularly attractive for educational settings requiring the analysis of extensive documents, research synthesis, and complex reasoning tasks. The competition among increasingly sophisticated large language models has contributed to a rapidly evolving ecosystem in which educational institutions face a growing number of opportunities and challenges.

The societal significance of these technologies extends far beyond individual learning activities. Large language models increasingly influence knowledge production, information access, and institutional decision-making processes. Universities are beginning to reconsider assessment methods, academic integrity policies, digital strategies, and even the competencies that graduates require in future labor markets (Cotton et al., 2023; Rudolph et al., 2023). Educational leaders and policymakers are therefore confronted with a complex task: embracing the opportunities presented by generative artificial intelligence while simultaneously addressing concerns related to ethics, governance, transparency, and human agency.

The emergence of large language models should therefore not be interpreted merely as another stage in educational technology development. Rather, it represents a broader transformation in the relationship between knowledge, intelligence, and learning. The growing integration of systems such as ChatGPT and Claude into educational environments suggests that higher education may be entering a new phase in which human and artificial intelligence increasingly coexist and collaborate within shared cognitive and organizational ecosystems.

1.2 Research Problem

The remarkable growth of research on large language models in education has generated an increasingly diverse body of literature that spans multiple disciplines, including educational technology, computer science, management, information systems, psychology, and public policy. This rapid expansion has produced important insights into the capabilities and limitations of generative artificial intelligence, yet it has also resulted in significant conceptual fragmentation. Existing studies often examine isolated aspects of large language models without establishing broader theoretical connections among technological developments, pedagogical practices, and institutional transformation.

A substantial proportion of educational research has concentrated on immediate classroom applications, particularly the use of ChatGPT and similar systems for learning support,

DOI: https://doi.org/10.65222/VIRAL.2026.7.41.61

writing assistance, assessment, and personalized tutoring (Kasneci et al., 2023; Lo, 2023; Tlili et al., 2023). While these studies provide valuable evidence regarding the pedagogical potential of large language models, they frequently adopt a technology-centric perspective that pays limited attention to wider organizational and governance implications.

Conversely, technological studies primarily focus on model architectures, reasoning capabilities, prompt engineering, and performance benchmarking (OpenAI, 2023; Bommasani et al., 2021). Although these contributions are essential for understanding the evolution of large language models, they often remain detached from educational theories and institutional realities. The result is a growing divide between research that examines what these systems can technically accomplish and research that investigates how educational institutions may meaningfully integrate them into their practices.

A further limitation concerns the relative absence of studies addressing educational leadership, institutional governance, and organizational transformation. The implementation of large language models raises important questions regarding accountability, digital strategy, policy development, quality assurance, and the future role of educational institutions in knowledge societies. However, these dimensions remain comparatively underdeveloped in current scholarship (Williamson & Eynon, 2020; Selwyn, 2019).

The fragmentation of the literature also reflects a broader theoretical gap. Research on large language models in education rarely integrates perspectives from socio-technical systems theory, complexity theory, organizational information processing theory, and human-AI collaboration research. Consequently, there is still limited understanding of how generative artificial intelligence may reshape educational systems as complex organizational and institutional arrangements rather than merely influencing individual learning activities.

This fragmentation creates both theoretical and practical challenges. From a scholarly perspective, the absence of integrative frameworks hinders cumulative knowledge development and limits the capacity to identify coherent research agendas. From an institutional perspective, universities and policymakers face increasing pressure to make strategic decisions regarding artificial intelligence adoption without a sufficiently consolidated evidence base capable of informing governance and long-term planning.

The present study addresses this problem by conducting an integrative literature review that synthesizes fragmented research streams and develops a conceptual understanding of large language models as transformative educational technologies embedded within broader organizational and institutional contexts.

1.3 Research Questions

The study is guided by four interrelated research questions designed to capture both the evolution of the literature and the emerging challenges associated with large language models in education:

RQ1. How has the literature on large language models in education evolved?

DOI: https://doi.org/10.65222/VIRAL.2026.7.41.61

RQ2. What are the principal themes and theoretical foundations of this literature?

RQ3. How are ChatGPT and Claude transforming educational processes and institutional governance?

RQ4. What research gaps and future directions emerge from the existing literature?

These questions recognize that large language models should not be viewed solely as technological artifacts. They are increasingly becoming institutional actors that influence teaching practices, knowledge production, decision-making processes, and organizational structures within higher education.

1.4 Contributions

This study makes several contributions to the emerging literature on generative artificial intelligence in education.

First, it provides a systematic synthesis of a rapidly expanding and fragmented body of scholarship that spans multiple disciplinary domains. By bringing together pedagogical, technological, organizational, and governance perspectives, the review offers a more comprehensive understanding of how large language models are reshaping educational systems.

Second, the study develops an integrative conceptual framework that connects antecedents, technological capabilities, educational mechanisms, and institutional outcomes associated with large language model adoption. This framework contributes to theory building by positioning generative artificial intelligence within broader processes of educational transformation and organizational adaptation.

Third, the review identifies emerging research themes that extend beyond classroom applications and include issues related to human-AI collaboration, educational leadership, digital transformation, institutional resilience, and AI governance. These themes suggest that the impact of large language models cannot be adequately understood through pedagogical perspectives alone.

Fourth, the study proposes a future research agenda intended to support cumulative knowledge development in this rapidly evolving field. The agenda highlights areas requiring further conceptual refinement and empirical investigation, including the role of trust, ethical governance, organizational capabilities, and long-term institutional consequences associated with artificial intelligence adoption.

Finally, the paper contributes to ongoing debates concerning the future of higher education by offering a theoretically informed perspective on the relationship between large language models and educational transformation. As universities increasingly navigate the opportunities and uncertainties generated by generative artificial intelligence, there is a

DOI: https://doi.org/10.65222/VIRAL.2026.7.41.61

growing need for integrative frameworks capable of informing both scholarly inquiry and institutional decision-making.

2. Literature Review

2.1 Evolution of Artificial Intelligence in Education

The integration of artificial intelligence into education did not begin with generative models or conversational agents. Rather, it represents the latest phase of a longer process of technological transformation that has progressively altered how knowledge is produced, disseminated, and evaluated within educational systems. Understanding the emergence of large language models requires situating them within this broader historical trajectory.

The earliest forms of educational technology were largely characterized by one-way information delivery systems that digitized existing teaching materials without fundamentally changing pedagogical practices. Educational Technology 1.0 focused primarily on computer-assisted instruction and electronic content repositories, emphasizing efficiency and access to information rather than personalization or intelligent adaptation (Selwyn, 2019).

The subsequent development of Computer-Assisted Instruction (CAI) introduced greater levels of interaction and automation. These systems employed predefined rules and programmed responses to guide learners through structured activities and exercises. Although pedagogically valuable, CAI remained constrained by limited adaptability and an inability to understand complex learner behaviors (Luckin et al., 2016).

During the 1980s and 1990s, advances in artificial intelligence led to the development of Intelligent Tutoring Systems (ITS). These systems represented a significant conceptual shift because they attempted to model student knowledge, identify misconceptions, and provide individualized learning pathways. Intelligent tutoring systems introduced adaptive learning principles that would later become central to contemporary educational technologies (Woolf, 2010). Nevertheless, their implementation often required highly specialized domain models and substantial development resources, limiting widespread adoption.

The emergence of big data and advances in computational analytics subsequently facilitated the growth of Learning Analytics. Educational institutions increasingly collected and analyzed large volumes of learner data to predict performance, identify at-risk students, and support evidence-based decision-making. Learning analytics expanded the role of artificial intelligence from individualized tutoring toward institutional decision support and educational governance (Siemens & Long, 2011).

The release of large-scale generative models fundamentally transformed this landscape. Generative Artificial Intelligence introduced systems capable of producing novel content, engaging in dialogue, and responding flexibly to diverse educational tasks. Unlike previous technologies that were designed for specific pedagogical functions, generative AI systems are general-purpose technologies that can support a wide range of educational activities, from tutoring and assessment to curriculum design and administrative processes (Kasneci et al., 2023).

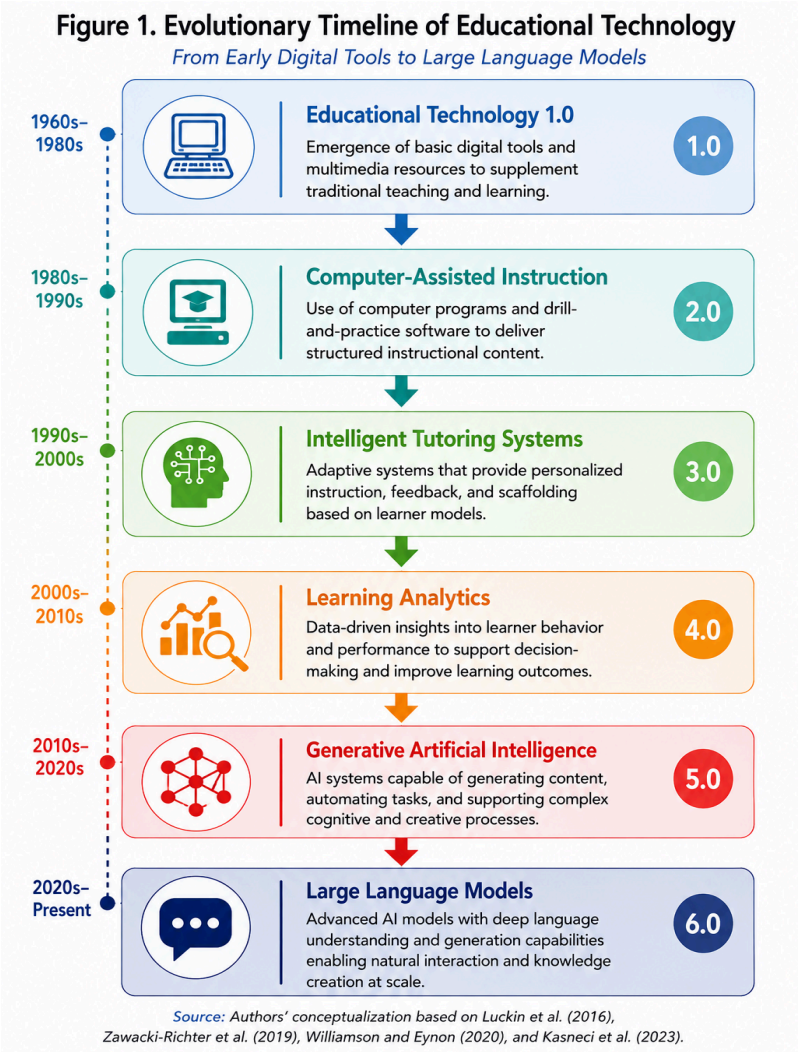
Large Language Models represent the most recent and arguably most disruptive stage of this evolution. Built upon transformer architectures and trained on massive datasets, these

DOI: https://doi.org/10.65222/VIRAL.2026.7.41.61

models exhibit unprecedented capabilities in natural language understanding, reasoning, and content generation (Bommasani et al., 2021). Systems such as ChatGPT and Claude have accelerated the diffusion of generative artificial intelligence by making advanced language technologies widely accessible to educators, students, and institutions. Their adoption has initiated a new phase in artificial intelligence in education characterized by conversational interaction, knowledge co-creation, and emerging forms of human-AI collaboration.

