

Towards Responsible Generative AI in Education: Developing an Ethical and Human-Centric Framework

Dr. Sruti Sruba Bharali
Discipline of Computer Science
Krishna Kanta Handiqui State
Open University
Guwahati, India

Abstract: This paper observes the ever-growing role of generative artificial intelligence (Gen AI) in education. It highlights the emergence of tools like ChatGPT and the usage of said tools by all stakeholders including students, teachers, administrators and educational institutions. It explores the various policies on using Gen AI in education worldwide. The study also discusses the ethical considerations with responsible use of Gen AI and advocates for the need of a policy dedicated to regulate the use of Gen AI in education. This study also proposes a human-centric framework for responsible use of Gen AI in education.

Keywords: Artificial intelligence, Generative AI, Education, Ethical, Framework

1. INTRODUCTION

Generative artificial intelligence (Gen AI) has emerged as one of the most transformative developments in the education sector in recent years. Unlike artificial intelligence that had been around for decades but mainly focused in areas involving pattern recognition, model creation and analysis of data, Gen AI can create new text, audio, video, programming code and educational materials transforming the education sector. Although generative models have been around for a few years, the widespread application in education sector was accelerated by the public release of conversational AI systems such as OpenAI ChatGPT. The accessibility of these tools has enabled both students and teachers to experiment with AI-generated content in teaching and learning materials.

The emergence of ChatGPT and similar tools like Google Gemini, Microsoft Copilot, Claude, Perplexity and other AI-powered platforms has changed the whole teaching learning environment. Students use these tools to find explanations of complex topics, brainstorm ideas for presentations, generate texts, translate content in own local language for better understanding and even develop programming code. Teachers too use these tools for preparing lesson plans, generating quizzes and assignments and creating personalized learning materials. Despite these applications, the integration of Gen AI into education sector creates important ethical, pedagogical and governance challenges. There is an increasing need to ensure that Gen AI is integrated in education environment in a responsible, transparent, inclusive, and human-centric manner. In response to this need, this paper proposes an ethical and human centric framework for the responsible adoption and use of Gen AI in education. The proposed framework aims to provide a responsible ethical framework for Gen AI use in education.

The contributions of the paper are as follows:

- (1) To examine the current policies associated with Gen AI in education sector.
- (2) To identify key ethical principles relevant to the responsible adoption of Gen AI in educational settings.
- (3) To develop a comprehensive ethical framework for the human centric use of Gen AI in education.

The rest of the paper is as follows. Section 2 offers a conceptual background of Gen AI. Section 3 explores the existing guidelines and policies on Gen AI use in education. Section 4 examines the many ethical challenges associated with responsible adoption of Gen AI in educational settings. Section 5 presents the responsible framework for Gen AI use in educational institutions. Section 6 presents the conclusion and future scope.

2. CONCEPTUAL BACKGROUND

2.1 Generative AI

Gen AI refers to a class of artificial intelligence technologies that are capable of creating new content including text, audio, images, simulations and video [1]. The rapid rise of technology and the access to internet has made Gen AI models increasingly accessible to students, teachers, researchers, administrators and other stakeholders in education. Gen AI models have the ability to generate new content as outputs as a response to prompts given by the user. These outputs are not always retrieved directly from a predefined database but are based on models that are trained on large amounts of data called Large Language Models (LLMs). LLMs are trained to predict the likelihood of word sequences or generate new text based on a given input [2]. LLMs are AI models that form the technological basis of Gen AI applications. The popularity of Gen AI tools like ChatGPT, Google Gemini, Microsoft Copilot, Claude, and many other similar systems prove the huge potential of Gen AI in education.

2.2 Generative AI in Education

In education, students use Gen AI to understand complex concepts, improve grammar, brainstorm ideas, learning programming and for receiving feedback to their queries on assignments etc. [3]. Teachers on the other hand, use Gen AI to assist with curriculum design, lesson planning, content development, creating quizzes and assignments, grading quizzes and codes and other time-consuming tasks [4]. Administrators use Gen AI models to generate speeches, create feedback questions and reports and also summarize various events held by institutions. Gen AI has the ability to generate digital content in different formats like text, audio, video and image generation [5]. Text generation is one of the most widely used application of Gen AI which includes

