International Journal of Web of Multidisciplinary Studies
(Peer-Reviewed, Open Access, Fully Refereed International Journal)
website: *http://ijwos.com*
Vol.02 No.12.
OPEN ACCESS
E-ISSN : 3049-2424
DOI: https://doi.org/10.71366/ijwos

Generative AI in Higher Education: A Systemic Analysis of Pedagogical Adaptation, Ethical Integrity, and Institutional Governance

Arihant Raj[1], Sagar Choudhary[2]
[1] *B.Tech Student, Computer Science and Engineering, Quantum University, Roorkee, India*
[2] *Assistant Professor, Computer Science and Engineering, Quantum University, Roorkee, India.*

Article Info

Article History:

Published:18 Dec 2025

Publication Issue:
Volume 2, Issue 12
December-2025

Page Number:
369-378

Corresponding Author:
Arihant Raj

Abstract:

The rapid spread of Generative AI (GenAI) across universities is reshaping how knowledge is produced and evaluated. Large Language Models (LLMs) now participate in tasks once understood as exclusively human—analyzing information, generating ideas, and constructing written work. This development unsettles long-standing assumptions about how students learn and how institutions assess that learning. Drawing from sociotechnical and dataset-focused perspectives, this paper examines the multiple ways AI integrates into educational environments and proposes a structured framework for categorizing student–AI interactions. We also outline a comprehensive institutional protocol for evaluating the impact of GenAI on academic integrity and assessment design. Ultimately, we argue that the challenges posed by GenAI stem not from student misconduct alone but from deeper curricular dependencies on product-based assessment.

Keywords: Generative AI, Higher Education, Academic Integrity, Large Language Models, Algorithmic Bias, Assessment Security

1. INTRODUCTION

GenAI is rapidly transforming higher education, influencing both day-to-day academic work and broader institutional practices. LLMs now produce text with a level of fluency and responsiveness that enables them to assist—or fully substitute for—student cognitive effort. Unlike earlier educational technologies that primarily stored or retrieved information, contemporary models generate original written material, participate in dialogue, and adapt to prompts with ease. As a result, traditional markers of student learning, such as take-home essays or unproctored assignments, are increasingly misaligned with what these tools can produce.

Survey data indicates widespread adoption across student populations, with many learners viewing AI tools as essential components of their academic workflow. However, the integration of AI is uneven: some students use these tools to support their learning processes, while others rely on them as substitutes for original work. This divergence raises questions about fairness, assessment validity, and the broader meaning of academic credentials.

The fundamental issue is that higher education has long relied on submitted products as evidence of internal learning. GenAI disrupts this assumption because high-quality outputs can now be generated

Copyright © 2025 The Author(s) : This is an open access article under the CC BY license (http://creativecommons.org/licenses/by-nc/4.0/)

without corresponding cognitive development. As recent scholarship has noted, educational risks arise where data quality, assignment structure, and incentive systems intersect.

TABLE I
STUDENT ADOPTION AND USAGE STATISTICS (2024-2025)
Data aggregated from Digital Education Council and HEPI surveys [4],[5]

Metric	Statistic	Description
Overall Usage	86%	Percentage of students reporting regular use of AI in studies.
Daily Frequency	24%	Students using AI tools on a daily basis.
Primary Tool	66%	Usage dominance of ChatGPT compared to other tools.
Writing Aid	25%	Usage of Grammarly and similar editorial augments.
Perceived Necessity	67%	Students who believe AI is "essential" in today's world.

The presence of GenAI in higher education manifests in a complex plurality of forms. Students may utilize these systems to bypass the productive struggle of essay writing, or they may employ them as sophisticated dialectical partners. Conversely, instructors and administrators may employ these tools to generate syllabi or personalize feedback. These behaviors raise profound ethical concerns, particularly as universities serve as gatekeepers for professional credentialing. The stakes of unregulated AI adoption extend beyond individual instances of plagiarism; they pose the risk of a systemic devaluation of the university degree.

At the core of this issue lies the role of *process* versus *product*. Traditional higher education has relied on the tangible output—the essay, the code block—as a proxy for internal learning. GenAI shatters this proxy relationship. An LLM can produce a high-quality output without the user possessing the corresponding internal knowledge state. As noted in recent dataset-centric studies, errors and harms enter systems at many points; in education, the disruption enters at the intersection of assignment design and student incentive structures.

Furthermore, the data-centric perspective highlighted in bias research is directly applicable to the educational context. The models entering the classroom are not neutral; they are culturally and historically situated artifacts. They carry the "representational biases" of their training data into the diverse environment of the global university. A dataset may contain imbalanced demographic representation, which, when deployed as an educational tutor, can marginalize non-dominant epistemologies.

In this paper, we adopt a pedagogy-centric and sociotechnical perspective to the problem of GenAI in higher education. We argue that any effort to address GenAI must begin with a reassessment of *why* and *how* we assess students.