Importantly, each stage in this historical progression has expanded the role of technology within educational systems. Earlier technologies primarily supported information transmission, whereas contemporary large language models increasingly participate in cognitive and organizational processes that were previously considered inherently human. This evolution suggests that artificial intelligence is no longer merely an educational tool but is becoming an active participant in educational ecosystems and institutional transformation.

Figure 1. Evolution of Artificial Intelligence in Education



Source: *Authors' conceptualization based on Luckin et al. (2016), Zawacki-Richter et al. (2019), Williamson and Eynon (2020), and Kasneci et al. (2023).*

DOI: https://doi.org/10.65222/VIRAL.2026.7.41.61

The historical trajectory presented in Figure 1 also illustrates an important conceptual transition. Earlier educational technologies focused primarily on supporting instructional delivery and administrative efficiency. Contemporary large language models increasingly influence knowledge production, educational governance, and institutional decision-making. Consequently, the study of artificial intelligence in education requires analytical frameworks capable of explaining not only technological adoption but also broader organizational and societal transformations associated with human-AI collaboration.

The following sections therefore examine the nature of large language models as educational technologies and explore the theoretical perspectives that can support a deeper understanding of their transformative potential.

2.2 Understanding Large Language Models

Large language models represent a significant departure from previous generations of artificial intelligence systems. Rather than being designed to perform narrowly defined tasks, they belong to a class of foundation models trained on extremely large datasets and capable of adapting to a broad range of activities through natural language interaction (Bommasani et al., 2021). Their emergence has fundamentally altered how researchers conceptualize human-computer interaction, knowledge work, and educational technology.

The technological foundation of modern large language models lies in the transformer architecture proposed by Vaswani et al. (2017). Unlike earlier recurrent neural networks, transformer models employ self-attention mechanisms that enable them to process contextual relationships among words more effectively and to scale efficiently to massive datasets. This architectural innovation facilitated the development of increasingly sophisticated language models capable of generating coherent text, performing reasoning tasks, and responding to complex prompts.

The concept of foundation models further expanded the significance of large language models. Foundation models are trained on diverse and extensive corpora and can subsequently be adapted to numerous downstream applications through fine-tuning or prompt engineering (Bommasani et al., 2021). In educational settings, this flexibility allows a single model to support tutoring, assessment, content generation, translation, feedback provision, and administrative decision support.

Prompt engineering has consequently emerged as a critical dimension of human-AI interaction. Unlike traditional software systems that require explicit programming, large language models rely on carefully designed instructions that shape their outputs and determine the quality of generated responses. Effective interaction with these systems therefore requires new forms of digital literacy and communicative competence, often described as prompt literacy or AI literacy (Kasneci et al., 2023).

DOI: https://doi.org/10.65222/VIRAL.2026.7.41.61

Another defining characteristic of large language models is their conversational nature. Conversational AI systems enable users to engage in iterative dialogue, refine requests, and receive explanations tailored to specific contexts. This capability has considerable educational implications because it facilitates personalized learning experiences and creates opportunities for continuous learner support (Lo, 2023).

Recent developments have also introduced multimodal capabilities. Contemporary systems increasingly integrate text, images, audio, and documents into unified models capable of processing and generating information across different modalities. These developments considerably expand the educational applications of large language models by enabling richer forms of interaction and supporting increasingly complex learning activities.

The issue of reasoning and contextual understanding remains the subject of considerable debate. Although contemporary large language models demonstrate impressive performance across many tasks, scholars continue to disagree regarding the extent to which these systems genuinely reason or merely generate statistically plausible outputs based on learned patterns (Bubeck et al., 2023; Floridi & Chiriatti, 2020). From an educational perspective, this distinction is particularly important because it influences how these technologies should be integrated into teaching and learning processes and shapes expectations regarding their pedagogical role.

Consequently, large language models should not be viewed simply as advanced search engines or text generators. They represent a new class of socio-technical systems that mediate knowledge production, facilitate human-AI collaboration, and increasingly participate in educational and organizational processes.

2.3 ChatGPT and Claude as Educational Technologies

Among the rapidly growing ecosystem of large language models, ChatGPT and Claude have emerged as two of the most influential systems in educational contexts. Although both systems share common technological foundations and rely on transformer-based architectures, they differ in their design philosophies, capabilities, and potential applications within educational environments.

ChatGPT, developed by OpenAI, rapidly became the most widely adopted generative AI system in education due to its accessibility, conversational fluency, and broad functionality (OpenAI, 2023). Its capabilities include content generation, tutoring support, coding assistance, assessment feedback, and instructional design. The widespread availability of ChatGPT has led to extensive experimentation across universities and schools and has generated substantial scholarly interest concerning its pedagogical and organizational implications (Dwivedi et al., 2023).

Claude, developed by Anthropic, emerged somewhat later but has increasingly attracted attention due to its emphasis on safety, interpretability, and document analysis capabilities. The model's development is strongly influenced by the concept of Constitutional AI, which seeks to align system behavior with explicit normative principles and reduce harmful outputs

DOI: https://doi.org/10.65222/VIRAL.2026.7.41.61

(Bai et al., 2022). Claude's larger context windows and ability to process extensive documents make it particularly suitable for research activities, literature analysis, and complex educational tasks involving lengthy materials.

The comparison between ChatGPT and Claude should not be reduced to technical specifications alone. Their differences also reflect broader debates concerning responsible artificial intelligence, human oversight, and the governance of generative technologies in educational environments.

Table 1. Comparative Characteristics of ChatGPT and Claude in Educational Contexts

Dimension	ChatGPT	Claude
Context window	Large, continuously expanding depending on model version	Very large context windows designed for extensive document processing
Reasoning capabilities	Strong general reasoning and problem-solving performance	Strong analytical reasoning with emphasis on long-context understanding
Document analysis	Effective with uploaded files and structured materials	Particularly strong in processing lengthy documents and synthesis tasks
Educational applications	Tutoring, content generation, coding support, assessment feedback	Research assistance, document synthesis, analytical writing, policy analysis
Safety mechanisms	Reinforcement learning from human feedback and content filtering	Constitutional AI and explicit safety principles
Hallucination management	Ongoing improvements through model iterations and retrieval mechanisms	Strong emphasis on cautious responses and uncertainty acknowledgment
Ethical considerations	Extensive public debate concerning bias, privacy, and academic integrity	Greater emphasis on alignment, transparency, and responsible use
Institutional use cases	Teaching support, student assistance, curriculum development	Research support, governance analysis, academic administration

Source: *Authors' synthesis based on OpenAI (2023), Bai et al. (2022), Dwivedi et al. (2023), and Kasneci et al. (2023).*

DOI: https://doi.org/10.65222/VIRAL.2026.7.41.61

The educational significance of these systems lies not only in their individual functionalities but also in their broader implications for educational transformation. ChatGPT and Claude increasingly act as cognitive partners that participate in information processing, knowledge production, and decision support. Their growing integration into educational systems therefore requires analytical frameworks capable of explaining not merely technological adoption but also the evolving relationships among human expertise, institutional structures, and intelligent technologies.

2.4 Large Language Models and Educational Transformation

The growing integration of large language models into educational settings has generated changes that extend well beyond the adoption of another digital tool. Their diffusion is reshaping pedagogical practices, redefining the boundaries between teaching and learning, and challenging long-standing assumptions regarding knowledge creation and academic work. Unlike previous educational technologies that supported specific tasks or administrative functions, large language models increasingly participate in cognitive activities that have traditionally been performed by teachers and students.

One of the most visible transformations concerns teaching and learning processes. Conversational systems such as ChatGPT and Claude can provide explanations, generate examples, summarize complex concepts, and adapt responses to individual learner needs. These capabilities create opportunities for more personalized and flexible educational experiences. Students can engage in iterative dialogues, receive immediate feedback, and access forms of support that were previously constrained by time and resource limitations (Kasneci et al., 2023; Lo, 2023). At the same time, these developments raise important questions regarding dependency on artificial intelligence, the development of critical thinking, and the changing role of educators.

Curriculum design is also being reconsidered. The availability of systems capable of producing text, code, and analytical outputs challenges traditional assumptions regarding the knowledge and skills that graduates should possess. Educational institutions increasingly recognize that future professionals will require competencies associated with human-AI collaboration, critical evaluation of machine-generated content, and ethical reasoning rather than merely technical proficiency in information retrieval (Chiu, 2024; Rudolph et al., 2023).

Assessment practices have become a particularly contested area. Large language models have stimulated debates concerning plagiarism, authorship, and academic integrity because students can now generate essays, reports, and programming solutions with minimal effort. These developments have encouraged educators to reconsider the purposes and methods of assessment, leading to increased interest in authentic assessment, project-based evaluation, oral examinations, and process-oriented learning activities (Cotton et al., 2023; Tlili et al., 2023).

Feedback processes are undergoing significant changes as well. Large language models can provide immediate and personalized responses, suggest revisions, and support formative assessment. Several studies indicate that generative AI may improve feedback accessibility

DOI: https://doi.org/10.65222/VIRAL.2026.7.41.61

and responsiveness, particularly in large educational settings where individual instructor attention is limited (Mollick & Mollick, 2023). Nevertheless, concerns remain regarding the reliability, accuracy, and pedagogical appropriateness of automatically generated feedback.

Personalization constitutes another important dimension of educational transformation. The conversational and adaptive capabilities of large language models enable the creation of learning experiences tailored to individual preferences, prior knowledge, and learning pace. Such possibilities align with broader trends toward learner-centered education and adaptive learning environments (Zawacki-Richter et al., 2019). Yet they also raise questions concerning data governance, privacy, and the potential reinforcement of existing educational inequalities.

Beyond teaching and learning, large language models increasingly influence administrative processes and educational governance. Universities are experimenting with generative AI for student advising, communication, policy development, strategic planning, and institutional decision support. These applications suggest that large language models are becoming organizational technologies that shape information flows and institutional practices rather than merely classroom tools.

The transformative potential of these technologies therefore extends across multiple dimensions of educational systems. Their adoption requires not only technological readiness but also organizational adaptation, ethical reflection, and new forms of institutional governance capable of managing the opportunities and risks associated with increasingly intelligent educational environments.

2.5 Theoretical Foundations

The rapid emergence of large language models has created a need for theoretical perspectives capable of explaining their complex and multidimensional implications for education. No single theory is sufficient to capture the interactions among technological capabilities, human behavior, organizational processes, and institutional change. Consequently, this review adopts an interdisciplinary perspective that integrates several complementary theoretical frameworks.

Socio-Technical Systems Theory provides one of the most useful perspectives for understanding artificial intelligence in education. The theory proposes that organizational outcomes emerge through interactions between social and technical subsystems rather than through technological factors alone (Trist & Bamforth, 1951; Bostrom & Heinen, 1977). From this perspective, large language models should be viewed as components of broader educational ecosystems in which technologies, educators, students, and institutional structures continuously influence one another.

