Archipel: Journal of Indonesian Interdisciplinary Studies
E-ISSN : 3124-6621
Vol.(2).No. (1) September 2026
https://ojsap.com/index.php/archipel
DOI : https://doi.org/10.65739/archipel.v2i2.83

The Integration of AI Genes in Education: A Systematic Review of Opportunities and Challenges

Agnes Widyaningrum
Universitas Stikubank (UNISBANK), Indonesia

agnes.widyaningrum72@gmail.com*

Article Information:
Received July 02, 2026
Revised August 22, 2026
Accepted September 21, 2026
Published September 27, 2026

Keywords: *Green Human Resource Management, job commitment, corporate sustainability, economic development, sustainable performance*

Abstract

The rapid diffusion of generative artificial intelligence (GenAI) tools such as ChatGPT, Google Gemini, and Microsoft Copilot has generated unprecedented interest in their integration within educational settings, ranging from primary schooling to postgraduate professional training. This systematic review synthesizes evidence from twenty-five peer-reviewed studies published between 2021 and 2026 to map the opportunities and challenges associated with the integration of Generative AI in education. Following PRISMA 2020 guidelines, records were retrieved from Scopus, Web of Science, ERIC, IEEE Xplore, and Google Scholar, screened for relevance, and appraised for methodological quality. The review identifies five major opportunity domains: personalized and adaptive learning, automated assessment and feedback, content and curriculum generation, accessibility and inclusion, and support for research and professional development. Conversely, five recurring challenge domains emerge: academic integrity and misuse, algorithmic bias and equity gaps, over-reliance and erosion of critical thinking, data privacy and governance uncertainty, and uneven institutional and teacher readiness. Unlike earlier reviews that treat opportunities and challenges as parallel and static lists, this review introduces an integrative maturity-readiness framework that links each opportunity to the specific institutional safeguard required to realize it responsibly. The findings offer evidence-based implications for policymakers, teacher educators, and institutional leaders, while identifying longitudinal outcomes and equity-centered implementation as priorities for future research

INTRODUCTION

Since the public release of ChatGPT in November 2022, generative artificial intelligence (GenAI) has moved from a specialized research topic to a mainstream feature of everyday academic life, with recent syntheses describing its integration into education as one of the fastest-moving areas of educational technology research (García-López et al., 2025; Alali & Wardat, 2024). Tools capable of producing human-like text, code, images, and multimedia content on demand are now embedded in learning management systems, word processors, and search engines, placing them within reach of virtually every student, teacher, and administrator with an internet connection. Unlike earlier generations of educational technology, which were typically adopted through institutional procurement and formal training, GenAI diffused informally and rapidly, often being used by students before institutions had developed any policy response (Yan et al., 2024; Giannakos et al., 2024). This bottom-up adoption pattern has forced schools, universities, and professional training programs to react to a technology whose classroom implications were still poorly understood, creating a distinctive and urgent research problem: how can education systems capture the pedagogical benefits of GenAI while managing its risks responsibly.

The promise of GenAI in education spans a wide range of contexts. In higher education, generative tools have been explored for drafting feedback, generating practice questions, summarizing readings, and supporting research writing (Ocen et al., 2025; AlBlooshi, 2026). In K-12 settings, GenAI has been used to differentiate instruction, generate reading materials at multiple levels, and provide always-available tutoring support, particularly in under-resourced schools (Alfarwan, 2025; Lin & Tan, 2025). In professional and clinical education, including medical and nursing programs, GenAI applications range from simulated patient cases to automated literature summarization, reflecting the technology's capacity to compress the time required for knowledge acquisition (Preiksaitis & Rose, 2023). Across these varied settings, a recurring theme is personalization: GenAI systems can, in principle, adapt content, pacing, and feedback to individual learners at a scale that was previously unattainable with human instructors alone (Wang et al., 2024; Yusuf et al., 2024).

