Int. J. of AI and Education Technology

♦ ISSN: 2835-2432 Vol: 03, No. 01, 2024 ♦ PP. 20–28
Article DOI: https://doi.org/10.54216/IJAIET.030103



FULL LENGTH ARTICLE

Generative AI Chatbots in Education Technology: A Critical Review of Feedback, Assessment, and Governance

Mahshid Manouchehri[1,*]

[1] Amity University Dubai Campus, Dubai International Academic City, Dubai, 345019, United Arab Emirates

Email: mmanouchehri@amityuniversity.ae

Received: January 03, 2024 Revised: March 05, 2024 Accepted: May 29, 2024 ⋆ Corresponding author

ABSTRACT

Generative artificial intelligence (AI) chatbots have become a disruptive education technology because they combine dialogue, content generation, reasoning support, and immediate feedback in a single interface. Their rapid adoption has created a strategic challenge for schools, universities, and professional education providers: the same tools that can expand formative support and learner agency can also weaken assessment validity, privacy protection, academic integrity, and equitable participation. This review synthesises peer-reviewed literature published from 2020 to 2023 to examine how generative AI chatbots should be understood within education technology. The paper focuses on three interdependent areas: feedback and learner support, assessment redesign, and institutional governance. The review finds that the most defensible use of generative AI is not tool substitution but learning-design augmentation: chatbots can support explanation, drafting, questioning, and formative feedback when teachers define learning goals, evidence requirements, and acceptable use conditions. The main risks are not limited to plagiarism or inaccurate outputs; they also include hidden inequity, over-reliance, weakened disciplinary judgement, and policy-practice misalignment. The paper proposes a responsible adoption framework that links pedagogical affordances, assessment redesign, AI literacy, and governance controls. The contribution is a publication-ready conceptual synthesis for institutions seeking to adopt generative AI in education technology without reducing learning to automated content production.

Keywords: Generative artificial intelligence ▪ Education technology ▪ AI chatbots ▪ Formative feedback ▪ Assessment redesign ▪ Academic integrity ▪ AI governance

1. INTRODUCTION

Generative AI chatbots have altered the trajectory of education technology more quickly than most earlier digital tools. Learning management systems, adaptive platforms, and analytics dashboards generally entered institutions through procurement, pilot studies, and gradual policy development. Generative AI systems entered through everyday public access. Students and teachers could use them before institutions had developed policies, assessment rules, or professional development routines. This mismatch between adoption speed and governance readiness explains why generative AI has become a central issue in education technology strategy rather than only a new classroom application.

The debate should not be reduced to whether chatbots should be allowed or prohibited. Recent AI-in-education scholarship shows that the field has already moved from tool-level experimentation toward questions of pedagogy, agency, human judgement, and institutional implementation [1, 2, 3]. Generative AI strengthens this shift because it is not confined to a single function. It can draft explanations, translate language, propose examples, critique text, simulate dialogue,

© 2024 The Author(s). Published by ASPG. This article is licensed under CC BY 4.0.

summarise readings, generate code, and produce assessment-like responses. These features can support learning, but they also challenge the meaning of student work and teacher judge-ment.

This review examines generative AI chatbots as an education technology for feedback, assessment, and governance. The fo-cus is intentionally specific. Feedback and assessment are the areas where generative AI is most immediately visible to stu-dents and teachers, while governance determines whether use remains educationally purposeful, equitable, and accountable. The paper therefore asks: how should educational institutions conceptualise, adopt, and govern generative AI chatbots in ways that protect learning quality and assessment validity?

The contribution is threefold. First, the paper synthesises literature published between 2020 and 2023 without relying on post-2023 references. Second, it organises the literature into a review framework connecting pedagogical affordances, assessment design, and institutional governance. Third, it proposes a responsible adoption model that can guide schools, universities, and professional education providers.

2. REVIEW SCOPE AND METHOD

This paper is a critical narrative review supported by a structured conceptual synthesis. It is not presented as a meta-analysis because the field contains heterogeneous arti-cle types, including systematic reviews, conceptual papers, early empirical studies, policy-oriented analyses, and domain-specific evaluations. It applies a transparent selection logic while limiting the synthesis to literature directly relevant to AI education technology, chatbot applications, feedback, as-sessment, and governance.

The review includes peer-reviewed journal articles from 2020 to 2023. Earlier foundational work is intentionally excluded to match the publication window and to keep the synthesis focused on recent AI education technology developments. The final synthesis draws on studies from *Computers and Education: Artificial Intelligence*, *International Journal of Educational Technology in Higher Education*, *Education and Information Technologies*, *Education Sciences*, *Applied Sci-ences*, *SN Social Sciences*, *Learning, Media and Technology*, and other indexed journals. The included literature covers AI-in-education roles, chatbot applications, online higher education, K-12 education, assessment redesign, ethics, gen-erative AI, and institutional governance.

The synthesis proceeded in four analytical steps. First, re-view papers on AI in education were used to establish the broader trajectory of the field [2, 4, 5, 6, 7]. Second, chatbot-specific reviews and generative AI studies were examined to identify pedagogical affordances and practical concerns [8, 9, 10, 11, 12, 13, 14, 15]. Third, student, teacher, and institutional perspectives were analysed to capture adoption conditions [16, 17, 18, 19, 20]. Fourth, assessment-focused and governance-oriented discussions were examined to clar-ify implications for validity, disclosure, privacy, and learner support [21, 22, 23, 24, 25, 26]. The result is a conceptual review rather than a statistical aggregation.