generating explanations to complex topics, summaries, reports, essays and answers to questions. In the field of education, these tools can be used by teachers to provide personalized explanations or tutoring service for students at various levels of understanding. It can also be used to provide feedback on assignment and quizzes and for the development of teaching learning materials for classes. Image generation can be used by teachers and students alike to generate explanations, illustrations, simulations and animations to understand complex topics. Image generation models help to create images from text or modify existing images as per the instructions of the user. Audio generation includes generation of audio content which could help in accessibility support for differently abled students. Audio generation includes creation of speech, music, sounds and other forms of audio. In the context of education, Gen AI models can be used to support language learning, translation, narration and for development of multimedia content for teaching learning resources. Another important application of Gen AI is code generation. Gen AI conversational systems can help students with AI-assisted learning of various programming languages like Python, Java etc. Gen AI systems can generate programs, identify syntax and run time errors in the programming code and suggest corrective approaches. Code generation can also be used by teachers to explain the complex topics with lots of examples based on the programming skill level of the students.

2.3 Gen AI – Human Interaction in Education

Teachers, students and administrators all interact with Gen AI systems in different ways to obtain different goals. For teachers, Gen AI is mostly used for making lesson plans, development of content, numerous examples and question banks, grading of quizzes and assignments among many other academic duties. Gen AI can be used to reduce the workload of teachers by carrying out the more routine tasks and allowing teachers to devote their time and attention to creativity, research, mentoring and working on individual student needs. For students, Gen AI is mostly used as conversational tutors, for writing assistants, coding assistants, getting personalized feedback, and getting explanations on complex topics. Students also use the tools to brainstorm and for translating course materials to local languages and many other activities. Though there needs to be a balance on using Gen AI tools as overdependence on AI may reduce thinking, problem solving, reasoning and creativity of the students. For administrators too, Gen AI is mostly used for making survey questions, generating reports and analysis of the reports. They also use it for creating beautiful presentations and generating speeches. But the outputs produced for students, teachers and administrators by these Gen AI systems require critical evaluation because the models may generate factually incorrect responses, reflect biases in their training data, or provide information without sufficient contextual understanding [6]. As such, use of the Gen AI tools require not only human oversight but also ethical framework for appropriate use of Gen AI in education.

3. EXISTING GUIDELINES ON GEN AI USE

Prior to the release of ChatGPT, national and international institutions and governments had been developing frameworks for use and adoption of AI systems across various sectors including education. However, after the release of Gen AI models, most of the responses in the form of policies from

government and educational organizations have been to either urgently formulate new regulations or to ban use of Gen AI altogether to discourage faculty and students from using it. Though there are a lot of guidelines at individual institutional level from educational institutions or research specific guidelines from journal publishing companies but very limited number of guidelines that encompass all major criteria on Gen AI use in education.

Some of the visible international and national guidelines on Gen AI use come from UNESCO, OCED or are country specific. The UNESCO guidelines offers concrete recommendations for policy-makers and educational institutions on how the uses of Gen AI tools can be designed to protect human agency and genuinely benefit learners, teachers and researchers [7]. The guidelines propose a framework with six perspectives based on domain of knowledge and expected outcomes and also the requirements of users. The first potential of Gen AI use is for research where it can act as an AI advisor or as data explorer and literature reviewer. The second potential of Gen AI is to facilitate teachers and teaching by helping in curriculum or course designer and also application of Gen AI as chatbot or teaching assistants. The third potential is for using Gen AI as a coach for learning languages and arts like conversation practice for a new language or for coaching basic arithmetic or coding. Gen AI can also be used as an advisor to facilitate inquiry or project-based learning. The next potential of Gen AI is to support learners with special needs or with learning difficulties by giving access to AI-powered accessibility tools. The OECD, Recommendation of the Council on Artificial Intelligence guidelines include policies for governments on investing in AI research and development [8]. The policy asks governments to invest and promote innovation in trustworthy AI that focuses on AI-related social, legal and ethical implications and policy issues. The Council also recommends governments to either invest or encourage investment on AI research that is free of harmful bias and respect privacy and data protection.