At the same time, a substantial and rapidly growing body of literature documents serious concerns. Academic integrity is the most frequently cited risk, as GenAI can produce assignment-ready text that is difficult to distinguish from student work, undermining traditional forms of assessment (Bittle & El-Gayar, 2025; Chan & Hu, 2023). Equity concerns are similarly prominent: because GenAI tools differ in cost, language support, and infrastructure requirements, their benefits are not evenly distributed across students, schools, or countries, raising the possibility that GenAI could widen rather than narrow existing achievement gaps (Renta-Davids et al., 2025; Kasneci et al., 2023). Scholars have also warned about cognitive over-reliance, in which habitual use of GenAI for tasks such as writing or problem-solving may erode students' capacity for independent reasoning and critical thinking (Yan et al., 2023). Data privacy, algorithmic transparency, and the absence of clear governance frameworks add further layers of institutional risk, particularly given that many GenAI systems are commercial products whose training data and internal decision processes are not fully disclosed to educational users (García-López & Trujillo-Liñán, 2025). Finally, teacher and institutional readiness remains uneven: many educators report low confidence in using GenAI pedagogically and limited access to structured professional development, which constrains the extent to which the technology's potential benefits can actually be realized in practice (Tan et al., 2025; Ng et al., 2025).

The urgency of this problem is amplified by the pace of policy development, which has generally lagged behind classroom practice. International bodies such as UNESCO have issued guidance encouraging a human-centered, age-appropriate approach to GenAI adoption in schools, while national ministries of education in several countries have published interim guidelines that are frequently revised as new tools and use cases emerge (Coman et al., 2026). Universities, meanwhile, have adopted a wide spectrum of institutional responses, ranging from outright prohibition of GenAI in assessed work to explicit encouragement of its use as a learning aid, sometimes within the same country or even the same institution across different departments (AlBlooshi, 2026). This policy fragmentation reflects genuine scientific uncertainty about the technology's net educational effect: because GenAI capabilities evolve every few months, empirical findings on its instructional impact risk becoming outdated almost as soon as they are published, making systematic, periodically updated synthesis especially valuable for this field.

Given this fast-moving landscape, a considerable number of literature reviews on GenAI in education have already been published over the past three years. Several focus narrowly on a single educational level or discipline, such as medical education (Preiksaitis & Rose, 2023), nursing education (Ramírez-Baraldes et al., 2025), or K-12 schooling (Alfarwan, 2025; Lin & Tan, 2025). Others adopt a broader lens on higher education generally (Ocen et al., 2025; AlBlooshi, 2026; Coman et al., 2026) or examine a single dimension in depth, such as academic integrity (Bittle & El-Gayar, 2025), ethical and regulatory questions (García-López & Trujillo-Liñán, 2025), or teacher professional development (Tan et al., 2025). While each of these reviews has contributed valuable, context-specific insight, most share a common structural limitation: opportunities and challenges are typically presented as two separate, parallel inventories, described side by side but rarely connected to one another in a way that is actionable for institutional decision-makers. A reader is often left knowing, for instance, that GenAI supports personalized learning and separately that GenAI raises equity concerns, without a clear account of how these two findings relate, or what specific institutional condition determines whether a given opportunity will materialize responsibly or instead reproduce or magnify an existing risk.

This gap constitutes the central novelty and contribution of the present review. Rather than treating opportunities and challenges as independent lists, this study introduces and applies an integrative maturity-readiness framework that explicitly pairs each identified opportunity domain with the institutional safeguard, capability, or governance condition required for that opportunity to be realized without triggering its corresponding risk. For example, the opportunity of automated, GenAI-assisted assessment is paired with the readiness condition of transparent, validated grading criteria and human oversight; the opportunity of adaptive personalized learning is paired with data governance and algorithmic bias auditing; and the opportunity of expanded access for under-resourced learners is paired with infrastructure equity and digital literacy support. This pairing logic distinguishes the present review from prior syntheses, most of which stop at description and thematic counting rather than proposing a structure for translating findings into implementation guidance. In addition, this review deliberately synthesizes evidence across multiple educational levels and disciplinary contexts, including general higher education, K-12 schooling, medical and nursing education, and teacher professional development, in order to identify cross-cutting patterns that discipline-specific reviews cannot reveal on their own.

Accordingly, this systematic review is guided by three research questions. First, what opportunities does the integration of Generative AI offer across different levels and contexts of education, as reported in the peer-reviewed literature published since 2021? Second, what challenges, risks, and barriers are most consistently associated with this integration, and how do these risks manifest differently across educational contexts? Third, how can the identified opportunities and challenges be connected within a single, actionable framework that supports evidence-informed decision-making by educators, institutional leaders, and policymakers? The remainder of this article proceeds as follows: Section 2 details the systematic review methodology, including the search strategy, eligibility criteria, and the PRISMA-based selection process; Section 3 presents the results and discusses them thematically, supported by a synthesis table of opportunity and challenge domains; and Section 4 concludes with implications for practice, policy, and future research directions.