Activity Theory offers a second important lens. Originating in the work of Vygotsky and later developed by Engeström (1987), Activity Theory conceptualizes human action as mediated by tools and artifacts. Large language models can therefore be understood as mediating

DOI: https://doi.org/10.65222/VIRAL.2026.7.41.61

technologies that transform educational activities, redistribute cognitive tasks, and alter relationships between teachers, learners, and knowledge.

Complexity Theory provides additional insights into educational transformation. Universities increasingly operate as complex adaptive systems characterized by interdependence, nonlinearity, and continuous adaptation (Davis & Sumara, 2006). The introduction of large language models into such environments may generate unintended consequences and emergent forms of organizational change that cannot be fully understood through linear models of technological adoption.

Human-AI Collaboration Theory represents another essential perspective. Rather than framing artificial intelligence as a substitute for human capabilities, this theory emphasizes complementarities between human judgment and machine intelligence (Wilson & Daugherty, 2018; Jarrahi, 2018). The concept of collaborative intelligence is particularly relevant in educational settings where effective outcomes depend on the quality of interactions between educators, learners, and intelligent systems.

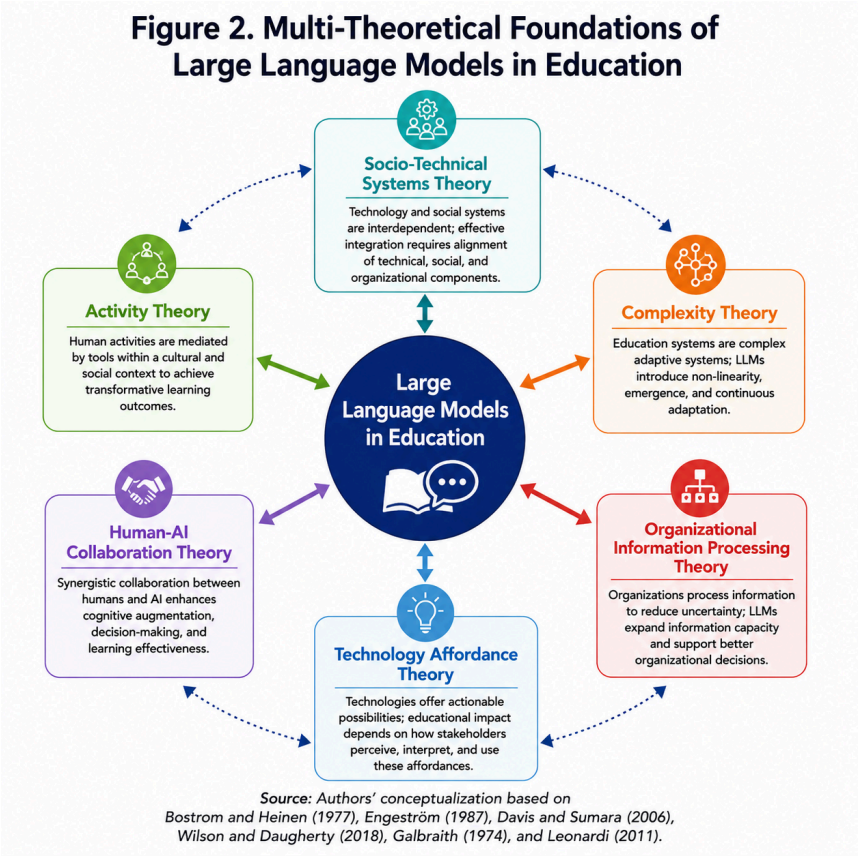
Organizational Information Processing Theory (Galbraith, 1974) also provides a valuable analytical lens. Educational institutions face increasing levels of complexity and uncertainty generated by digital transformation, regulatory pressures, and rapidly changing learner expectations. Artificial intelligence may function as an information-processing mechanism that expands organizational capacity to collect, interpret, and utilize information.

Finally, Technology Affordance Theory contributes to understanding how large language models enable new forms of educational practice. Affordances refer to the possibilities for action that emerge through interactions between users and technologies (Leonardi, 2011). Large language models create new educational affordances related to content generation, personalized support, knowledge synthesis, and collaborative learning, while simultaneously constraining other forms of educational practice.

Taken together, these theoretical perspectives suggest that large language models should not be understood simply as technological innovations. They represent socio-technical phenomena that reshape educational activities, organizational processes, and institutional arrangements.

Figure 2. Multi-Theoretical Foundations of Large Language Models in Education

DOI: https://doi.org/10.65222/VIRAL.2026.7.41.61



Source: Authors' conceptualization based on Galbraith (1974), Engeström (1987), Leonardi (2011), Davis and Sumara (2006), Jarrahi (2018), and Wilson and Daugherty (2018).

The multi-theoretical perspective presented in Figure 2 provides the conceptual foundation for the remainder of this review. It enables a more comprehensive understanding of how large language models influence educational processes and institutional transformation and offers an integrative lens through which the rapidly expanding literature can be synthesized.

3. Methodology

3.1 Research Design

The study adopts an **integrative literature review** approach to synthesize the rapidly expanding body of research on large language models in education. Integrative reviews are particularly appropriate for emerging and interdisciplinary fields because they enable the consolidation of fragmented knowledge, facilitate theory development, and allow the integration of conceptual, empirical, and methodological contributions from diverse research traditions (Torraco, 2005; Snyder, 2019).

Unlike bibliometric studies that focus primarily on citation patterns and publication networks, integrative reviews seek to develop a comprehensive understanding of a phenomenon by identifying dominant themes, theoretical foundations, conceptual tensions, and future

DOI: https://doi.org/10.65222/VIRAL.2026.7.41.61

research directions. The choice of this methodology is justified by the highly heterogeneous nature of research on ChatGPT, Claude, and generative artificial intelligence in education. Existing studies originate from multiple disciplines, employ different methodological traditions, and examine educational transformation from pedagogical, technological, organizational, and governance perspectives.

The review therefore follows a theory-building orientation and seeks to answer broader questions concerning how large language models are reshaping educational systems and what conceptual frameworks may be required to understand these developments.

3.2 Search Strategy

The literature search was conducted using four major academic databases that collectively provide broad coverage of educational, technological, and interdisciplinary research:

- Scopus
- Web of Science
- ERIC
- ScienceDirect

The search strategy was designed to capture the rapidly evolving literature on large language models and their educational applications. The following search string was employed:

("ChatGPT" OR "Claude" OR "Large Language Models" OR "Generative AI")

AND

("Education" OR "Teaching" OR "Learning" OR "Higher Education")

Additional searches were conducted using related terms such as:

- "Artificial Intelligence in Education"
- "Generative Artificial Intelligence and Higher Education"
- "Human-AI Collaboration in Education"
- "Educational Leadership and Artificial Intelligence"
- "AI Governance in Education"

Backward and forward citation tracking was subsequently employed to identify additional influential studies that may not have been captured through database searches.

Because the field has developed rapidly following the public release of ChatGPT in late 2022, particular attention was given to recent publications appearing between 2022 and 2025, although seminal works in artificial intelligence in education and educational technology were also included to provide historical and theoretical context.

DOI: https://doi.org/10.65222/VIRAL.2026.7.41.61

3.3 Inclusion Criteria

To ensure conceptual relevance and academic rigor, studies were included according to the following criteria:

- peer-reviewed journal articles;
- peer-reviewed conference papers;
- English-language publications;
- studies explicitly addressing large language models, generative AI, ChatGPT, Claude, or related technologies in educational settings;
- studies discussing pedagogical, organizational, leadership, governance, or policy implications.

The review intentionally adopted broad inclusion criteria because the objective was to capture the diversity of perspectives emerging within this rapidly developing field.

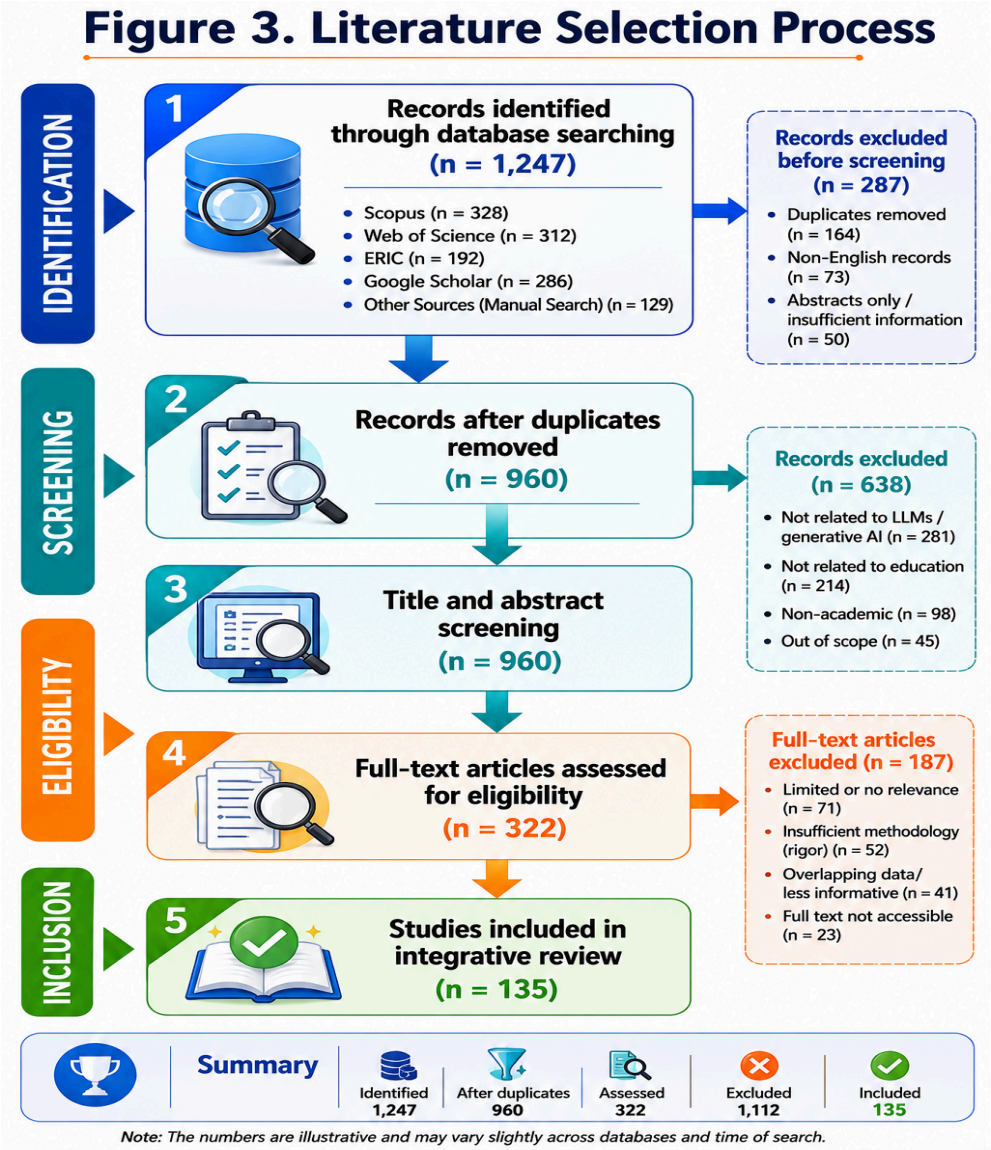
3.4 Exclusion Criteria

The following categories of publications were excluded:

- editorials and opinion pieces without substantive analysis;
- non-academic reports and magazine articles;
- studies focusing exclusively on medical education without broader educational implications;
- engineering papers centered solely on algorithmic performance and lacking educational relevance;
- duplicate publications and non-English documents.