3. RELATED WORK SYNTHESIS

Table 1 summarises the main contribution of prior work used in this review. The table is included to make clear how the literature informs the present synthesis rather than treating all sources as equivalent.

3.1 The changing role of AI in education

The recent AI-in-education literature shows a clear transition from algorithmic support tools to broader educational ecosystems. Hwang et al. [1] described the opportunities and challenges of AI in education, including intelligent tutoring, adaptive learning, analytics, and human-centred design. Chen et al. [2] emphasised the gap between application growth and theoretical development. This gap becomes more important with generative AI because its outputs often appear fluent and convincing even when they are pedagogically shallow or factually unreliable.

Ouyang and Jiao [3] offered a useful distinction between AI-directed, AI-supported, and AI-empowered learning. Generative AI chatbots operate across all three categories. When students ask a chatbot for an answer, the system may become AI-directed. When students use it to request hints, examples, counterarguments, or formative comments, it becomes AI-supported. When teachers design tasks that require students to critique, verify, and improve AI output, the technology can support AI-empowered learning. The same tool can therefore produce very different educational outcomes depending on task design. This helps explain why Rospigliosi [29] argued that educators need to ask better questions about ChatGPT rather than treat it as a fixed instructional solution.

3.2 Chatbots before and after generative AI

Chatbots were used in education before the emergence of large language models, but earlier systems were usually nar-row in scope. They answered frequently asked questions, supported administrative guidance, or simulated scripted di-alogue. Okonkwo and Ade-Ibijola [8] reviewed educational chatbot applications and showed that chatbots were already being used for teaching support, learner engagement, and administrative interaction. Generative AI changed the scale of this discussion because the systems are not restricted to pre-authored flows. They can produce open-ended responses and adapt linguistically to user prompts.

This flexibility is also the source of risk. Earlier chatbots were limited, but their limitations were often visible. Generative chatbots produce polished language that can mask uncertainty, bias, or weak reasoning. For education technology, this means that usability is no longer sufficient. Institutions need to judge the quality of outputs, the appropriateness of use, and the evidential value of student work produced with AI assistance.

4. PEDAGOGICAL AFFORDANCES

4.1 Formative feedback and tutoring support

The most defensible educational use of generative AI chat-bots is formative support. Feedback is valuable when it is timely, specific, understandable, and linked to learning goals. Teachers often struggle to provide such feedback at scale, especially in large classes or writing-intensive courses. Gen-erative AI can help students receive immediate comments on drafts, request examples, test explanations, and explore

Table 1. Selected 2023 studies and their contributions to generative AI and education technology.

Reference	Focus area	Main contribution	Relevance to this review	Year
Crompton and Burke [7]	Higher education AI	Mapped applications of AI in higher education and identified the need for clearer implementation strategies.	Frames generative AI as part of a wider institutional AI agenda.	2023
Lo [9]	ChatGPT rapid review	Synthesised early evidence on ChatGPT's educational capabilities, uses, and risks.	Provides the baseline for the feedback, assessment, and integrity discussion.	2023
Rahman and Watanobe [10]	Education and research	Discussed opportunities, threats, and strategic responses for ChatGPT use.	Supports the balanced treatment of benefits and governance controls.	2023
Chiu [11]	Policy and practice	Analysed generative AI implications for educational practice, policy, and research direction.	Informs the policy-practice alignment argument.	2023
Chan and Hu [16]	Student perceptions	Reported students' perceived benefits and challenges of generative AI in higher education.	Adds the learner voice to the adoption discussion.	2023
Su and Yang [13]	Adoption framework	Proposed a framework for applying generative AI in education.	Supports the paper's responsible integration model.	2023
Cotton et al. [21]	Academic integrity	Examined academic integrity risks and institutional responses in the ChatGPT era.	Directly informs the assessment validity section.	2023
Sullivan et al. [22]	Integrity and learning	Connected ChatGPT with academic integrity and student learning concerns.	Helps distinguish misconduct risk from legitimate learning support.	2023
Rudolph et al. [23]	Assessment change	Questioned whether traditional assessments remain defensible under generative AI.	Supports the argument for redesign rather than simple detection.	2023
Pradana et al. [24]	Review and bibliometrics	Reviewed ChatGPT in education using literature and bibliometric evidence.	Locates the topic within a rapidly expanding research field.	2023
Grassini [15]	Future of education	Reviewed potential and consequences of AI and ChatGPT in educational settings.	Supports the paper's balanced opportunity-risk framing.	2023
Dempere et al. [17]	Higher education impact	Analysed ChatGPT's impact on higher education institutions.	Reinforces the need for institutional readiness and guidance.	2023
Cooper [27]	Science education	Explored ChatGPT responses to science education prompts.	Provides domain evidence for accuracy and pedagogical judgement risks.	2023
Kohnke et al. [28]	Language learning	Reviewed ChatGPT affordances and drawbacks for language teaching and learning.	Shows how feedback and writing support depend on learner competence.	2023
Michalon and Camacho-Zuniga [20]	Competency development	Described classroom integration of ChatGPT to strengthen enduring competencies.	Supports use cases that require student reasoning rather than answer copying.	2023
van den Berg and du Plessis [19]	Teacher education	Discussed lesson planning, critical thinking, and openness in teacher education.	Links generative AI to teacher capability and learning design.	2023
Adeshola and Adepoju [14]	Opportunities and challenges	Reviewed major benefits and limitations of ChatGPT in education.	Supports the paper's structured opportunity-risk synthesis.	2023
Fuchs [18]	Higher education pedagogy	Examined whether NLP models in higher education are beneficial or risky.	Reinforces the need for contextual and pedagogical judgement.	2023
Mills et al. [25]	Open educational practice	Proposed open educational practices as a response to generative AI disruption.	Supports transparent, iterative, and student-facing governance.	2023
Halaweh [26]	Responsible implementation	Proposed strategies for responsible ChatGPT implementation in education.	Directly informs institutional governance and rollout recommendations.	2023