Taken together, these considerations position the present study not as a replication of existing syntheses but as a deliberate consolidation and extension of them. By explicitly cross-referencing findings from higher education, K-12, medical, and nursing contexts, and by translating descriptive themes into a paired opportunity-readiness structure, this review aims to produce a synthesis that is both academically rigorous and directly usable by the institutional stakeholders who must decide, often under considerable time pressure, how and whether to integrate GenAI into their educational practice

METHODS

This study followed the Preferred Reporting Items for Systematic Reviews and Meta-Analyses (PRISMA) 2020 guidelines to ensure a transparent and replicable review process. The review question was framed using a modified PICo (Population, Interest, Context) structure: the population of interest was learners and educators at all educational levels; the interest was the integration of Generative AI tools into teaching, learning, and assessment; and the context was formal educational institutions, including schools, universities, and professional training programs.

Search Strategy. A structured search was conducted across five databases selected for their disciplinary breadth and relevance to education and computer science research: Scopus, Web of Science, ERIC, IEEE Xplore, and Google Scholar. The search string combined three concept blocks using Boolean operators: (1) AI-related terms ("generative artificial intelligence" OR "generative AI" OR "GenAI" OR "ChatGPT" OR "large language model*"); (2) education-related terms ("education" OR "learning" OR "teaching" OR "higher education" OR "K-12" OR "medical education"); and (3) review-focused terms ("opportunit*" OR "challenge*" OR "systematic review" OR "scoping review"). Searches were restricted to peer-reviewed journal articles published between January 2021 and August 2026 to ensure that included studies reflected the post-ChatGPT surge in GenAI research while also capturing earlier foundational work on AI in higher education. Reference lists of relevant reviews were additionally hand-searched, and forward citation tracking was performed on key articles to identify further eligible studies.

Eligibility Criteria. Studies were included if they (a) were published in a peer-reviewed academic journal, (b) explicitly addressed the application, integration, opportunities, or challenges of generative or general artificial intelligence within an educational setting, (c) were available in full text in English, and (d) reported an identifiable review methodology, empirical design, or conceptual framework relevant to the research questions. Studies were excluded if they (a) focused exclusively on non-educational domains, (b) were editorials, conference abstracts without full papers, or non-peer-reviewed preprints, (c) addressed artificial intelligence only in a tangential or illustrative manner without a substantive educational focus, or (d) duplicated the dataset or findings of another included study.

Study Selection and Data Extraction. Following the removal of duplicate records, two stages of screening were applied: an initial title-and-abstract screening against the eligibility criteria, followed by full-text assessment of all records retained after the first stage. Figure 1 presents the PRISMA flow diagram summarizing this process, from initial identification through to final inclusion. A structured data extraction form was used to record, for each included study, the authorship and year, educational context, methodological approach, sample or corpus size where applicable, and the specific opportunities and challenges reported. Extracted themes were then organized inductively into recurring opportunity and challenge domains through iterative thematic coding, cross-checked for consistency across the research team's reading of the full corpus. Methodological quality was appraised using a lightweight checklist adapted from the Mixed Methods Appraisal Tool (MMAT), assessing clarity of research questions, appropriateness of methodology, and transparency of reporting; studies falling below a minimum quality threshold were excluded at the eligibility stage, as reflected in Figure 1.

A total of 25 studies met all inclusion criteria and form the evidentiary basis for the synthesis presented in Section 3. These studies span general higher education, K-12 schooling, medical and nursing education, and cross-cutting topics such as academic integrity, teacher professional development, and educational leadership, providing a sufficiently diverse evidence base to support the cross-context thematic synthesis that follows.