These exclusion criteria were designed to maintain conceptual coherence and ensure that the final corpus remained directly relevant to educational transformation and institutional change.

DOI: https://doi.org/10.65222/VIRAL.2026.7.41.61



Source: Adapted from Page et al. (2021) and authors' methodological design.

3.5 Analytical Procedure

The analytical process followed four sequential stages.

The first stage involved a **descriptive analysis** of the selected literature, including publication trends, educational contexts, and principal topics addressed by existing studies.

The second stage consisted of **thematic coding**. Studies were coded inductively and iteratively to identify recurrent themes, conceptual categories, and emerging areas of inquiry. This process enabled the identification of dominant research streams and the relationships among them.

DOI: https://doi.org/10.65222/VIRAL.2026.7.41.61

The third stage focused on **conceptual synthesis**. The objective was not merely to summarize previous findings but to integrate diverse perspectives into a coherent theoretical understanding of large language models in education. Particular attention was given to connections among pedagogical innovation, organizational transformation, educational leadership, and governance.

The fourth stage involved **theory development**. Building on the thematic synthesis, the review developed an original conceptual framework that explains how large language models influence educational systems through interactions among institutional antecedents, technological capabilities, educational mechanisms, and governance processes.

The methodological procedure therefore moves beyond descriptive mapping and adopts an interpretive and theory-building orientation. This approach is particularly appropriate for a rapidly evolving field in which conceptual integration remains limited and where theoretical frameworks capable of guiding future empirical research are still emerging.

The following section presents the evolution of research on large language models in education and examines how scholarly attention has shifted from technological capabilities toward broader concerns regarding human-AI collaboration, institutional transformation, and educational governance.

4. Evolution of Large Language Models - LLM in Education

The emergence of large language models has generated one of the fastest-growing streams of research in the history of educational technology. Although artificial intelligence in education has been an established field for several decades, the public release of generative AI systems fundamentally altered both the scale and direction of scholarly attention. The literature evolved from relatively specialized discussions concerning intelligent tutoring systems and learning analytics toward broader debates involving academic integrity, institutional governance, educational leadership, and the future of higher education itself.

This evolution did not occur in a linear manner. Rather, it unfolded through several overlapping phases characterized by different research priorities, methodological approaches, and theoretical concerns. Early studies focused primarily on technological capabilities and immediate pedagogical applications. Subsequent research increasingly examined broader organizational and societal implications, including questions of trust, governance, and human-AI collaboration. Understanding this evolution is essential because it reveals how scholarly perspectives have shifted and highlights areas where theoretical development remains incomplete.

4.1 Early Research on Generative AI in Education

Prior to the emergence of ChatGPT, research on artificial intelligence in education concentrated largely on adaptive learning systems, intelligent tutoring environments, and learning analytics. Studies during this period emphasized personalization, automated

DOI: https://doi.org/10.65222/VIRAL.2026.7.41.61

feedback, predictive analytics, and student support mechanisms (Zawacki-Richter et al., 2019; Holmes et al., 2019). Artificial intelligence was generally conceptualized as a tool that could support educational processes by increasing efficiency and enabling more individualized learning experiences.

Early investigations of generative artificial intelligence remained relatively limited because available systems lacked the capabilities necessary for widespread educational adoption. Research often focused on narrow applications of natural language processing and automated text generation, with relatively little attention devoted to broader institutional implications.

The theoretical orientation of this early literature was predominantly technological and pedagogical. Researchers explored how intelligent systems could improve teaching and learning outcomes, but questions concerning organizational transformation, governance, and human-AI collaboration remained largely peripheral.

4.2 The ChatGPT Explosion

The release of ChatGPT in November 2022 represented a watershed moment in the history of artificial intelligence in education. Within a few months, the platform became one of the fastest-growing digital technologies ever adopted and triggered an extraordinary increase in academic publications. Universities, educators, and policymakers were suddenly confronted with a technology capable of generating essays, solving problems, producing code, and engaging in sophisticated dialogue.

The initial wave of research was characterized by a strong focus on opportunities and risks. Studies examined the potential of ChatGPT for tutoring, assessment, feedback, and content generation while simultaneously raising concerns regarding plagiarism, misinformation, academic integrity, and student dependency (Kasneci et al., 2023; Cotton et al., 2023; Lo, 2023).

Another notable feature of this period was the emergence of normative and ethical debates. Researchers increasingly questioned whether existing educational models remained appropriate in environments where students had immediate access to sophisticated generative technologies. The literature also began to address issues related to trust, bias, transparency, and the responsibilities of educational institutions in governing artificial intelligence.

The rapid diffusion of ChatGPT transformed artificial intelligence from a specialized research topic into a central concern for educational policy and institutional strategy. Consequently, the literature expanded considerably beyond classroom applications and started to address broader organizational and societal implications.

4.3 The Emergence of Claude and Advanced Educational Assistants

The introduction of Claude and other advanced large language models marked a second phase in the evolution of research. While early studies primarily focused on the novelty and

DOI: https://doi.org/10.65222/VIRAL.2026.7.41.61

disruptive potential of ChatGPT, later research began to examine differences among models and their suitability for specific educational contexts.

Claude attracted particular interest because of its emphasis on safety, interpretability, and long-context reasoning capabilities. Its ability to analyze extensive documents and engage in sophisticated analytical tasks opened new possibilities for research support, literature synthesis, policy analysis, and advanced educational applications.

The emergence of multiple high-performing models also shifted scholarly attention toward comparative analyses and questions concerning institutional choice. Educational institutions increasingly recognized that the adoption of generative AI involves strategic considerations related to governance, privacy, data security, and alignment with organizational values.

This phase of the literature also introduced a stronger focus on human-AI collaboration. Rather than treating large language models merely as productivity tools, researchers began conceptualizing them as cognitive partners capable of participating in knowledge creation and organizational learning processes (Mollick & Mollick, 2023; Jarrahi, 2018).

4.4 Current Research Trends

The current literature on large language models in education is characterized by increasing diversification and conceptual maturity. Several dominant themes have emerged.

The first concerns the transformation of teaching and learning processes through personalized and conversational educational experiences. Research increasingly investigates how large language models influence learner autonomy, metacognition, and adaptive learning.

The second theme focuses on assessment and academic integrity. Scholars are actively exploring alternative forms of assessment capable of preserving meaningful learning outcomes in environments where generative AI is widely available.

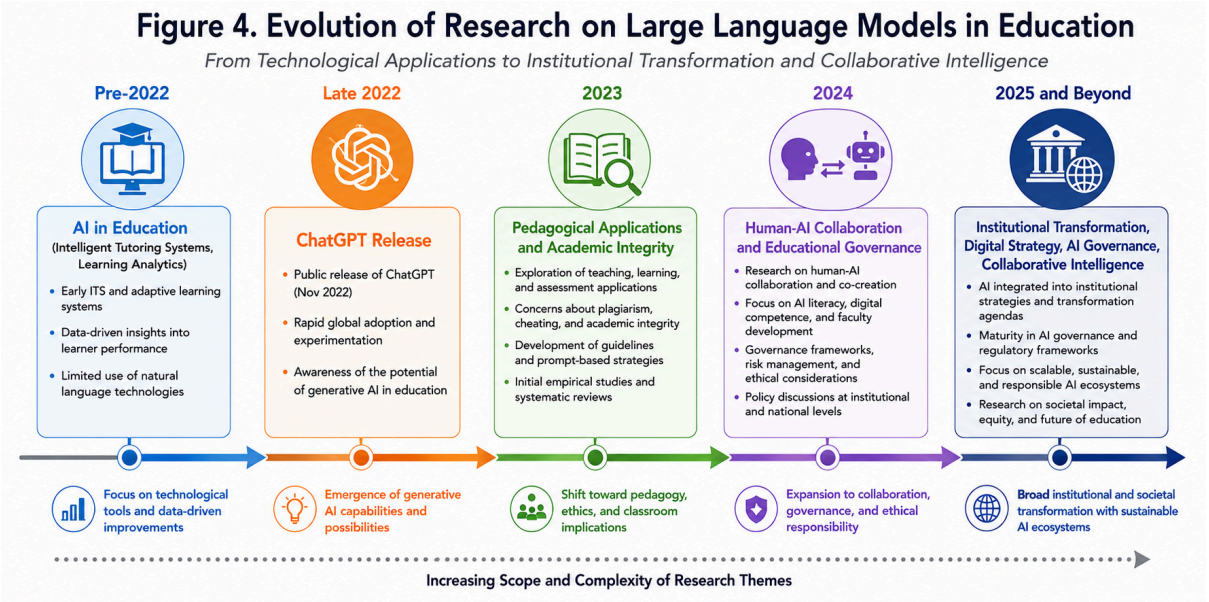
A third theme examines educational leadership and governance. Universities are beginning to develop institutional strategies and regulatory frameworks that address the opportunities and risks associated with large language models. This line of research remains comparatively underdeveloped but is growing rapidly.

The fourth theme concerns human-AI collaboration and the emergence of collaborative intelligence. Rather than emphasizing technological replacement, contemporary studies increasingly investigate how humans and intelligent systems can work together effectively and responsibly.

Finally, recent research has expanded toward broader questions of institutional transformation and digital strategy. Large language models are increasingly viewed as catalysts for organizational change that influence governance structures, information flows, and long-term educational planning.

DOI: https://doi.org/10.65222/VIRAL.2026.7.41.61

These developments suggest that the literature is moving from an initial period of technological enthusiasm and concern toward a more nuanced understanding of generative artificial intelligence as a complex socio-technical phenomenon embedded within educational systems.



Source: *Authors' synthesis based on Zawacki-Richter et al. (2019), Williamson and Eynon (2020), Kasneci et al. (2023), Lo (2023), and Dwivedi et al. (2023).*

The historical evolution presented in Figure 4 demonstrates that research on large language models in education is rapidly transitioning from a technology-centered field toward a broader interdisciplinary domain concerned with institutional transformation, governance, and the future relationship between humans and intelligent systems. This transition provides the foundation for the thematic synthesis developed in the following section.

5. Thematic Synthesis of the Literature

The rapid expansion of research on large language models in education has generated a highly diverse body of scholarship characterized by multiple disciplinary perspectives and methodological approaches. Although the literature initially concentrated on the technological novelty of systems such as ChatGPT and Claude, recent contributions increasingly examine broader educational, organizational, and societal implications. The thematic synthesis conducted in this review reveals seven dominant themes that collectively illustrate the evolving nature of research in this area.