alternative ways of understanding difficult concepts. Lo [9], Rahman and Watanobe [10], Su and Yang [13], and Adiguzel et al. [30] identify feedback, learner support, content generation, and academic assistance as important opportunities for large language models in education.

However, formative support is not automatically meaningful. A chatbot can provide fluent but generic comments. It can also encourage surface-level revision if students simply accept its phrasing. Effective use therefore depends on instructional design. A productive activity might ask students to submit a draft, request chatbot feedback against a rubric, identify which feedback is valid, justify accepted changes, and reflect on what was learned. In this form, the chatbot becomes a dialogic resource rather than an invisible ghost writer.

4.2 Learning design, questioning, and metacognition

Generative AI can support questioning and metacognition when the task requires students to explain their reasoning. For example, a chatbot can be used to generate counterexamples, ask Socratic questions, compare solution strategies, or simulate peer critique. These uses align with AI-supported and AI-empowered learning rather than simple answer pro-duction [3]. The value emerges when students must evaluate and regulate the AI interaction.

The literature also suggests that the pedagogical effects of AI depend heavily on context, including learner readiness, teacher guidance, and disciplinary expectations [16, 31, 18]. Reviews of AI in online higher education identify prediction, recommendation, assessment, and personalised support as common functions, but also show that pedagogical integration varies widely [6]. In K-12 contexts, the challenges include teacher readiness, curriculum fit, student maturity, and eq-uity of access [5]. These differences warn against universal claims about generative AI. The same system may support advanced learners in one context while creating dependency or confusion in another.

4.3 Specialised domains and professional education

Generative AI has attracted particular attention in professional education because it can answer exam-style questions, produce explanations, and mimic domain reasoning. Recent education-focused reviews show that assessment tasks based mainly on recall, generic explanations, or conventional short answers are increasingly vulnerable to automation [9, 10]. This does not mean that chatbots can replace professional judgement. Instead, it shows that assessment must ask students to justify decisions, evaluate evidence, and explain the limits of generated responses.

Professional education therefore needs tasks that evaluate judgement, justification, ethical reasoning, and contextual application; this is consistent with science, physics, language, and teacher-education studies that treat ChatGPT as a prompt for reasoning rather than a replacement for expertise [27, 32, 28, 19]. A medical, engineering, business, or teacher education programme may use generative AI as a case partner, but students should still demonstrate why a recommendation is valid, what assumptions are being made, what evidence is missing, and what risks are present. In this sense, generative AI can raise the standard of assessment if institutions redesign tasks around defensible reasoning rather than output production.

5. ASSESSMENT REDESIGN AND ACADEMIC INTEGRITY

5.1 Why prohibition is insufficient

Many institutional responses to generative AI began with academic integrity concerns. This response is understandable because generative AI can produce essays, code, summaries, and answers that resemble student work. However, prohibition alone is insufficient for three reasons. First, detection is uncertain and can create unfair accusations. Second, students may use AI outside institutional systems. Third, blanket prohibition prevents legitimate uses such as brainstorming, feedback, accessibility support, and language refinement.

Lo's rapid review [9] shows that early literature on ChatGPT in education already identified both learning opportunities and integrity risks. Rahman and Watanobe [10] similarly frame ChatGPT as a source of opportunities, threats, and strategic responses. Integrity-focused papers further argue that assessment design, disclosure rules, and teacher guidance are more defensible than relying only on detection [21, 22, 23, 33]. The implication is that institutions need differentiated assessment policies. The key question is not simply whether AI was used, but whether its use is consistent with the learning outcome being assessed.

5.2 Assessment validity in the chatbot era

Assessment validity depends on whether the task measures the intended construct. If an assignment is intended to measure writing fluency, independent synthesis, or domain reasoning, unacknowledged chatbot use can undermine validity. If the task is intended to measure AI critique, prompt design, evidence evaluation, or revision judgement, chatbot use may be central to validity. This distinction changes how educators should design assessments.

A practical approach is to classify assessment tasks into four categories. First, AI-restricted tasks assess unaided recall, closed-book reasoning, or live performance. Second, AI-permitted tasks allow limited support but require disclosure. Third, AI-integrated tasks require students to use AI as part of a structured process and submit evidence of interaction and revision. Fourth, AI-critical tasks assess the ability to evaluate, challenge, and improve AI outputs. Figure 2 presents this continuum.