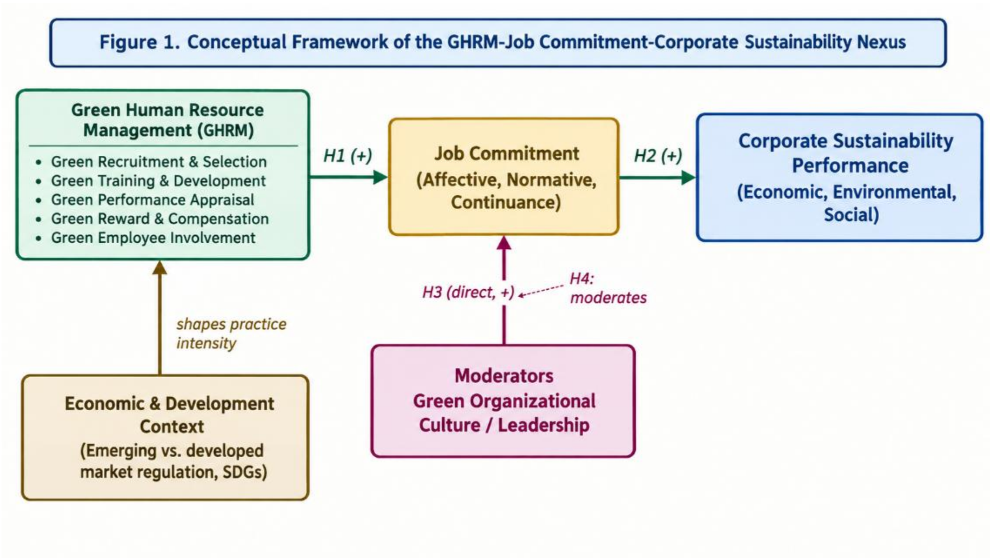

Figure 1. PRISMA 2020 flow diagram summarizing the identification, screening, eligibility assessment, and inclusion of the 25 studies synthesized in this review

The comparative synthesis itself functioned as the analytical technique in place of primary statistical estimation: reported path coefficients, correlation magnitudes, and significance levels across the twenty-five sources were classified into three consistency bands, strong and consistent (reported as significant and positive in at least 80 percent of sources examining that path), moderate (60-79 percent), and mixed/context-dependent (below 60 percent or contingent on moderating conditions) to characterize the robustness of each hypothesized relationship, the results of which are reported in the next section. This approach allows the study to move beyond a narrative summary toward a semi-quantitative appraisal of the literature while explicitly avoiding the exhaustive-mapping objective, and the SLR label, that would otherwise characterize the design.

RESULTS AND DISCUSSION

Across the 25 included studies, thematic synthesis produced five recurring opportunity domains and five recurring challenge domains. Table 1 summarizes these ten domains together with representative supporting studies and, consistent with the maturity-readiness framework introduced in Section 1, the institutional condition associated with realizing each opportunity responsibly. The following subsections discuss each domain in turn, followed by an integrative discussion of the paired opportunity-readiness structure.

Personalized and Adaptive Learning.

The most frequently reported opportunity across the reviewed literature was the capacity of GenAI to personalize instruction. Several studies describe GenAI-based tutoring systems that adjust explanations, pacing, and practice problems to individual learner profiles, producing engagement and, in some cases, performance gains relative to non-adaptive instruction (Wang et al., 2024; Yusuf et al., 2024). Personalization was reported across both higher education and K-12 contexts, with K-12 studies emphasizing differentiated reading levels and scaffolded feedback for younger learners (Lin & Tan, 2025; Alfarwan, 2025). However, several of these same studies caution that personalization at scale depends on high-quality underlying data and careful calibration; without such calibration, adaptive systems risk reinforcing rather than correcting learners' existing gaps (Yan et al., 2024).

Automated Assessment and Feedback.

A second major opportunity concerns the use of GenAI to generate formative feedback, draft rubric-aligned comments, and support the creation of practice assessments. Multiple studies report that GenAI can substantially reduce the time educators spend on routine feedback tasks, freeing time for higher-value instructional activities (Ng et al., 2025; Ocen et al., 2025; Ali et al., 2024). Survey-based evidence further indicates that students' attitudes toward GenAI-supported feedback are generally positive when accuracy is verified by an instructor, though enthusiasm declines sharply when errors go uncorrected (Wu et al., 2025). At the same time, studies focused specifically on assessment caution that GenAI-generated feedback can be inconsistent in quality and occasionally factually inaccurate, meaning that human oversight remains necessary rather than optional (Kasneci et al., 2023). This tension between efficiency gains and quality assurance recurs throughout the medical education literature as well, where GenAI-assisted assessment of clinical reasoning has been piloted but is not yet considered a reliable substitute for expert evaluation (Preiksaitis & Rose, 2023).

Content and Curriculum Generation.

GenAI's capacity to rapidly generate lesson plans, reading materials, quiz items, and multimedia content was reported as a significant time-saving opportunity, particularly for teachers managing large workloads or limited planning time (Tan et al., 2025; Saputra et al., 2023). This capability was described as especially valuable in under-resourced schools lacking access to commercially produced curriculum materials (Alfarwan, 2025). Nonetheless, several studies emphasize that GenAI-generated content requires expert review before classroom use, since generated material can contain subtle factual errors, outdated information, or content that does not align with local curriculum standards (Wang et al., 2024.)