The India AI Governance Guidelines present a broad, cross-sector framework that is highly relevant to the responsible use of AI and Gen AI in education. Rather than proposing a separate AI law, the Guidelines advocate a balanced, flexible and pro-innovation governance approach, built around seven sutras: Trust, People First, Innovation over Restraint, Fairness and Equity, Accountability, Understandable by Design, and Safety, Resilience and Sustainability [9]. For the education sector, this implies that AI systems should support rather than replace human decision making, provide understandable information about their use and limitations, and avoid discrimination or exclusion. For Gen AI specifically, the guidelines highlight issues that are directly applicable to educational institutions, including misinformation, deep fakes, transparency, privacy, bias, copyright, and accountability. The framework recommends risk assessment based on the nature and severity of potential harm, stronger protection for vulnerable groups such as children, content authentication and provenance mechanisms, and appropriate human oversight at critical decision points. In an educational context, the seven principles can provide the foundation for institutional policies covering responsible Gen AI use by teachers and students, disclosure of AI-generated content, protection of personal data, academic integrity, human review of AI outputs, copyright compliance, and mechanisms for reporting AI-related harm. However, the institutions and organization need to translate these principles into specific education policies and guidelines. Hence, there is a need of a universal education

framework for responsible use of Gen AI in the education sector as proposed in this paper.

4. ETHICAL CHALLENGES TO GEN AI IN EDUCATION

Some of the key ethical challenges to be kept in mind while designing the framework for responsible Gen AI use in education are: academic integrity, bias and fairness, privacy, data protection, transparency, intellectual property and copyright [10].

4.1 Academic integrity

Gen AI has created new possibilities for academic dishonesty by enabling students to generate answers, essays, computer programs, and presentations. Educational institutions should be able to distinguish between plagiarism, unauthorized AI assistance, and legitimate AI-supported learning. Students should be taught how to acknowledge AI assistance appropriately and how to critically evaluate and transform AI-generated material rather than simply reproducing it. Academic integrity policies should provide explicit guidance on when to declare use of Gen AI models for AI-generated text, images, code etc. Such policies would encourage responsible AI use from students.

4.2 Bias and Fairness

Bias and fairness are significant ethical challenges in the use of Gen AI in education [11]. Since Gen AI systems are trained on large datasets that may contain social, cultural, gender, and linguistic biases, their outputs could unintentionally include such existing inequalities [12]. Hence, the Gen AI systems used in education should be regularly evaluated for bias, tested across diverse learner groups, and designed to promote fairness, inclusivity, and equitable access to educational opportunities.

4.3 Privacy and Data Protection

Privacy and data protection are also major ethical challenges in the use of Gen AI in education. AI systems may process sensitive student data, personal information, learning analytics, and prompts containing confidential academic or personal details. When such information is entered into third-party AI platforms, there may be risks related to unauthorized access, data storage, profiling, or secondary use of information [13]. Educational institutions must therefore establish clear guidelines for data minimization, informed consent, secure handling of information, and responsible use of third-party AI tools. Students and teachers should also be encouraged to avoid entering personally identifiable or confidential information into AI systems unless appropriate safeguards are in place.

4.4 Transparency

Students and teachers should have transparency on how AI-generated answers are produced, including the limitations and possibility of errors or biases in the output [14]. When AI supports assessment or grading, institutions should ensure that the basis for AI-supported decisions is understandable, traceable and subject to human oversight.

4.5 Accountability and Responsibility

The use of Gen AI in education requires accountability to be established among students, teachers, administrators, educational institutions and AI platform providers [15]. The clearly mentioned responsibilities would ensure that Gen AI supports education without weakening human oversight and accountability.

4.6 Human Agency and Autonomy

Human agency and autonomy are important ethical considerations as overdependence on AI tools may reduce student's motivation to think independently, potentially weakening critical thinking, creativity, problem-solving, and decision-making skills. Similarly, the increasing use of Gen AI raises concerns about teacher displacement, particularly when AI is used to automate teaching, assessment, and student support. Therefore, Gen AI should complement rather than replace human educators, with meaningful human oversight and opportunities for learners to exercise independent judgment [16].

4.7 Intellectual Property and Copyright

Gen AI systems are often trained on large datasets that may contain copyrighted books, articles, images, videos, and other educational resources. As such, questions may arise regarding the ownership and copyright status of AI-generated content, particularly when students or teachers use AI tools to create assignments, teaching materials, presentations etc. [17]. Therefore, educational institutions should promote responsible use of AI by encouraging proper attribution, respecting copyright restrictions, verifying the originality of AI-generated content, and establishing clear guidelines for the use of copyrighted educational materials.

5. PROPOSED ETHICAL FRAMEWORK FOR GEN AI USE IN EDUCATION

The proposed ethical framework takes into consideration each of the ethical challenge for responsible Gen AI use in education. The framework includes five dimensions for responsible Gen AI use in education: Transparency, Human Oversight, Privacy and Data protection, Fairness and Responsibility. The functionality of each of these dimensions in the framework are described below.