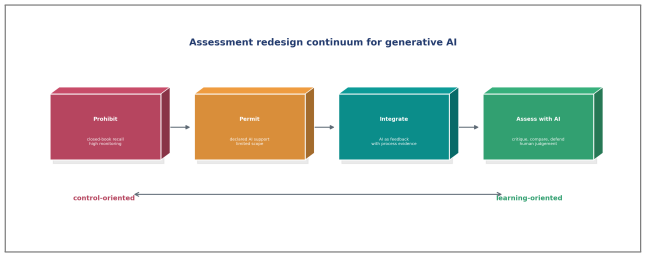


Figure 2. Assessment redesign continuum. Institutions can move from prohibition to structured integration when the learning outcome requires critique, judgement, and reflective use.

5.3 Evidence of process

Generative AI makes final products less reliable as sole evidence of learning. Assessment design should therefore include process evidence, a position consistent with open and transparent educational practice responses to generative AI [25, 29]. This may include annotated drafts, oral defence, version history, reflective commentary, prompt logs, peer discussion, or in-class application. The goal is not to surveil students but to restore the connection between work submitted and learning demonstrated.

Process evidence also supports fairness. Students who use AI responsibly can show how they used it. Students who choose not to use AI can demonstrate their reasoning in other ways. Teachers can assess the learning process rather than only the polished output. This aligns with the argument that AI use should be guided by pedagogy rather than treated as a separate technology problem.

6. ETHICAL AND GOVERNANCE CONCERNS

6.1 Accuracy, bias, and over-reliance

The most visible risk of generative AI is inaccurate output. Chatbots can fabricate references, produce plausible but incorrect claims, and hide uncertainty behind confident language. In educational settings, this can mislead learners who lack the expertise to evaluate responses. However, accuracy is only one part of the problem. Bias, cultural assumptions, and uneven representation can also shape the quality of outputs. Governance-focused AI education research stresses that educational institutions need transparency, accountability, human oversight, and clear policy-practice alignment [34, 11, 26].

Over-reliance is a related pedagogical risk. If students use chatbots to replace effort, struggle, and revision, they may lose opportunities to develop expertise. The risk is not that students receive help; learning has always involved help. The risk is that students outsource the cognitive work that the task was designed to develop. Teachers therefore need to make explicit which parts of the learning process may be supported by AI and which parts must remain student-owned.