The first theme concerns personalized learning and adaptive education. A substantial number of studies argue that large language models can facilitate individualized instruction by providing immediate feedback, tailoring explanations to learner needs, and supporting

DOI: https://doi.org/10.65222/VIRAL.2026.7.41.61

self-regulated learning processes (Kasneci et al., 2023; Lo, 2023; Mollick & Mollick, 2023). These studies generally frame generative artificial intelligence as a mechanism for increasing educational accessibility and responsiveness. At the same time, concerns remain regarding the potential development of cognitive dependency and the risk that excessive reliance on conversational AI may weaken independent problem-solving capacities.

A second major theme involves assessment and feedback. The ability of large language models to generate high-quality written content has fundamentally challenged conventional assessment practices in higher education. Researchers increasingly argue that traditional forms of assessment are becoming less effective in environments where students have immediate access to generative technologies (Cotton et al., 2023; Rudolph et al., 2023). Consequently, there is growing interest in authentic assessment, process-based evaluation, and forms of assessment that emphasize critical reflection and higher-order thinking.

The third theme focuses on human-AI co-teaching and collaborative intelligence. The literature increasingly rejects deterministic narratives that portray artificial intelligence as a replacement for educators. Instead, many scholars conceptualize large language models as cognitive partners capable of augmenting teaching practices and supporting knowledge creation (Jarrahi, 2018; Wilson & Daugherty, 2018). This perspective has stimulated discussions concerning new pedagogical models in which educators and intelligent systems operate collaboratively.

A fourth and highly visible theme concerns academic integrity and ethics. The emergence of generative artificial intelligence has intensified longstanding debates regarding plagiarism, authorship, transparency, and the responsible use of technology in educational environments (Tlili et al., 2023; Cotton et al., 2023). Ethical discussions have also expanded to include issues related to bias, misinformation, privacy, and algorithmic accountability.

Educational leadership and governance constitute a fifth theme that has gained increasing prominence. Universities are beginning to recognize that large language models influence not only classroom activities but also organizational structures, policy development, and strategic decision-making. Research in this area examines institutional readiness, governance frameworks, and leadership strategies capable of supporting responsible artificial intelligence adoption (Williamson & Eynon, 2020; Selwyn, 2019).

The sixth theme concerns institutional transformation and digital strategy. Large language models are increasingly viewed as catalysts of organizational change that affect administrative processes, information management, and long-term institutional planning. Several studies suggest that educational institutions are entering a new phase of digital transformation characterized by the integration of generative artificial intelligence into both academic and administrative functions (Dwivedi et al., 2023; Chiu, 2024).

Finally, a seventh theme focuses on student and teacher competencies in the age of large language models. Researchers increasingly argue that future educational systems must cultivate new forms of AI literacy, critical evaluation skills, and collaborative capabilities that enable individuals to work effectively alongside intelligent systems (Kasneci et al., 2023; UNESCO, 2023).

Collectively, these themes suggest that the literature has progressively moved beyond narrow discussions of technological capabilities toward broader questions concerning institutional adaptation, governance, and human-AI collaboration.

DOI: https://doi.org/10.65222/VIRAL.2026.7.41.61

Table 2. Thematic Synthesis of the Literature

Theme	Main Findings	Theoretical Foundations	Research Gaps	Future Directions
Personalized Learning and Adaptive Education	LLMs support individualized learning pathways and immediate feedback	Technology Affordance Theory; Self-Regulated Learning Theory	Limited evidence regarding long-term learning outcomes	Longitudinal studies on adaptive learning and learner autonomy
Assessment and Feedback	Traditional assessment methods are increasingly challenged by generative AI	Constructivism; Assessment Theory	Lack of validated AI-compatible assessment frameworks	Development of authentic and process-oriented assessment models
Human-AI Co-Teaching	AI increasingly functions as a cognitive partner rather than a replacement for teachers	Human-AI Collaboration Theory; Socio-Technical Systems Theory	Limited understanding of collaborative teaching models	Studies of hybrid pedagogical ecosystems and collaborative intelligence
Academic Integrity and Ethics	Concerns regarding plagiarism, transparency, bias, and responsible use continue to grow	Ethics of Technology; Responsible AI Frameworks	Lack of consensus regarding acceptable AI use in education	Development of institutional ethics and governance frameworks
Educational Leadership and Governance	LLM adoption requires strategic leadership and organizational adaptation	Organizational Information Processing Theory; Complexity Theory	Governance mechanisms remain underdeveloped	Comparative studies of institutional AI governance models
Institutional Transformation and Digital Strategy	Generative AI influences administrative processes and organizational structures	Digital Transformation Theory; Socio-Technical Systems Theory	Limited evidence concerning organizational consequences	Research on institutional resilience and AI-enabled transformation
Student and Teacher Competencies	New forms of AI literacy and critical evaluation skills are increasingly required	Competency Theory; Human Capital Theory	Unclear models of AI-related competencies	Development of frameworks for AI literacy and professional capability
Faculty Development and Professional Learning	Educators require new competencies and pedagogical strategies	Adult Learning Theory; Communities of Practice	Limited research on faculty adaptation processes	Studies of professional learning and organizational capability building
Educational Equity and Inclusion	LLMs may increase accessibility but also reproduce digital inequalities	Social Justice Theory; Digital Divide Theory	Limited understanding of differential impacts across contexts	Research on inclusive AI strategies and equitable access
Student Agency and Cognitive Dependency	AI support may simultaneously increase efficiency and reduce independent reasoning	Cognitive Load Theory; Activity Theory	Lack of evidence regarding long-term cognitive consequences	Studies examining critical thinking and cognitive resilience

DOI: https://doi.org/10.65222/VIRAL.2026.7.41.61

Theme	Main Findings	Theoretical Foundations	Research Gaps	Future Directions
Data Governance and Privacy	Increased use of LLMs raises concerns regarding data security and institutional accountability	Information Governance Theory; Institutional Theory	Governance frameworks remain fragmented	Development of policy and regulatory approaches for educational AI
Future of Higher Education	LLMs are reshaping assumptions regarding knowledge, expertise, and educational purpose	Complexity Theory; Organizational Learning Theory	Limited theoretical integration	Development of systemic models of AI-enabled educational transformation

Source: *Authors' synthesis based on Kasneci et al. (2023), Lo (2023), Dwivedi et al. (2023), Tlili et al. (2023), Chiu (2024), UNESCO (2023), Williamson and Eynon (2020), and related literature.*

The thematic synthesis reveals two important patterns. First, the literature is becoming increasingly interdisciplinary, drawing simultaneously upon educational technology, organizational studies, leadership research, and information systems. Second, the focus of scholarship is gradually shifting from questions concerning what large language models can do toward broader questions regarding how educational institutions should govern, integrate, and coexist with these technologies.

This shift provides the foundation for developing a more comprehensive theoretical framework capable of explaining the mechanisms through which large language models contribute to educational transformation. The next section therefore proposes an original conceptual model that integrates antecedents, technological capabilities, educational mechanisms, and institutional outcomes associated with large language model adoption.

6. Toward a Conceptual Framework of LLM-Enabled Educational Transformation

The thematic synthesis demonstrates that research on large language models in education has evolved from isolated studies of technological capabilities toward a more complex understanding of educational transformation as a multi-level and socio-technical process. Yet the literature remains theoretically fragmented. Most studies focus on specific educational applications, such as assessment, feedback, or content generation, while relatively few attempts have been made to integrate technological, organizational, and governance dimensions into a coherent explanatory framework.

To address this limitation, the present study proposes the **Large Language Model Educational Transformation Framework (LLM-ETF)**. The framework conceptualizes educational transformation as an emergent process generated by interactions among institutional antecedents, technological capabilities, educational mechanisms, and contextual moderators. Rather than treating large language models as isolated technological artifacts, the framework positions them within broader organizational and institutional systems.

DOI: https://doi.org/10.65222/VIRAL.2026.7.41.61

The model begins with a set of antecedent conditions that influence the capacity of educational institutions to adopt and integrate large language models effectively. Digital readiness constitutes a critical prerequisite because institutions differ significantly in their technological infrastructures, data management capabilities, and levels of digital maturity (Bond et al., 2018). Leadership support also plays a central role. The implementation of artificial intelligence frequently requires strategic direction, resource allocation, and organizational legitimacy that can only be provided through effective educational leadership.

Data capabilities represent another important antecedent. Large language models increasingly rely on data-rich environments in which information can be collected, processed, and utilized to support educational processes and decision-making. Similarly, faculty competencies significantly influence implementation outcomes because educators must possess not only technical knowledge but also pedagogical and ethical capacities related to artificial intelligence use.

Institutional culture completes this set of antecedents. Universities vary considerably in their openness toward innovation, experimentation, and technological change. Institutional norms and values therefore shape the extent to which large language models are accepted and integrated into educational practices.

These antecedent conditions influence the educational value generated by the capabilities of large language models themselves. The framework identifies five principal capabilities: content generation, reasoning, personalization, knowledge assistance, and conversational interaction.

Content generation enables the creation of educational materials, summaries, examples, and assessment items. Reasoning capabilities support analytical tasks, problem-solving activities, and knowledge synthesis. Personalization allows the adaptation of educational experiences to individual learner characteristics and needs. Knowledge assistance facilitates information retrieval, explanation, and tutoring functions, while conversational interaction creates opportunities for iterative dialogue and continuous learner support.

The framework proposes that these technological capabilities influence educational transformation through several mediating mechanisms.

Teaching augmentation refers to the capacity of large language models to support educators in instructional design, content preparation, and classroom activities. Student support includes tutoring, feedback, and personalized learning assistance. Assessment transformation captures the growing need to redesign evaluation methods in response to generative artificial intelligence. Administrative efficiency reflects the increasing use of large language models in institutional communication, advising, and decision support processes. Finally, knowledge creation recognizes the role of generative artificial intelligence in research activities and collaborative knowledge production.

The interaction of these mechanisms produces several organizational and educational outcomes. Learning effectiveness may improve through increased personalization and access

DOI: https://doi.org/10.65222/VIRAL.2026.7.41.61

to support. Educational innovation emerges through experimentation with new pedagogical models and human-AI collaboration practices. Educational quality may increase through enhanced responsiveness and efficiency, while institutional resilience reflects the capacity of universities to adapt to technological and environmental change. At a broader level, these developments may contribute to organizational performance by strengthening adaptability and knowledge management capabilities.

Importantly, the framework assumes that these relationships are not deterministic. Their strength and direction are influenced by a series of contextual moderators.

Ethics constitutes a fundamental moderating factor because concerns related to fairness, transparency, and accountability influence both adoption and acceptance of artificial intelligence. Trust also plays a critical role because educators and students are unlikely to engage meaningfully with technologies that they perceive as unreliable or opaque. Governance mechanisms shape institutional capacity to regulate and coordinate artificial intelligence implementation, while regulatory environments influence acceptable forms of use and institutional responsibilities. Finally, digital inequality may significantly affect outcomes by creating unequal access to technological resources and competencies.

The proposed framework may also be represented conceptually through the following relationships:

$$ET=f(A,L,M)$$

where:

- **ET** = educational transformation;
- **A** = antecedent conditions;
- **L** = large language model capabilities;
- **M** = educational mechanisms.