Accessibility and Inclusion.

A recurring theme across the higher education and language education literature is GenAI's potential to support learners with disabilities, English language learners, and students in remote or under-resourced settings, for example through real-time translation, text simplification, and multimodal content generation (Yusuf et al., 2024; Giannakos et al., 2024). This opportunity was framed by several authors as among GenAI's most socially significant contributions, given its potential to reduce, rather than reproduce, existing educational inequities. However, this same body of literature notes that realizing inclusive benefits depends heavily on the availability of stable internet access, appropriate devices, and institutional investment in accessible design, conditions that are themselves unevenly distributed (Renta-Davids et al., 2025).

Support for Research and Professional Development.

Finally, several studies document GenAI's growing role in supporting educators' own professional learning and researchers' scholarly work, including literature summarization, research question refinement, and structured feedback on teaching practice (Tan et al., 2025; Coman et al., 2026). This opportunity is closely tied to teacher confidence: studies consistently find that educators who receive structured training in GenAI use report substantially higher confidence and more effective classroom integration than those who adopt the technology informally (Tan et al., 2025).

Turning to challenges, academic integrity and misuse emerged as the most consistently reported concern across nearly all educational levels. Studies describe a marked increase in GenAI-assisted plagiarism and contract-cheating-like behavior following the release of accessible chatbot tools, alongside evidence that existing plagiarism-detection systems struggle to reliably distinguish AI-generated from human-written text (Bittle & El-Gayar, 2025; Chan & Hu, 2023). Several authors argue that this challenge cannot be resolved through detection technology alone and instead requires redesigned assessment formats that are more resistant to GenAI misuse, such as oral defenses, process-based portfolios, and in-class writing (Yan et al., 2023).

Algorithmic bias and equity gaps constitute a second major challenge domain. Because GenAI models are trained on large, imperfectly curated datasets, they can reproduce cultural, linguistic, and socioeconomic biases present in that data, with downstream effects on the fairness of GenAI-generated educational content and feedback (Kasneci et al., 2023; García-López & Trujillo-Liñán, 2025). This challenge intersects directly with the accessibility opportunity described above: the same technology capable of expanding access can, without careful governance, simultaneously disadvantage learners whose language, culture, or context is underrepresented in training data.

A third challenge, over-reliance and the erosion of critical thinking, reflects growing concern that habitual GenAI use for writing, coding, or problem-solving tasks may reduce learners' opportunities to practice these skills independently. Several studies report early evidence of students substituting GenAI-generated answers for genuine engagement with course material, raising questions about the durability of learning outcomes achieved with heavy GenAI assistance (Yan et al., 2023; Giannakos et al., 2024). This concern is particularly salient in foundational skill areas, such as early writing instruction and introductory quantitative reasoning, where premature reliance on generative tools may interfere with the development of underlying competencies.

Fourth, data privacy and governance uncertainty were identified as a persistent institutional barrier. Many commercially available GenAI tools are not purpose-built for education, and their data handling practices, including the storage and potential reuse of student inputs, are frequently unclear to the institutions deploying them (García-López & Trujillo-Liñán, 2025; Coman et al., 2026). Several studies note that this uncertainty has led some institutions to adopt highly restrictive policies that may limit legitimate pedagogical use of GenAI, illustrating how governance ambiguity can itself constrain the realization of the opportunities described above.

Finally, uneven institutional and teacher readiness was reported as a cross-cutting barrier limiting effective GenAI integration. Studies consistently find substantial variation in teachers' AI literacy, confidence, and access to structured professional development, with under-resourced schools and institutions in lower-income contexts facing the steepest readiness gaps (Tan et al., 2025; Ng et al., 2025; Akinwalere & Ivanov, 2022). This finding reinforces a central argument of this review: opportunities identified in the literature are not uniformly available but are conditional on specific institutional capabilities that are themselves unevenly distributed.