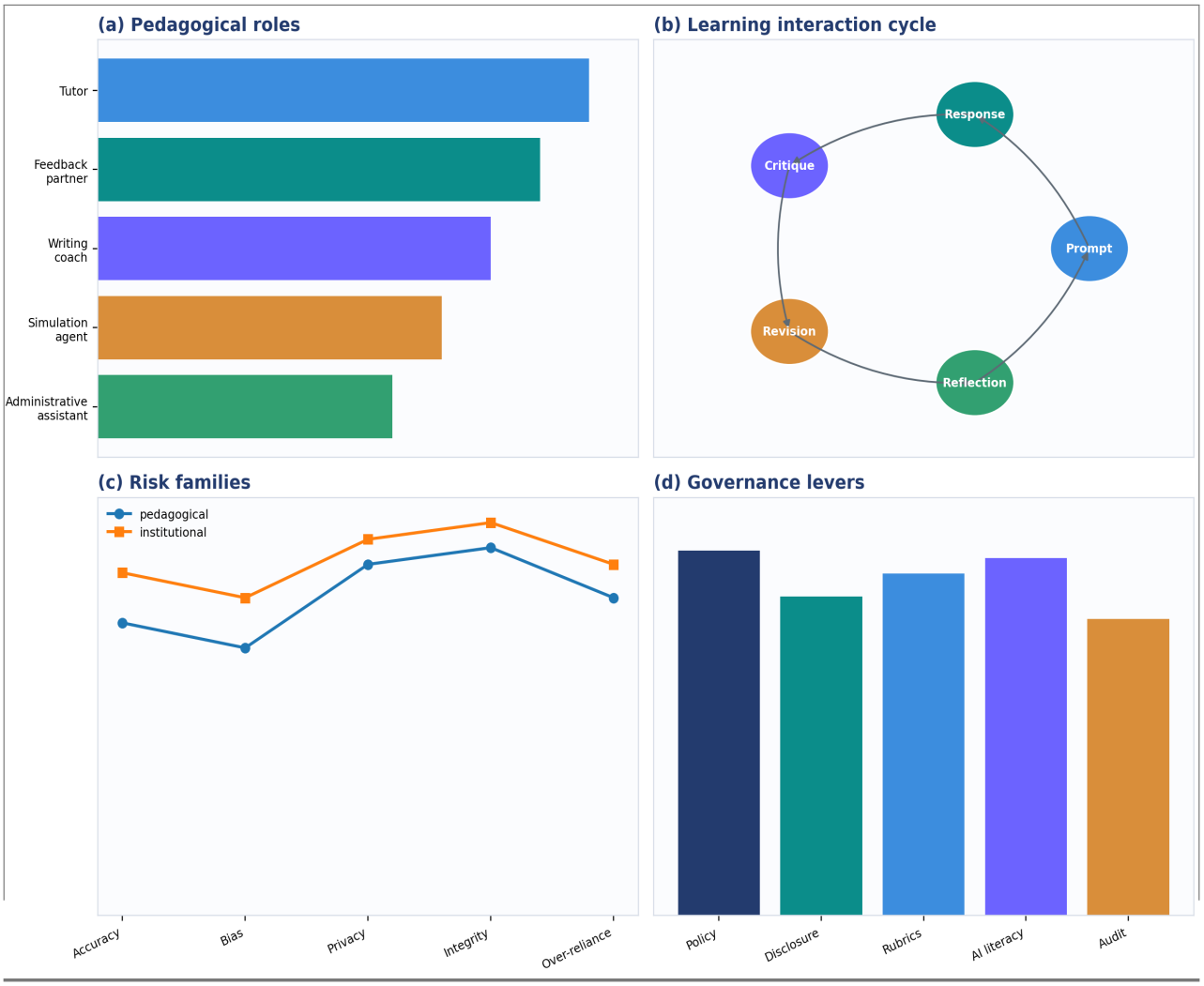


Figure 1. Four-part taxonomy for generative AI chatbots in education technology: pedagogical roles, interaction cycle, risk families, and governance levers. The figure is a conceptual synthesis of the review

6.2 Privacy and data protection

Generative AI chatbots raise privacy concerns because in-teractions may include personal data, student writing, as-sessment materials, or institutional information. Even when tools appear free or convenient, data handling may not align with institutional obligations.

Education technology gover-nance should therefore include data minimisation, approved platforms, contractual review, and clear guidance on what students and staff should not upload.

Privacy concerns are also equity concerns. Students may differ in access to paid tools, reliable internet, institutional guidance, and digital literacy. A policy that assumes equal access can unintentionally widen gaps. Responsible adoption requires institutions to define baseline access, provide alter-natives, and ensure that AI-enhanced tasks do not privilege students with better tools or stronger informal support.

6.3 Policy-practice alignment

Institutional policy often lags behind classroom practice. A policy may state that AI use must be disclosed, but teach-

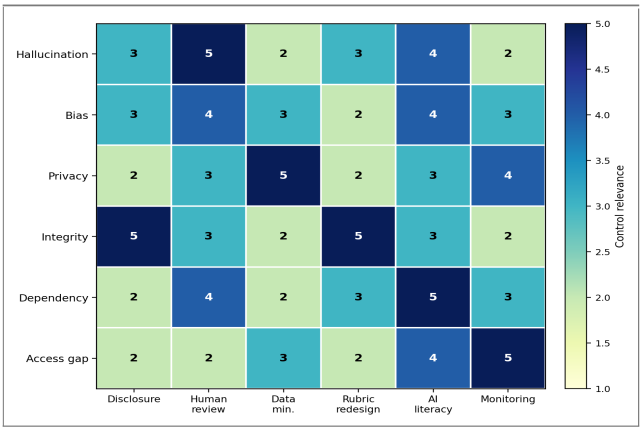


Figure 3. Governance risk-control matrix for educational AI chatbots. Values indicate relative relevance of each control to each risk category.

ers may not have rubrics for evaluating disclosure. A policy may ban AI use, but students may still use it privately. A policy may encourage innovation, but teachers may not re-ceive training. The governance problem is therefore not only writing rules; it is aligning policy, assessment design, teacher practice, student guidance, and quality assurance.

Chiu [11] describes generative AI as a development with direct implications for educational practice, policy, and research direction; Halaweh [26] similarly stresses responsible implementation strategies, while Pradana et al. [24] show how rapidly the literature on ChatGPT in education expanded during 2023. In education, this means that governance must be both practical and pedagogical. Rules must be understandable, enforceable, and connected to learning design. Otherwise, institutions risk creating symbolic policies that do not shape actual behaviour.

7. A RESPONSIBLE ADOPTION FRAMEWORK

7.1 Principle 1: Pedagogy before tools

The first principle is that educational goals should determine AI use. A course team should begin with learning outcomes, assessment evidence, and student needs. Only then should it decide whether a chatbot can support explanation, feedback, practice, drafting, or critique. This reverses the common adoption pattern in which a tool is introduced first and pedagogy is adjusted afterward.

A pedagogically grounded approach also helps distinguish between acceptable and unacceptable use. For example, AI-generated examples may be helpful in a formative activity but unacceptable in a final independent analysis unless disclosed and critically evaluated. The same technology can therefore be legitimate or illegitimate depending on the purpose of the task.

7.2 Principle 2: Human judgement remains central

Generative AI can assist with feedback, but it should not become the sole judge of learning. Teachers need to remain responsible for assessment standards, ethical decisions, and student support. Students also need to learn how to judge AI outputs. This human-centred position is consistent with the broader AI-in-education literature, which warns against treating AI systems as neutral or self-sufficient educational agents [34, 1].

Human judgement is especially important when AI feedback conflicts with teacher expectations, disciplinary norms, or local context. Students should be encouraged to ask: Is the AI response accurate? Is the reasoning complete? Does it match the assignment criteria? What evidence supports it? What has been omitted? These questions turn chatbot use into an exercise in evaluative judgement.

7.3 Principle 3: Assessment must include transparency

Disclosure should be normalised rather than treated only as confession. If AI use is permitted, students should state what tool was used, for what purpose, and how the output was evaluated. Disclosure can be brief for low-stakes formative work and more detailed for high-stakes assignments. The aim is to make the boundary between student contribution and machine assistance visible.