A more detailed representation may be expressed as:

$$ET=\alpha+\beta_1 DR+\beta_2 LS+\beta_3 DC+\beta_4 FC+\beta_5 IC+\varepsilon$$

where:

- **DR** = digital readiness;
- **LS** = leadership support;
- **DC** = data capabilities;
- **FC** = faculty competencies;
- **IC** = institutional culture;
- **α** = baseline institutional capacity;
- **β** coefficients represent theoretical relationships;
- **ε** captures contextual influences.

DOI: https://doi.org/10.65222/VIRAL.2026.7.41.61

The framework further assumes that outcomes are moderated by institutional and environmental conditions:

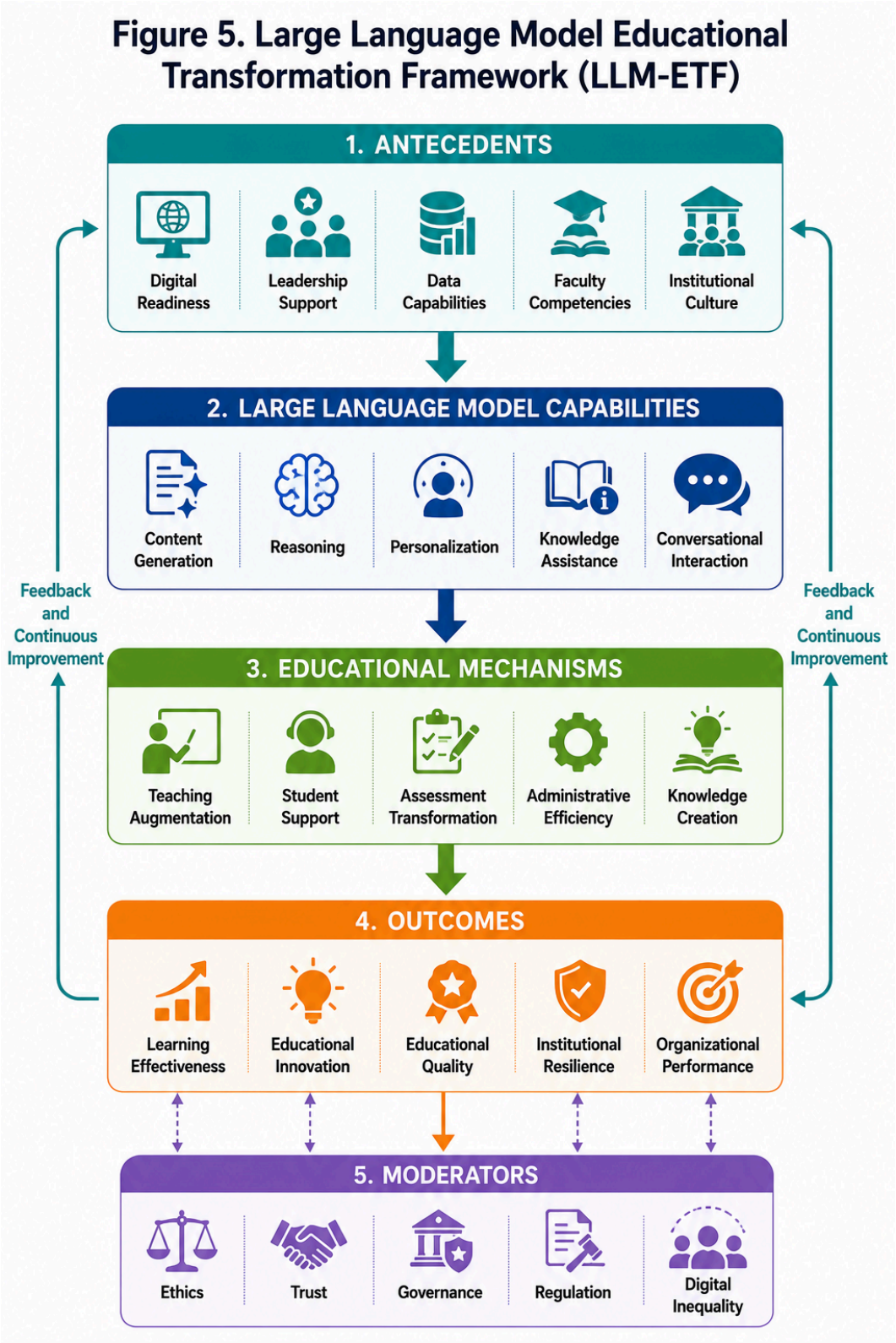
$$O=f(ET,E,T,G,R,DI)$$

where:

- **O** = organizational and educational outcomes;
- **E** = ethics;
- **T** = trust;
- **G** = governance;
- **R** = regulation;
- **DI** = digital inequality.

These formulations are purely conceptual and do not imply empirical estimation. Their purpose is to formalize relationships identified through the literature synthesis and provide a foundation for future empirical investigations.

DOI: https://doi.org/10.65222/VIRAL.2026.7.41.61



Source: Authors' conceptualization.

The LLM-ETF contributes to the literature in three important ways. First, it integrates fragmented research streams into a single explanatory model that captures technological,

DOI: https://doi.org/10.65222/VIRAL.2026.7.41.61

pedagogical, and organizational dimensions simultaneously. Second, it positions educational transformation as a dynamic and systemic process rather than a simple consequence of technology adoption. Third, it provides a theoretical foundation for future empirical studies examining the antecedents, mechanisms, and outcomes associated with large language model implementation in educational settings.

7. Research Agenda

The rapid expansion of research on large language models in education has produced a substantial body of scholarship within a remarkably short period of time. Yet the thematic synthesis conducted in this review indicates that the field remains at an early stage of theoretical consolidation. Much of the existing literature is exploratory, descriptive, and focused on immediate educational applications. There is comparatively limited evidence concerning the long-term consequences of large language model adoption for educational institutions, organizational structures, and governance arrangements.

Future research would benefit from moving beyond questions of technological capability and adoption toward broader investigations of how generative artificial intelligence reshapes educational systems and professional practices. Several areas appear particularly important.

The first concerns institutional digital maturity. Existing studies frequently assume that the educational effects of large language models are determined primarily by the characteristics of the technologies themselves. However, institutional contexts vary considerably in terms of technological infrastructure, leadership capacity, and organizational readiness. Future research should therefore examine how these contextual factors influence implementation outcomes and whether differences in institutional maturity explain variations in educational effectiveness.

A second area concerns human-AI collaboration. Although numerous studies discuss the potential of large language models to support teaching and learning, relatively little is known about the mechanisms through which educators and intelligent systems may collaborate effectively over extended periods of time. Questions regarding trust, task allocation, professional identity, and cognitive dependency remain largely unresolved.

Educational leadership represents a third area requiring further investigation. The literature increasingly recognizes that successful implementation of artificial intelligence depends on institutional leadership and governance, yet there remains limited understanding of the capabilities and strategies required to lead educational transformation in AI-rich environments. Comparative studies of universities and educational systems may provide valuable insights into these issues.

Trust and governance also emerge as critical research priorities. Generative artificial intelligence raises significant questions regarding accountability, transparency, privacy, and algorithmic bias. Future research should therefore examine the conditions under which

DOI: https://doi.org/10.65222/VIRAL.2026.7.41.61

educational stakeholders develop trust in large language models and the governance mechanisms that support responsible implementation.

Finally, there is a need to better understand the competencies required in educational environments increasingly characterized by human-AI collaboration. Existing research has only begun to address issues related to AI literacy, prompt literacy, and collaborative intelligence. More comprehensive theoretical and empirical models of these competencies remain necessary.

Building upon these observations, the present review proposes five propositions intended to guide future investigations.

Proposition 1 - Institutional digital maturity positively influences the effectiveness of large language model adoption in educational environments.

Institutions with stronger digital infrastructures, higher levels of technological readiness, and greater organizational capabilities are likely to integrate large language models more effectively and derive greater educational value from their implementation.

Proposition 2 - Human-AI collaboration positively contributes to educational innovation and the development of new pedagogical practices.

Educational systems that successfully integrate human expertise and artificial intelligence capabilities may demonstrate higher levels of experimentation, adaptability, and instructional innovation.

Proposition 3 - Educational leadership mediates the relationship between large language model adoption and institutional transformation.

Leadership capacities related to strategic vision, governance, and change management are expected to play a critical role in determining whether artificial intelligence implementation produces meaningful organizational transformation.

Proposition 4 - Trust and governance moderate the educational impact of large language models.

The benefits associated with generative artificial intelligence are likely to depend significantly on stakeholder trust, ethical frameworks, and institutional governance arrangements.

Proposition 5 - Faculty competencies positively influence educational outcomes generated by large language models.

Educators who possess stronger AI literacy, pedagogical flexibility, and technological competencies are expected to use large language models more effectively and generate better learning outcomes.

DOI: https://doi.org/10.65222/VIRAL.2026.7.41.61

Table 3. Future Research Agenda

Research Gap	Relevant Theory	Suggested Methodology	Context
Institutional determinants of LLM adoption	Digital Transformation Theory; Institutional Theory	Structural Equation Modeling; Multi-case studies	Universities and higher education systems
Human-AI collaboration mechanisms	Human-AI Collaboration Theory; Activity Theory	Mixed methods; Longitudinal studies	Teaching and learning environments
Leadership and organizational transformation	Organizational Information Processing Theory; Complexity Theory	Comparative case studies; Multi-level modeling	Educational leadership and governance
Trust and acceptance of LLMs	Technology Acceptance Theory; Trust Theory	Surveys; Experimental designs	Students, faculty, and administrators
AI literacy and faculty competencies	Competency Theory; Human Capital Theory	Survey research; Scale development	Higher education institutions
Long-term educational outcomes	Socio-Technical Systems Theory	Longitudinal studies	Digital transformation in education
Governance and regulation of generative AI	Institutional Theory; Governance Theory	Comparative policy analysis	National and international contexts
Academic integrity in AI-rich environments	Ethics of Technology; Responsible AI Frameworks	Qualitative studies; Design-based research	Assessment and quality assurance
Digital inequality and access	Digital Divide Theory; Social Justice Theory	Comparative studies; Cross-country analyses	Developing and developed economies
Organizational resilience and AI adoption	Organizational Learning Theory; Complexity Theory	Systems dynamics; Multi-level analysis	University transformation and strategic management
Collaborative intelligence development	Human-AI Collaboration Theory	Longitudinal and mixed-method designs	Human-AI educational ecosystems
AI-enabled institutional performance	Resource-Based View; Dynamic Capabilities Theory	Structural Equation Modeling	Universities and educational organizations

Source: *Authors' conceptualization based on the integrative review.*

DOI: https://doi.org/10.65222/VIRAL.2026.7.41.61

The proposed agenda illustrates that research on large language models in education is entering a new phase. The initial focus on technological novelty and immediate pedagogical applications is gradually giving way to more complex questions concerning organizational transformation, governance, leadership, and the long-term relationship between human and artificial intelligence in educational systems.

Addressing these questions will require greater theoretical integration and methodological diversity. Multi-level studies, longitudinal designs, comparative analyses, and theory-driven empirical investigations are likely to become increasingly important as the field matures. Such research may contribute not only to a deeper understanding of large language models in education but also to broader discussions concerning the future of knowledge, work, and human-AI collaboration in society.