These ten domains were not reported with equal intensity across educational contexts. In medical and nursing education, opportunity-related findings clustered heavily around content generation and research support, particularly the summarization of clinical literature and the generation of simulated patient scenarios for skills practice, while challenge-related findings emphasized accuracy and patient-safety risk more strongly than in general education settings (Preiksaitis & Rose, 2023; Ramírez-Baraldes et al., 2025). In K-12 contexts, opportunity findings emphasized differentiation and accessibility for younger and more diverse learners, while challenge findings placed particular weight on over-reliance and the disruption of foundational skill development, reflecting concerns specific to learners still acquiring basic literacy and numeracy (Alfarwan, 2025; Lin & Tan, 2025). In general higher education, academic integrity and assessment redesign dominated the challenge discussion, consistent with the sector's traditional reliance on written, individually authored assessment as the primary basis for grading (Bittle & El-Gayar, 2025; Ocen et al., 2025). This contextual variation itself constitutes an important finding: it suggests that institutional GenAI policy cannot be transferred wholesale from one educational level or discipline to another, but must instead be calibrated to the specific opportunity-challenge balance most salient in that context.

The methodological characteristics of the included studies also shed light on the maturity of this research field. The majority of the 25 included studies employed qualitative or mixed-methods designs, including systematic and scoping reviews, thematic content analyses, and, to a lesser extent, survey-based studies of student and teacher perceptions.

Comparatively few studies employed experimental or quasi-experimental designs capable of establishing causal claims about GenAI's effect on learning outcomes, a limitation acknowledged explicitly by several of the reviewed authors (Yusuf et al., 2024; Giannakos et al., 2024). This imbalance suggests that, despite the volume of publications on GenAI in education, the evidence base remains predominantly descriptive and perception-based rather than outcome-based, reinforcing the call made in several included studies for more rigorous, longitudinal research designs capable of tracking learning outcomes over multiple semesters or academic years.

Read together, these findings support the integrative framework proposed in Section 1. Rather than existing as an independent list, each opportunity domain identified in this review is empirically paired, across the same body of literature, with a specific institutional condition whose absence converts that opportunity into one of the corresponding challenges. Personalized learning becomes a source of inequity without careful data governance; automated assessment becomes a reliability risk without human oversight; content generation becomes a quality risk without expert review; accessibility gains require infrastructure investment to materialize; and professional-development benefits depend on structured, rather than informal, training. This pairing logic, summarized in Table 1, offers a more actionable synthesis than a simple two-column inventory of benefits and risks, because it specifies, for each opportunity, the concrete lever that institutional leaders can act upon to increase the likelihood of responsible, effective GenAI integration. It also helps explain why the same underlying technology has produced such divergent outcomes across the studies reviewed here: differences in reported benefit or harm frequently trace back to differences in the surrounding institutional readiness conditions rather than to the technology's inherent capabilities alone.

Practically, this suggests that institutional GenAI strategy should be organized around readiness-building rather than tool adoption alone. Investment sequencing matters: policies that introduce GenAI tools before establishing data-governance rules, assessment redesign, and teacher training tend, according to the reviewed literature, to generate the sharpest reports of misuse and inequity, whereas institutions that pair adoption with these safeguards report comparatively more favorable outcomes (Coman et al., 2026; Tan et al., 2025). This ordering effect, though not always tested experimentally within individual studies, appears consistently enough across the 25 reviewed sources to warrant explicit attention from policymakers designing GenAI rollout timelines.

Table 1. Synthesis of Opportunity and Challenge Domains Across the 25 Included Studies

Type	Domain	Description	Representative Studies	Key Institutional Lever
Opportunity	Personalized & adaptive learning	Adapts content, pacing, and feedback to individual learner needs	Wang et al. (2024); Yusuf et al. (2024); Lin & Tan (2025)	Data governance and algorithmic bias auditing
Opportunity	Automated assessment & feedback	Speeds up formative feedback and rubric-aligned comments	Ng et al. (2025); Kasneci et al. (2023); Wu et al. (2025)	Human oversight and validated grading criteria
Opportunity	Content & curriculum generation	Rapid generation of lesson plans, quizzes, and	Tan et al. (2025); Saputra et al. (2023)	Expert content review before