Transparency also protects teachers. Without disclosure norms, teachers must guess whether work reflects student ability. With disclosure norms, teachers can evaluate the quality of student decision-making. A student who simply accepts AI output without critique may show weak judgement, while a student who interrogates and revises AI output may show strong learning.

7.4 Principle 4: Governance must be iterative

Generative AI tools change rapidly. A policy written once and left unchanged will become outdated. Institutions should treat AI governance as an iterative cycle involving monitoring, feedback, revision, and quality assurance. Figure 4 summarises the implementation roadmap proposed by this review.

8. IMPLICATIONS FOR TEACHERS, STUDENTS, AND INSTITUTIONS

8.1 Implications for teachers

Teachers need practical guidance rather than abstract warnings. They should be supported to redesign assignments, build AI-use statements, create rubrics that evaluate process evidence, and design activities that require critique of AI output. Professional development should include hands-on exploration of chatbot strengths and weaknesses. Teachers who do not understand the tools are unlikely to design effective tasks or detect inappropriate reliance.

At the same time, teachers should not be expected to solve institutional governance problems alone. Course-level innovation requires programme-level consistency. If one course bans AI, another encourages it, and a third ignores it, students receive mixed signals. Institutions need shared baseline policies that still allow disciplinary adaptation.

8.2 Implications for students

Students need AI literacy. This includes technical understanding, but also ethical, epistemic, and practical judgement. Students should learn that a chatbot is not an authority; it is a probabilistic language system that can support exploration but must be checked. They should also learn how to cite or disclose AI assistance, how to protect their data, and how to avoid over-reliance.

AI literacy should be integrated into normal learning rather than offered as a one-time orientation. Students in writing courses need guidance on drafting and revision. Students in computing need guidance on code generation and debugging. Students in professional programmes need guidance on domain judgement and risk. AI literacy is therefore not a generic add-on; it must be contextualised by discipline.

8.3 Implications for institutions

Institutional leaders should avoid both panic and uncritical enthusiasm. The review literature suggests that generative AI has genuine educational potential, but only when adoption is governed by learning design, equity, privacy, and quality assurance. Institutions should establish approved-use categories, provide guidance for disclosure, support teachers with examples, and monitor student experience.

A useful institutional policy should answer six questions: Which tools are approved? What data may be entered? Which uses are permitted, restricted, or prohibited? How should students disclose assistance? How will assessment be redesigned? How will the institution evaluate impact? Without answers to these questions, adoption becomes fragmented and risk management becomes reactive.

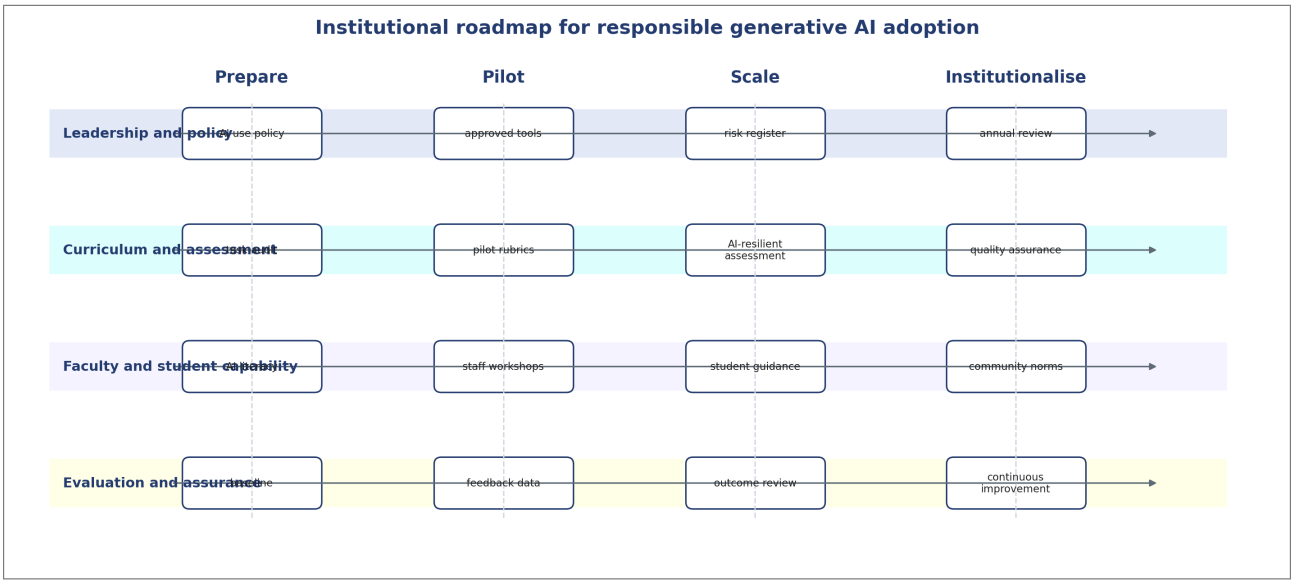


Figure 4. Institutional roadmap for responsible generative AI adoption in education technology. The roadmap links policy, curriculum, capability development, and evaluation across preparation, pilot, scale-up, and institutionalisation phases.