8. Implications

Theoretical Implications

The emergence of large language models in education has important implications for several streams of scholarship, including Artificial Intelligence in Education, Educational Technology, Human-AI Collaboration, Educational Leadership, and Digital Transformation research. The findings of this review suggest that existing theoretical perspectives remain only partially capable of explaining the complexity of educational change generated by generative artificial intelligence.

First, the study contributes to Artificial Intelligence in Education by moving beyond a predominantly pedagogical and technology-centered perspective. Much of the existing literature examines large language models as instructional tools that support tutoring, feedback, and content generation (Kasneci et al., 2023; Lo, 2023). The present review argues that such an approach captures only part of the phenomenon. Large language models increasingly function as organizational technologies that influence information processing, decision-making, and institutional governance. Consequently, future theory development should conceptualize generative artificial intelligence as an integral component of educational ecosystems rather than merely as a teaching aid.

Second, the review contributes to Educational Technology research by highlighting the importance of socio-technical interactions. Educational outcomes cannot be understood solely through technological characteristics or user acceptance. Instead, they emerge from complex interactions among technological capabilities, institutional conditions, leadership practices, and human agency. This perspective aligns with socio-technical systems theory and supports calls for more systemic approaches to educational innovation.

Third, the study extends Human-AI Collaboration Theory by introducing the concept of collaborative educational intelligence. Existing discussions frequently frame artificial intelligence either as a productivity tool or as a potential substitute for human labor. The evidence synthesized in this review suggests a more nuanced interpretation. In educational

DOI: https://doi.org/10.65222/VIRAL.2026.7.41.61

contexts, effective outcomes appear to depend on complementarities between human expertise and machine capabilities. Large language models are therefore best understood as cognitive partners that augment, rather than replace, educational processes.

Fourth, the review contributes to research on Educational Leadership by emphasizing that artificial intelligence adoption constitutes an organizational and governance challenge rather than a purely technological one. Leadership influences institutional readiness, strategic direction, and the legitimacy of technological transformation. Consequently, future research should pay greater attention to the role of leadership capabilities in shaping educational responses to generative artificial intelligence.

Finally, the proposed Large Language Model Educational Transformation Framework contributes to Digital Transformation research by offering an integrative model that links antecedents, technological capabilities, educational mechanisms, and institutional outcomes. The framework provides a theoretical basis for future empirical studies and may facilitate cumulative knowledge development in this rapidly evolving field.

Practical Implications

The findings of this review also have important implications for educational practice. The rapid adoption of large language models creates opportunities for innovation but simultaneously generates new responsibilities for educational institutions and stakeholders.

For teachers, the findings suggest that professional roles are undergoing significant transformation. Educators increasingly operate in environments where students have immediate access to sophisticated conversational AI systems. Teaching practices may therefore need to shift from information transmission toward activities that emphasize critical thinking, problem-solving, reflective judgment, and the responsible use of artificial intelligence. Faculty development programs should support the acquisition of AI literacy, prompt design capabilities, and competencies related to human-AI collaboration.

For educational leaders, the review highlights the importance of strategic preparedness and organizational adaptability. Decisions regarding the adoption of large language models cannot be reduced to technological procurement. They involve questions concerning governance, institutional values, risk management, and long-term educational objectives. Universities therefore require comprehensive strategies capable of integrating technological innovation with educational missions and ethical responsibilities.

For higher education institutions, the findings indicate that curriculum redesign may become increasingly necessary. Traditional educational models built primarily around information acquisition and routine knowledge reproduction may be less effective in environments characterized by widespread access to generative artificial intelligence. Curricula may need to place greater emphasis on analytical reasoning, interdisciplinary thinking, creativity, and collaborative intelligence.

The review also suggests that universities should invest in digital infrastructures and

DOI: https://doi.org/10.65222/VIRAL.2026.7.41.61

organizational capabilities that support responsible and sustainable artificial intelligence adoption. Institutions that fail to develop such capabilities may encounter increasing difficulties in responding to technological change and evolving student expectations.

Policy Implications

The rapid diffusion of large language models has created important challenges for educational governance and public policy. The findings indicate that regulatory frameworks and institutional policies have not evolved at the same pace as technological innovation, generating uncertainty regarding acceptable forms of use and appropriate governance mechanisms.

A first implication concerns AI governance. Educational institutions increasingly require formal frameworks capable of addressing issues related to transparency, accountability, and ethical responsibility. Policies should clearly define expectations regarding the use of generative artificial intelligence by students, faculty, and administrative personnel while maintaining sufficient flexibility to accommodate technological evolution.

A second implication relates to educational regulation and quality assurance. Existing accreditation and evaluation frameworks were largely developed before the emergence of large language models and may not adequately address the challenges associated with human-AI collaboration. Policymakers and quality assurance agencies may therefore need to reconsider standards related to assessment, academic integrity, and digital competencies.

The review also highlights the importance of digital ethics. Large language models raise significant concerns regarding bias, privacy, misinformation, intellectual property, and the potential concentration of technological power. Educational policy should therefore promote principles of fairness, transparency, and human oversight while encouraging innovation and experimentation.

Finally, the findings emphasize issues of digital inequality. Access to advanced artificial intelligence technologies remains uneven across institutions, regions, and socioeconomic groups. The educational benefits associated with large language models may therefore be distributed unequally, potentially reinforcing existing disparities in educational opportunity. Policymakers should pay particular attention to ensuring equitable access to technological resources and the competencies required to use them effectively.

Taken together, these implications suggest that the future of large language models in education will depend not only on technological progress but also on the capacity of educational institutions and policymakers to develop governance arrangements that balance innovation with responsibility, efficiency with equity, and technological capability with human values.

DOI: https://doi.org/10.65222/VIRAL.2026.7.41.61

9. Conclusions

The rapid emergence of large language models has introduced one of the most significant transformations in the history of educational technology. Systems such as ChatGPT and Claude have moved artificial intelligence from specialized applications toward mainstream educational practice and have stimulated extensive debates concerning teaching, learning, assessment, governance, and the future of higher education. The literature reviewed in this study demonstrates that the educational implications of these technologies extend far beyond their technical capabilities. Large language models increasingly participate in knowledge production, support organizational decision-making, and contribute to the reconfiguration of relationships among learners, educators, and institutions.

Several important patterns emerge from the analysis. First, research on large language models in education has evolved rapidly from an initial focus on technological novelty and pedagogical experimentation toward broader concerns related to human-AI collaboration, institutional transformation, and governance. Second, despite the growing volume of publications, the literature remains fragmented across disciplinary boundaries. Educational studies frequently concentrate on classroom applications, while technological research emphasizes model performance and computational capabilities. Comparatively fewer studies examine the organizational and governance implications of generative artificial intelligence or seek to integrate these perspectives into coherent theoretical frameworks.

The present review contributes to addressing this fragmentation by proposing an integrative conceptual perspective on large language models in education. The Large Language Model Educational Transformation Framework conceptualizes educational change as a systemic process generated through interactions among institutional antecedents, technological capabilities, educational mechanisms, and contextual moderators. This perspective shifts the analytical focus from technology adoption alone toward broader questions concerning institutional readiness, governance, and long-term organizational adaptation.

The review also contributes to theory development in several ways. It extends research in Artificial Intelligence in Education by conceptualizing large language models as socio-technical and organizational phenomena rather than solely pedagogical tools. It contributes to Human-AI Collaboration Theory by introducing the notion of collaborative educational intelligence and emphasizing complementarities between human expertise and machine capabilities. Furthermore, it enriches research on educational leadership and digital transformation by demonstrating that successful implementation of generative artificial intelligence depends not only on technological sophistication but also on institutional capacity, governance arrangements, and strategic leadership.

The findings carry important implications for educational practice and policy. Universities increasingly require governance frameworks capable of addressing ethical concerns, accountability requirements, and questions of institutional legitimacy associated with artificial intelligence adoption. Faculty development initiatives and curriculum redesign efforts will also become increasingly important as educational systems seek to prepare learners for environments characterized by continuous interaction with intelligent

DOI: https://doi.org/10.65222/VIRAL.2026.7.41.61

technologies. Equally important is the need to address issues of digital inequality and ensure that the benefits associated with generative artificial intelligence do not exacerbate existing disparities in educational opportunity.

Several limitations should nevertheless be acknowledged. The study employs an integrative literature review methodology and therefore depends on the scope and quality of existing scholarship. The field of large language models in education is evolving exceptionally rapidly, and new developments may alter current understandings and research priorities. In addition, the proposed conceptual framework remains theoretical and has not been empirically validated. The relationships identified in the model should therefore be interpreted as propositions requiring further investigation rather than established causal mechanisms.

Future research may proceed in several directions. Empirical studies could test the relationships proposed in the Large Language Model Educational Transformation Framework using structural equation modeling, longitudinal designs, or multi-level analyses. Comparative investigations across educational systems and national contexts may also improve understanding of how governance arrangements and institutional conditions shape artificial intelligence adoption. Additional research is needed on trust, digital ethics, academic integrity, collaborative intelligence, and the long-term consequences of human-AI interaction for educational institutions and professional practice.

The emergence of large language models marks a critical moment in the evolution of higher education. The central challenge is no longer whether these technologies will become part of educational systems, but rather how educational institutions can develop forms of governance, leadership, and human-AI collaboration that preserve human agency while taking advantage of the opportunities created by increasingly capable intelligent systems. The future of education will likely depend on the capacity to build institutions in which artificial and human intelligence coexist in ways that are ethically responsible, pedagogically meaningful, and organizationally sustainable.

References

1. Bai, Y., Kadavath, S., Kundu, S., Askell, A., Kernion, J., Jones, A., Chen, A., Goldie, A., Mirhoseini, A., McKinnon, C., Chen, C., Olsson, C., Olah, C., Hernandez, D., Drain, D., Ganguli, D., Li, D., Tran-Johnson, E., Perez, E., et al. (2022). Constitutional AI: Harmlessness from AI feedback. *arXiv Preprint arXiv:2212.08073*. https://doi.org/10.48550/arXiv.2212.08073

[2] 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.

[3] Bearman, M., Ryan, J., & Ajjawi, R. (2023). Discourses of artificial intelligence in higher education: A critical literature review. *Higher Education, 86*(2), 369-385.

DOI: https://doi.org/10.65222/VIRAL.2026.7.41.61

[4] Bond, M., Marín, V. I., Dolch, C., Bedenlier, S., & Zawacki-Richter, O. (2018). Digital transformation in German higher education: Student and teacher perceptions and usage of digital media. *International Journal of Educational Technology in Higher Education, 15*(1), 48. https://doi.org/10.1186/s41239-018-0130-1

[5] Bommasani, R., Hudson, D. A., Adeli, E., Altman, R., Arora, S., von Arx, S., Bernstein, M. S., Bohg, J., Bosselut, A., Brunskill, E., Brynjolfsson, E., Buch, S., Card, D., Castellon, R., Chatterji, N., Chen, A., Creel, K., Davis, J. Q., Demszky, D., et al. (2021). *On the opportunities and risks of foundation models*. Stanford Institute for Human-Centered Artificial Intelligence.