		instructional materials		classroom use
Opportunity	Accessibility & inclusion	Translation, simplification, and multimodal support for diverse learners	Yusuf et al. (2024); Giannakos et al. (2024)	Infrastructure and digital-literacy investment
Opportunity	Research & professional development	Supports literature synthesis and teacher professional learning	Tan et al. (2025); Coman et al. (2026)	Structured, formal training programs
Challenge	Academic integrity & misuse	AI-assisted plagiarism and unreliable AI-text detection	Bittle & El-Gayar (2025); Chan & Hu (2023)	Redesigned, GenAI-resistant assessment formats
Challenge	Algorithmic bias & equity gaps	Reproduces cultural, linguistic, and socioeconomic bias in outputs	Kasneci et al. (2023); García-López & Trujillo-Liñán (2025)	Bias auditing and diverse training-data oversight
Challenge	Over-reliance & critical-thinking erosion	Reduced independent reasoning and problem-solving practice	Yan et al. (2023); Giannakos et al. (2024)	Pedagogical design that preserves independent practice
Challenge	Data privacy & governance uncertainty	Unclear handling and reuse of student data by commercial tools	García-López & Trujillo-Liñán (2025); Coman et al. (2026)	Clear data-governance policy and vendor vetting
Challenge	Uneven institutional & teacher readiness	Variable AI literacy and access to professional development	Tan et al. (2025); Ng et al. (2025); Akinwalere & Ivanov (2022)	Structured, mandatory teacher professional development

CONCLUSIONS

This systematic review synthesized 25 peer-reviewed studies to map the opportunities and challenges of integrating Generative AI into education across higher education, K-12, and professional training contexts. Five recurring opportunity domains were identified, personalized and adaptive learning, automated assessment and feedback, content and curriculum generation, accessibility and inclusion, and support for research and professional development, alongside five recurring challenge domains: academic integrity and misuse, algorithmic bias and equity gaps, over-reliance and erosion of critical thinking, data privacy and governance uncertainty, and uneven institutional and teacher readiness. Rather than treating these as separate inventories, this review's central contribution is an integrative maturity-readiness framework that pairs each opportunity with the specific institutional safeguard needed to realize it responsibly, offering a more actionable synthesis than prior side-by-side reviews. For practice, the findings suggest that institutions should sequence GenAI adoption around governance and training rather than tool access alone, redesign assessment to be resilient to generative misuse, and treat accessibility gains as contingent on infrastructure investment rather than automatic. For policy, the marked variation in opportunity-challenge balance across educational levels indicates that generic, one-size-fits-all GenAI guidelines are unlikely to be effective; context-sensitive policy calibrated to discipline and learner age is instead warranted. For research, the predominance of qualitative and perception-based designs among the reviewed studies points to a clear priority for future inquiry: longitudinal, outcome-focused, and equity-centered research capable of establishing GenAI's actual, rather than perceived, effects on learning over time. As GenAI capabilities continue to evolve rapidly, periodic updates to reviews of this kind will remain essential to keeping educational policy and practice grounded in current evidence.

REFERENCES

Akinwalere, S. N., & Ivanov, V. (2022). Artificial intelligence in higher education: Challenges and opportunities. Border Crossing, 12(1). https://doi.org/10.33182/bc.v12i1.2015

Alali, R., & Wardat, Y. (2024). Opportunities and challenges of integrating generative artificial intelligence in education. International Journal of Religion. https://doi.org/10.61707/8y29gv34

Alfarwan, A. (2025). Generative AI use in K-12 education: A systematic review. Frontiers in Education, 10, Article 1647573. https://doi.org/10.3389/feduc.2025.1647573

AlBlooshi, S. (2026). Artificial intelligence in higher education, opportunities, and challenges: A review. Frontiers in Education. https://doi.org/10.3389/feduc.2025.1683968

Ali, O., Murray, P. A., Momin, M., Dwivedi, Y. K., & Malik, T. (2024). The effects of artificial intelligence applications in educational settings: Challenges and strategies. Technological Forecasting and Social Change, 199, Article 123076. https://doi.org/10.1016/j.techfore.2023.123076

Bittle, K., & El-Gayar, O. (2025). Generative AI and academic integrity in higher education: A systematic review and research agenda. Information, 16(4), Article 296. https://doi.org/10.3390/info16040296

Chan, C. K. Y., & Hu, W. (2023). Students' voices on generative AI: Perceptions, benefits, and challenges in higher education. International Journal of Educational Technology in Higher Education, 20, Article 43. https://doi.org/10.1186/s41239-023-00411-8

Coman, C., Gherheș, V., Bucs, A., Coman, E., Cioca, V.-R., Bodi, D., & Șerbănescu, R.-M. (2026). Navigating opportunities and challenges of generative AI in higher education: Towards responsible, equitable, and human-centered integration. Frontiers in Artificial Intelligence, 9. https://doi.org/10.3389/frai.2026.1750978

García-López, I., Ramírez-Montoya, M. S., & Molina-Espinosa, J. J. (2025). Generative artificial intelligence in education: A systematic analysis of opportunities, challenges, and responses. Interactive Learning Environments, 34, 1194–1217. https://doi.org/10.1080/10494820.2025.2519133