9. DISCUSSION

This review suggests that generative AI chatbots are best un-derstood as socio-technical education technologies. Their ed-ucational effect is not determined by model capability alone. It is shaped by task design, teacher judgement, student lit-eracy, institutional rules, and the quality of feedback loops. This explains why the same chatbot can appear beneficial in one setting and harmful in another.

The first major finding is that generative AI expands formative support but weakens product-only assessment. The more a task relies on final written output as sole evidence of learning, the more vulnerable it becomes. The solution is not to aban-don writing or independent work, but to redesign assessment so that process, judgement, and defence become visible.

The second finding is that AI literacy is central to respon-sible adoption. Students and teachers need to understand prompting, verification, disclosure, bias, data protection, and the limits of generated responses. AI literacy should not be left to informal experimentation because informal learning may reproduce inequity. It should become part of education technology strategy.

The third finding is that governance must be operational. General statements about ethical AI are insufficient if they do not translate into rubrics, approved tools, staff training, student guidance, and quality assurance. The governance matrix proposed in Figure 3 shows that different risks require different controls. No single measure addresses all concerns.

The fourth finding is that generative AI requires a shift from detection culture to design culture. Detection tools may have a role, but they cannot carry the burden of educational in-tegrity. More sustainable responses include assessment re-design, transparent use rules, oral or in-class components, reflective commentary, and tasks that evaluate critique of AI outputs.

10. FUTURE RESEARCH AGENDA

The literature up to 2023 is rich in conceptual and early review work but still limited in longitudinal evidence. Future research should examine how generative AI affects learning outcomes over time, not only immediate student satisfaction or task completion. Studies should compare different levels of AI integration, examine effects on writing development and disciplinary reasoning, and evaluate whether AI-supported feedback improves revision quality.

Another priority is equity. Research should examine whether generative AI narrows or widens educational gaps. Students with stronger prior knowledge may benefit more because they can evaluate outputs, while less prepared students may be more vulnerable to inaccuracies. Institutions need evidence on how scaffolding, teacher guidance, and access models affect these differences.

Assessment research should move beyond plagiarism. Future studies should test AI-resilient assignment designs, disclosure practices, oral defence formats, process portfolios, and rubric-based evaluation of AI critique. The central question should be how assessment can remain valid when generative support is available.

Governance research should examine implementation mod-els. Which policy designs are understandable to students? Which staff-development models change teaching practice? How should institutions audit AI use without creating surveil-lance cultures? How can quality assurance systems review AI-integrated assessment? These questions are essential for moving from emergency response to mature adoption.

11. CONCLUSION

Generative AI chatbots have become one of the most conse-quential education technologies of the current period. Their significance lies not only in content generation, but in their ca-pacity to reshape feedback, assessment, learner support, and institutional governance. This review shows that the educa-tional value of generative AI depends on how it is embedded in learning design. Chatbots can support feedback, question-ing, and revision, but they can also undermine assessment

validity and student agency when used without transparency or pedagogical purpose.

The review argues for a responsible adoption model built around pedagogy, human judgement, transparency, and iterative governance. Educational institutions should not respond with simple prohibition or unconditional adoption. They should define permitted uses, redesign assessment, teach AI literacy, protect student data, and evaluate impact. The core challenge is to ensure that generative AI strengthens learning rather than merely accelerating production. When integrated carefully, generative AI chatbots can become part of a more reflective and feedback-rich education technology ecosystem; when integrated poorly, they can weaken trust, integrity, and educational quality.

REFERENCES

[1] G.-J. Hwang, H. Xie, B. W. Wah, and D. Gasevic, "Vision, challenges, roles and research issues of artificial intelligence in education," *Computers and Education: Artificial Intelligence*, vol. 1, p. 100001, 2020.

[2] X. Chen, H. Xie, D. Zou, and G.-J. Hwang, "Application and theory gaps during the rise of artificial intelligence in education," *Computers and Education: Artificial Intelligence*, vol. 1, p. 100002, 2020.

[3] F. Ouyang and P. Jiao, "Artificial intelligence in education: The three paradigms," *Computers and Education: Artificial Intelligence*, vol. 2, p. 100020, 2021.

[4] K. Zhang and A. B. Aslan, "AI technologies for education: Recent research and future directions," *Computers and Education: Artificial Intelligence*, vol. 2, p. 100025, 2021.

[5] H. Crompton and D. Burke, "Artificial intelligence in K-12 education," *SN Social Sciences*, vol. 2, no. 7, p. 113, 2022.

[6] F. Ouyang, L. Zheng, and P. Jiao, "Artificial intelligence in online higher education: A systematic review of empirical research from 2011 to 2020," *Education and Information Technologies*, vol. 27, pp. 7893–7925, 2022.

[7] H. Crompton and D. Burke, "Artificial intelligence in higher education: the state of the field," *International Journal of Educational Technology in Higher Education*, vol. 20, no. 1, p. 22, 2023.

[8] C. W. Okonkwo and A. Ade-Ibijola, "Chatbots applications in education: A systematic review," *Computers and Education: Artificial Intelligence*, vol. 2, p. 100033, 2021.

[9] C. K. Lo, "What is the impact of ChatGPT on education? a rapid review of the literature," *Education Sciences*, vol. 13, no. 4, p. 410, 2023.