[6] Bostrom, R. P., & Heinen, J. S. (1977). MIS problems and failures: A socio-technical perspective. Part I: The causes. *MIS Quarterly, 1*(3), 17-32.

[7] Bubeck, S., Chandrasekaran, V., Eldan, R., Gehrke, J., Horvitz, E., Kamar, E., Lee, P., Lee, Y. T., Li, Y., Lundberg, S., Nori, H., Palangi, H., Ribeiro, M. T., & Zhang, Y. (2023). Sparks of artificial general intelligence: Early experiments with GPT-4. *arXiv Preprint arXiv:2303.12712*. https://doi.org/10.48550/arXiv.2303.12712

[8] Chiu, T. K. F. (2024). Future research recommendations for transforming higher education with generative AI. *Computers and Education: Artificial Intelligence, 6*, 100197.

[9] Crompton, H., & Burke, D. (2023). Artificial intelligence in higher education: The state of the field. *International Journal of Educational Technology in Higher Education, 20*(1), 22. https://doi.org/10.1186/s41239-023-00392-8

[10] Cotton, D. R. E., Cotton, P. A., & Shipway, J. R. (2023). Chatting and cheating: Ensuring academic integrity in the era of ChatGPT. *Innovations in Education and Teaching International, 61*(2), 228-239. https://doi.org/10.1080/14703297.2023.2190148

[11] Davis, B., & Sumara, D. (2006). *Complexity and education: Inquiries into learning, teaching, and research*. Lawrence Erlbaum Associates.

[12] Dwivedi, Y. K., Kshetri, N., Hughes, L., Slade, E. L., Jeyaraj, A., Kar, A. K., Baabdullah, A. M., Koohang, A., Raghavan, V., Ahuja, M., Albanna, H., Albashrawi, M., Al-Busaidi, A. S., Balakrishnan, J., Barlette, Y., Basu, S., Bose, I., Brooks, L., Buhalis, D., et al. (2023). So what if ChatGPT wrote it? Multidisciplinary perspectives on opportunities, challenges and implications of generative conversational AI for research, practice and policy. *International Journal of Information Management, 71*, 102642. https://doi.org/10.1016/j.ijinfomgt.2023.102642

[13] Engeström, Y. (1987). *Learning by expanding: An activity-theoretical approach to developmental research*. Orienta-Konsultit.

[14] Fashogbon, Bamikole Abel, Raimi Olalekan Adeleke, and Omobonike Adejoke Olowe. "The Application of Artificial Intelligence in economics: A review of current trends and future directions." *International Journal of Education, Leadership, Artificial Intelligence, Computing, Business, Life Sciences, and Society* 2.02 (2025): 67-89. https://doi.org/10.65222/VIRAL.2025.9.10

[15] Floridi, L., & Chiriatti, M. (2020). GPT-3: Its nature, scope, limits, and consequences. *Minds and Machines, 30*(4), 681-694. https://doi.org/10.1007/s11023-020-09548-1

DOI: https://doi.org/10.65222/VIRAL.2026.7.41.61

[16] Galbraith, J. R. (1974). Organization design: An information processing view. *Interfaces, 4*(3), 28-36. https://doi.org/10.1287/inte.4.3.28

[17] Holmes, W., Bialik, M., & Fadel, C. (2019). *Artificial intelligence in education: Promises and implications for teaching and learning*. Center for Curriculum Redesign.

[18] Holmes, W., & Tuomi, I. (2022). State of the art and practice in AI in education. *European Journal of Education, 57*(4), 542-570. https://doi.org/10.1111/ejed.12533

[19] Jarrahi, M. H. (2018). Artificial intelligence and the future of work: Human-AI symbiosis in organizational decision making. *Business Horizons, 61*(4), 577-586. https://doi.org/10.1016/j.bushor.2018.03.007

[20] Kasneci, E., Sessler, K., Küchemann, S., Bannert, M., Dementieva, D., Fischer, F., Gasser, U., Groh, G., Günnemann, S., Hüllermeier, E., Krusche, S., Kutyniok, G., Michaeli, T., Nerdel, C., Pfeiffer, F., Poquet, O., Sailer, M., Schmidt, A., Seidel, T., & Kasneci, G. (2023). ChatGPT for good? On opportunities and challenges of large language models for education. *Learning and Individual Differences, 103*, 102274. https://doi.org/10.1016/j.lindif.2023.102274

[21] Leonardi, P. M. (2011). When flexible routines meet flexible technologies: Affordance, constraint, and the imbrication of human and material agencies. *MIS Quarterly, 35*(1), 147-167. https://doi.org/10.2307/23043493

[22] Lim, W. M., Gunasekara, A. N., Pallant, J. L., Pallant, J. I., & Pechenkina, E. (2023). Generative AI and the future of education: Ragnarök or reformation? A paradoxical perspective from management educators. *International Journal of Management Education, 21*(2), 100790. https://doi.org/10.1016/j.ijme.2023.100790

[23] Lo, C. K. (2023). What is the impact of ChatGPT on education? A rapid review of the literature. *Education Sciences, 13*(4), 410. https://doi.org/10.3390/educsci13040410

[24] Luckin, R., Holmes, W., Griffiths, M., & Forcier, L. B. (2016). *Intelligence unleashed: An argument for AI in education*. Pearson.

[25] Lund, B. D., Wang, T., Mannuru, N. R., Nie, B., Shimray, S., & Wang, Z. (2023). ChatGPT and a new academic reality: Artificial intelligence-written research papers and the ethics of the large language models in scholarly publishing. *Journal of the Association for Information Science and Technology, 74*(5), 570-581. https://doi.org/10.1002/asi.24750

[26] Mollick, E., & Mollick, L. (2023). Assigning AI: Seven approaches for students, with prompts. *SSRN Electronic Journal*. https://doi.org/10.2139/ssrn.4475995

[27] Ouyang, L., Wu, J., Jiang, X., Almeida, D., Wainwright, C., Mishkin, P., Zhang, C., Agarwal, S., Slama, K., Ray, A., Schulman, J., Hilton, J., Kelton, F., Miller, L., Simens, M., Askell, A., Welinder, P., Christiano, P., Leike, J., & Lowe, R. (2022). Training language models to follow instructions with human feedback. *Advances in Neural Information Processing Systems, 35*, 27730-27744.

[28] OpenAI. (2023). *GPT-4 technical report. arXiv Preprint arXiv:2303.08774*. https://doi.org/10.48550/arXiv.2303.08774

DOI: https://doi.org/10.65222/VIRAL.2026.7.41.61

[29] Page, M. J., McKenzie, J. E., Bossuyt, P. M., Boutron, I., Hoffmann, T. C., Mulrow, C. D., Shamseer, L., Tetzlaff, J. M., Akl, E. A., Brennan, S. E., Chou, R., Glanville, J., Grimshaw, J. M., Hróbjartsson, A., Lalu, M. M., Li, T., Loder, E. W., Mayo-Wilson, E., McDonald, S., et al. (2021). The PRISMA 2020 statement: An updated guideline for reporting systematic reviews. *BMJ, 372*, n71. https://doi.org/10.1136/bmj.n71

[30] Rahman, M. M., Watanobe, Y., & Hasan, M. M. (2024). ChatGPT for education and research: Opportunities, threats, and strategies. *Applied Sciences, 14*(2), 572.

[31] Rudolph, J., Tan, S., & Tan, S. (2023). ChatGPT: Bullshit spewer or the end of traditional assessments in higher education? *Journal of Applied Learning and Teaching, 6*(1), 342-363.

[32] Sallam, M. (2023). ChatGPT utility in healthcare education, research, and practice: Systematic review on the promising perspectives and valid concerns. *Healthcare, 11*(6), 887. https://doi.org/10.3390/healthcare11060887

[33] Selwyn, N. (2019). *Should robots replace teachers? AI and the future of education*. Polity Press.

[34] Siemens, G., & Long, P. (2011). Penetrating the fog: Analytics in learning and education. *EDUCAUSE Review, 46*(5), 30-40.

[35] Snyder, H. (2019). Literature review as a research methodology: An overview and guidelines. *Journal of Business Research, 104*, 333-339. https://doi.org/10.1016/j.jbusres.2019.07.039

[36] Susnjak, T. (2022). ChatGPT: The end of online exam integrity? *arXiv Preprint arXiv:2212.09292*. https://doi.org/10.48550/arXiv.2212.09292

[37] Tlili, A., Alturise, F., Alghamdi, A., Aldahmani, F., Alshahrani, F., Alotaibi, S., Alharthi, N., Aljohani, N., Alswedani, S., Huang, R., Burgos, D., & Bozkurt, A. (2025). Generative artificial intelligence in education: Where we are and where we are going. *Computers and Education: Artificial Intelligence, 8*, 100311.

[38] Tlili, A., Shehata, B., Adarkwah, M. A., Bozkurt, A., Hickey, D. T., Huang, R., & Agyemang, B. (2023). What if the devil is my guardian angel: ChatGPT as a case study of using chatbots in education. *Smart Learning Environments, 10*(1), 15. https://doi.org/10.1186/s40561-023-00237-x

[39] Torraco, R. J. (2005). Writing integrative literature reviews: Guidelines and examples. *Human Resource Development Review, 4*(3), 356-367. https://doi.org/10.1177/1534484305278283

[40] UNESCO. (2023). *Guidance for generative AI in education and research*. UNESCO Publishing.

[41] Vaswani, A., Shazeer, N., Parmar, N., Uszkoreit, J., Jones, L., Gomez, A. N., Kaiser, Ł., & Polosukhin, I. (2017). Attention is all you need. In *Advances in Neural Information Processing Systems* (Vol. 30, pp. 5998-6008).

[43] Vulpe, M. I., Stancu, S., Munteanu, P., & Bărbulescu, R. (2026). Digital divide and technological exclusion: strategies for inclusive ict skill development in adult education. *International Journal of Education, Leadership, Artificial Intelligence, Computing, Business, Life Sciences, and Society*, *5*, 81-92. https://doi.org/10.65222/VIRAL.2026.2.12.32

DOI: https://doi.org/10.65222/VIRAL.2026.7.41.61

[44] Williamson, B., & Eynon, R. (2020). Historical threads, missing links, and future directions in AI in education. *Learning, Media and Technology, 45*(3), 223-235. https://doi.org/10.1080/17439884.2020.1798995

[45] Wilson, H. J., & Daugherty, P. R. (2018). Collaborative intelligence: Humans and AI are joining forces. *Harvard Business Review, 96*(4), 114-123.

[46] Woolf, B. P. (2010). *Building intelligent interactive tutors: Student-centered strategies for revolutionizing e-learning*. Morgan Kaufmann.

[47] Zawacki-Richter, O., Marín, V. I., Bond, M., & Gouverneur, F. (2019). Systematic review of research on artificial intelligence applications in higher education - Where are the educators? *International Journal of Educational Technology in Higher Education, 16*(1), 39. https://doi.org/10.1186/s41239-019-0171-0

[48] Zhao, Y. (2023). Two decades of havoc: A synthesis of criticism against PISA. *Journal of Educational Change, 24*(2), 169-190.