Dimension 1: Transparency

The first component of the framework is transparency. All stakeholders including students, teachers and administrators should disclose when Gen AI is being used. A mandatory disclosure is required from all when AI-assisted academic work is used. The information to be disclosed should also contain to what percentage of the work is contributed by the author and what percentage is Gen AI generated so that there could be no doubt about ownership and copyright infringement. The framework also includes an optional criterion to explain how the AI model generated that output.

Dimension 2: Human Oversight

The second component of the framework is human agency and oversight. The framework proposes that the final decision-making authority rests with the humans. The students would be responsible for their submitted work and the teachers would oversee the Gen AI generated content submitted. The goal of this component is to make sure humans are always in the loop of the academic process.

Dimension 3: Privacy and Data Protection

The next component of the framework is regarding privacy and data protection. The framework proposes that all users should limit the amount of personal data when using Gen AI models. There needs to be institutional policies to limit the amount of personal data shared with Gen AI systems. Personal data of students, teachers like address, medical information etc. should not be shared while using Gen AI for

various tasks like generating reports. Hence, this part of the framework applies appropriate data protection mechanisms.

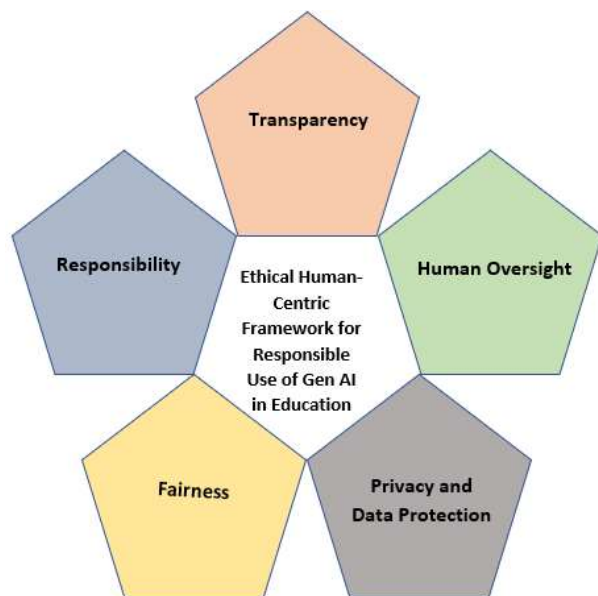


Figure. 1 Ethical Human-Centric Framework for Responsible Use of Gen AI in Education

Dimension 4: Fairness

The next component of the framework deals with fairness and bias of the Gen AI models. The Gen AI models that are used should be regularly evaluated for potential bias and to ensure that their outputs are fair, accurate and inclusive. AI generated content, feedback, and assessments may reflect biases present in training data or algorithmic design. Teachers and institutions should examine the AI outputs for gender, socioeconomic, cultural, racial, linguistic, or other forms of bias. Human oversight is necessary for identifying if similar quality of support and opportunities are provided to learners from different backgrounds. The AI systems should also be sensitive to linguistic and cultural diversity and AI-generated outputs should accommodate different languages, regional contexts and cultural practices.

Dimension 5: Responsibility

The last component of the framework defines the responsibilities of students, teachers, administrators, institutions and AI providers. Students should use Gen AI as a tool to support learning rather than as a substitute for their own intellectual effort. They are responsible for following institutional guidelines regarding AI use, maintaining academic integrity, and disclosing the use of AI where required. Teachers are responsible for guiding students in the appropriate and ethical use of Gen AI. They should design assessments that encourage critical thinking and authentic learning, clearly communicate expectations regarding acceptable AI use, and help students develop AI literacy. Teachers should also critically evaluate AI-generated content before incorporating it into teaching materials and remain aware of potential issues such as bias, inaccuracies, and hallucinations. Administrators should develop and implement clear policies and procedures governing the use of Gen AI. Their responsibilities include supporting faculty and students through training, addressing issues related to academic integrity and data privacy, and monitoring the implementation of AI practices across the institution. Educational institutions should establish governance frameworks, provide access to

appropriate AI tools and infrastructure, protect student and staff data, and ensure that AI adoption does not increase existing inequalities. Institutions should also establish mechanisms for accountability, grievance redressal, and continuous evaluation of the educational and social impacts of AI. AI providers should design and deploy tools that are safe, transparent, reliable, and inclusive. They should take reasonable measures to reduce bias, protect user privacy and data, provide clear information about the capabilities and limitations of their systems, and offer mechanisms to report any harm done from their outputs.