[10] M. M. Rahman and Y. Watanobe, "ChatGPT for education and research: Opportunities, threats, and strategies," *Applied Sciences*, vol. 13, no. 9, p. 5783, 2023.

[11] T. K. F. Chiu, "The impact of generative AI on practices, policies and research direction in education: a case of ChatGPT and Midjourney," *Interactive Learning Environments*, 2023.

[12] D. Baidoo-Anu and L. Owusu Ansah, "Education in the era of generative artificial intelligence (AI): Understanding the potential benefits of ChatGPT in promoting teaching and learning," *Journal of AI*, vol. 7, no. 1, pp. 52–62, 2023.

[13] J. Su and W. Yang, "Unlocking the power of ChatGPT: A framework for applying generative AI in education," *ECNU Review of Education*, vol. 6, no. 3, pp. 355–366, 2023.

[14] I. Adeshola and A. P. Adepoju, "The opportunities and challenges of ChatGPT in education," *Interactive Learning Environments*, 2023.

[15] S. Grassini, "Shaping the future of education: Exploring the potential and consequences of AI and ChatGPT in educational settings," *Education Sciences*, vol. 13, no. 7, p. 692, 2023.

[16] C. K. Y. Chan and W. Hu, "Students' voices on generative AI: perceptions, benefits, and challenges in higher education," *International Journal of Educational Technology in Higher Education*, vol. 20, no. 1, p. 43, 2023.

[17] J. Dempere, K. P. Modugu, A. Hesham, and L. K. Ramasamy, "The impact of ChatGPT on higher education," *Frontiers in Education*, vol. 8, p. 1206936, 2023.

[18] K. Fuchs, "Exploring the opportunities and challenges of NLP models in higher education: is Chat GPT a blessing or a curse?" *Frontiers in Education*, vol. 8, p. 1166682, 2023.

[19] G. van den Berg and E. du Plessis, "ChatGPT and generative AI: Possibilities for its contribution to lesson planning, critical thinking and openness in teacher education," *Education Sciences*, vol. 13, no. 10, p. 998, 2023.

[20] B. Michalon and C. Camacho-Zuñiga, "ChatGPT, a brand-new tool to strengthen timeless competencies," *Frontiers in Education*, vol. 8, p. 1251163, 2023.

[21] D. R. E. Cotton, P. A. Cotton, and J. R. Shipway, "Chatting and cheating: Ensuring academic integrity in the era of ChatGPT," *Innovations in Education and Teaching International*, 2023.

[22] M. Sullivan, A. Kelly, and P. McLaughlan, "ChatGPT in higher education: Considerations for academic integrity and student learning," *Journal of Applied Learning and Teaching*, vol. 6, no. 1, pp. 31–40, 2023.

[23] J. Rudolph, S. Tan, and S. Tan, "ChatGPT: Bullshit spewer or the end of traditional assessments in higher education?" *Journal of Applied Learning and Teaching*, vol. 6, no. 1, pp. 342–363, 2023.

[24] M. Pradana, H. P. Elisa, and S. Syarifuddin, "Discussing ChatGPT in education: A literature review and bibliometric analysis," *Cogent Education*, vol. 10, no. 2, p. 2243134, 2023.

[25] A. Mills, M. Bali, and L. Eaton, "How do we respond to generative AI in education? open educational practices give us a framework for an ongoing process," *Journal of Applied Learning and Teaching*, vol. 6, no. 1, pp. 16–30, 2023.

[26] M. Halaweh, "ChatGPT in education: Strategies for responsible implementation," *Contemporary Educational Technology*, vol. 15, no. 2, p. ep421, 2023.

[27] G. Cooper, "Examining science education in ChatGPT: An exploratory study of generative artificial intelligence," *Journal of Science Education and Technology*, vol. 32, no. 3, pp. 444–452, 2023.

[28] L. Kohnke, B. L. Moorhouse, and D. Zou, "ChatGPT for language teaching and learning," *RELC Journal*, vol. 54, no. 2, pp. 537–550, 2023.

[29] P. A. Rospigliosi, "Artificial intelligence in teaching and learning: what questions should we ask of ChatGPT?" *Interactive Learning Environments*, vol. 31, no. 1, pp. 1–3, 2023.

[30] T. Adiguzel, M. H. Kaya, and F. K. Cansu, "Revolutionizing education with AI: Exploring the transformative potential of ChatGPT," *Contemporary Educational Technology*, vol. 15, no. 3, p. ep429, 2023.

[31] A. Strzelecki, "To use or not to use ChatGPT in higher education? a study of students' acceptance and use of technology," *Interactive Learning Environments*, 2023.

[32] P. Bitzenbauer, "ChatGPT in physics education: A pilot study on easy-to-implement activities," *Contemporary Educational Technology*, vol. 15, no. 3, p. ep430, 2023.

[33] M. Farrokhnia, S. K. Banihashem, O. Noroozi, and A. Wals, "A SWOT analysis of ChatGPT: Implications for educational practice and research," *Innovations in Education and Teaching International*, 2023.

[34] B. Williamson and R. Eynon, "Historical threads, missing links, and future directions in AI in education," *Learning, Media and Technology*, vol. 45, no. 3, pp. 223–235, 2020.