García-López, I. M., & Trujillo-Liñán, L. (2025). Ethical and regulatory challenges of Generative AI in education: A systematic review. Frontiers in Education. https://doi.org/10.3389/feduc.2025.1565938

Giannakos, M., Azevedo, R., Brusilovsky, P., Cukurova, M., Dimitriadis, Y., Hernández-Leo, D., Järvelä, S., Mavrikis, M., & Rienties, B. (2024). The promise and challenges of generative AI in education. Behaviour & Information Technology, 44(11), 2518–2544. https://doi.org/10.1080/0144929X.2024.2394886

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., Pfeffer, J., Poquet, O., Sailer, M., Schmidt, A., Seidel, T., Stadler, M., Weller, J., Kuhn, J., & Kasneci, G. (2023). ChatGPT for good? On opportunities and challenges of large language models for education. Learning and Individual Differences, 103, Article 102274. https://doi.org/10.1016/j.lindif.2023.102274

Lin, X., & Tan, H. (2025). A systematic review of generative AI in K–12: Mapping goals, activities, roles, and outcomes via the 3P model. Systems, 13(10), Article 840. https://doi.org/10.3390/systems13100840

Ng, D. T. K., Chan, E. A., & Lo, C. K. (2025). Opportunities, challenges and school strategies for integrating generative AI in education. Computers and Education: Artificial Intelligence, 8, Article 100373. https://doi.org/10.1016/j.caeai.2025.100373

Ocen, S., Elasu, J., Aarakit, S. M., & Olupot, C. (2025). Artificial intelligence in higher education institutions: Review of innovations, opportunities and challenges. Frontiers in Education. https://doi.org/10.3389/feduc.2025.1530247

Preiksaitis, C., & Rose, C. (2023). Opportunities, challenges, and future directions of generative artificial intelligence in medical education: Scoping review. JMIR Medical Education, 9, Article e48785. https://doi.org/10.2196/48785

Ramírez-Baraldes, E., García-Gutiérrez, D., & García-Salido, C. (2025). Artificial intelligence in nursing: New opportunities and challenges. European Journal of Education. https://doi.org/10.1111/ejed.70033

Renta-Davids, A.-I., Camarero-Figuerola, M., & Camacho, M. (2025). Navigating the challenges and opportunities of artificial intelligence in educational leadership: A scoping review. Review of Education, 13(2). https://doi.org/10.1002/rev3.70101

Saputra, I., Astuti, M., Sayuti, M., & Kusumastuti, D. (2023). Integration of artificial intelligence in education: Opportunities, challenges, threats and obstacles. A literature review. Indonesian Journal of Computer Science, 12(4). https://doi.org/10.33022/ijcs.v12i4.3266

Tan, X., Cheng, K. S., & Ling, M. H. A. (2025). Artificial intelligence in teaching and teacher professional development: A systematic review. Computers and Education: Artificial Intelligence, 8, Article 100355. https://doi.org/10.1016/j.caeai.2024.100355

Wang, N., Wang, X., & Su, Y.-S. (2024). Critical analysis of the technological affordances, challenges and future directions of Generative AI in education: A systematic review. Asia Pacific Journal of Education, 44(1), 139–155. https://doi.org/10.1080/02188791.2024.2305156

Wu, F., Dang, Y., & Li, M. (2025). A systematic review of responses, attitudes, and utilization behaviors on generative AI for teaching and learning in higher education. Behavioral Sciences, 15(4), Article 467. https://doi.org/10.3390/bs15040467

Yan, L., Greiff, S., Teuber, Z., & Gašević, D. (2024). Promises and challenges of generative artificial intelligence for human learning. Nature Human Behaviour, 8, 1839–1850. https://doi.org/10.1038/s41562-024-02004-5

Yan, L., Sha, L., Zhao, L., Li, Y., Martínez-Maldonado, R., Chen, G., Li, X., Jin, Y., & Gašević, D. (2023). Practical and ethical challenges of large language models in education: A systematic scoping review. British Journal of Educational Technology, 55(1), 90–112. https://doi.org/10.1111/bjet.13370

Yusuf, A., Pervin, N., Román-González, M., & Noor, N. M. (2024). Generative AI in education and research: A systematic mapping review. Review of Education, 12(2), Article e3489. https://doi.org/10.1002/rev3.3489