6. CONCLUSION

Gen AI offers significant opportunities to enhance teaching, learning, assessment, and academic support in education; however, its responsible adoption requires careful attention to ethical, social, and educational concerns. This paper proposes a framework based on five key dimensions: Transparency, Human Oversight, Privacy and Data Protection, Fairness, and Responsibility. Together, these dimensions provide a practical foundation for guiding the ethical and human oversight on use of Gen AI in educational settings. The framework emphasizes that AI should complement rather than replace human judgment and should be implemented in ways that promote trust, equity, accountability, and student well-being. Future research can empirically validate the proposed framework across different educational contexts and develop measurable indicators for assessing responsible Gen AI adoption.

7. REFERENCES

- [1] Jonnala, A. 2024. Harnessing the power of generative AI. *International Journal of Computer Science and Engineering* 11, 8 (August 2024), 37–44. <https://doi.org/10.14445/23488387/IJCSE-V11I8P105>.
- [2] Chang, Y., Wang, X., Wang, J., Wu, Y., Yang, L., Zhu, K., et al. 2024. A survey on evaluation of large language models. *ACM Transactions on Intelligent Systems and Technology* 15, 3 (2024), 1–45.
- [3] Dwivedi, A., Pandey, P., and Mni, N. 2026. Revolutionizing education: Analyzing the transformative potential of generative AI in learning environment. *International Journal of Computer Information Systems and Industrial Management Applications* 18, 17s (2026), 518–532.
- [4] Memon, T. D. and Kwan, P. 2025. A collaborative model for integrating teacher and GenAI into future education. *TechTrends* 69, 6 (2025), 1192–1206.
- [5] Rizvi, N. 2024. From chalkboards to chatbots: The role of generative AI in education. *International Journal of Multicultural and Multireligious Understanding* 12 (2024), 683–696.
- [6] Fui-Hoon Nah, F., Zheng, R., Cai, J., Siau, K., and Chen, L. 2023. Generative AI and ChatGPT: Applications, challenges, and AI-human collaboration. *Journal of Information Technology Case and Application Research* 25, 3 (2023), 277–304.
- [7] Holmes, W. and Miao, F. 2023. *Guidance for Generative AI in Education and Research*. UNESCO Publishing.
- [8] OECD. 2025. *Recommendation of the Council on Artificial Intelligence*. OECD/LEGAL/0449.

- [9] Ministry of Electronics and Information Technology (MeitY), Government of India. 2025. India AI Governance Guidelines: Enabling Safe and Trusted AI Innovation. <https://static.pib.gov.in/WriteReadData/specificdocs/documents/2025/nov/doc2025115685601.pdf>.
- [10] Ren, X. and Wu, M. L. 2025. Examining teaching competencies and challenges while integrating artificial intelligence in higher education. *TechTrends* 69, 3 (2025), 519–538.
- [11] Yin, Z., Chinta, S. V., Wang, Z., Gonzalez, M., and Zhang, W. 2024. FairAIED: Navigating fairness, bias, and ethics in educational AI applications. *arXiv preprint arXiv:2407.18745*.
- [12] Gordon, R. 2023. Large Language Models Are Biased. Can Logic Help Save Them? MIT CSAIL.
- [13] Ye, X., Yan, Y., Li, J., and Jiang, B. 2024. Privacy and personal data risk governance for generative artificial intelligence: A Chinese perspective. *Telecommunications Policy* 48, 10 (2024), 102851.
- [14] Raman, R., Achuthan, K., and Nedungadi, P. 2026. Generative AI integration in education: Theoretical review and future directions informed by the ADO framework. *Information* 17, 3 (2026), 241.
- [15] García-López, I. M. and Trujillo-Liñán, L. 2025. Ethical and regulatory challenges of generative AI in education: A systematic review. *Frontiers in Education* 10 (June 2025), 1565938.
- [16] Cukurova, M. 2026. Agency as a system property in human–AI interaction in education. *British Journal of Educational Technology* 57, 4 (2026), 1065–1070.
- [17] Chesterman, S. 2025. Good models borrow, great models steal: Intellectual property rights and generative AI. *Policy and Society* 44, 1 (2025), 23–37.