2. RELATED WORK

The study of technology in higher education has evolved significantly, marked by cycles of hype and assimilation. To understand the specific impact of GenAI, it is necessary to situate it within the broader history of Educational Technology, academic integrity, and algorithmic bias.

A. The Evolution of Automated Instruction

Prior to the generative era, research focused on Intelligent Tutoring Systems (ITS) which sought to scaffold student learning through deterministic pathways. GenAI represents a rupture in this lineage; unlike ITS, LLMs are probabilistic rather than deterministic. They construct answers statistically based on training data patterns. This distinction shifts the risk profile from "rigid feedback" to "hallucinated feedback."

B. Algorithmic Bias and Epistemic Injustice

Scholars in AI ethics have proposed frameworks for understanding harm that are highly relevant to education. As noted in dataset-centric studies, models inherit and reproduce social biases. In an educational context, this means GenAI tutors may provide robust feedback on Western canonical texts while failing on Global South literature. This phenomenon, known as "representational bias" , translates into "epistemic bias" in the classroom. If universities rely on these tools for personalized learning, they risk reinforcing historical inequalities.

C. The Assessment-Validity Gap

Construct validity is compromised when the assessment method allows for "prompting ability" rather than "writing ability." Recent studies on GenAI integration have proposed frameworks for "Generative AI-Driven Assessment" that involve three key stakeholders: instructors, students, and control authorities. The literature suggests that the "arms race" of detection is largely futile due to the rapid evolution of models, necessitating a shift toward "assessment for learning".

3. METHODOLOGY

This study adopts a qualitative, sociotechnical methodology to analyze how GenAI affects instructional practices and institutional governance.

A. *Theoretical Framework and Data Sources*

We ground our analysis in dataset-centric approaches that link model behavior to the underlying data used for training. Technical reports from major model developers and educational policy documents inform the conceptual lens used here.

B. *Taxonomy Construction*

A taxonomy of student–AI interaction was developed by mapping common model capabilities to typical student workflows. Categories were defined by determining whether cognitive responsibility rested primarily with the student or with the AI tool.

C. *Protocol Development*

Our proposed evaluation protocol adapts fairness-auditing strategies from machine learning governance and applies them to educational assessment, offering institutions a structured starting point for policy development.

4. TAXONOMY OF GENAI-STUDENT INTERACTION

Understanding how GenAI manifests in higher education requires a taxonomy of the interactions between learner and model. We identify four major categories defined by the locus of cognition.

TABLE II
TAXONOMY OF AI-STUDENT INTERACTION TYPES & TOOL CAPABILITIES
Based on interaction analysis and tool capabilities [8], [9]

Interaction Type	Definition	Primary Tools	Educational Risk
Unauthorized Proxy	Total replacement of student cognition; AI generates full output.	ChatGPT-4, Jasper	Complete loss of learning; academic fraud.
Socratic Interlocutor	Iterative dialogue to clarify concepts or brainstorm.	Claude 3, Bing Chat	Epistemic dependence; internalization of hallucinations.
Editorial Augment	Refining syntax/grammar of human-generated text.	Grammarly, DeepL	Homogenization of voice; masking L2 distinctiveness.
Semantic Obfuscator	Adversarial modification to evade detection.	Quillbot, SpinBot	Institutional distrust; surveillance arms race.

A. The Unauthorized Proxy (Total Substitution)

This occurs when the student utilizes the LLM to generate the entirety of a deliverable. The student functions merely as a "relay," and there is near-zero cognitive engagement. The output reflects the "semantic leakage" of the training data—reproducing the most common arguments found in the corpus. This represents a total failure of the educational feedback loop.

B. The Socratic Interlocutor (Pedagogical Scaffolding)

In this mode, the student uses GenAI as a tutor or dialectical partner. The interaction is conversational and iterative. The locus of control remains with the student, who must synthesize the AI's responses. However, if the model exhibits bias , the student may uncritically internalize these errors.

C. The Editorial Augment (Linguistic Polishing)

This involves utilizing the LLM to refine syntax and grammar while retaining the core human argumentation. While this can level the playing field for L2 learners, it raises concerns about "representational bias" where only certain forms of expression are deemed valid by the model, leading to a homogenization of student voice.

D. The Semantic Obfuscator (Detection Evasion)

This is an adversarial interaction where the student uses tools to paraphrase text specifically to bypass detection algorithms. This transforms the educational environment into a surveillance state.

5. METHODS FOR DETECTION AND ASSESSMENT VERIFICATION

Educational institutions have aggressively deployed a wide array of technical strategies and software solutions to identify the presence of Generative AI in student submissions. However, scrutiny reveals that each detection method currently available suffers from significant limitations, often resulting in unreliable accuracy, alarming false-positive rates, and easy evasion by students.

A. Statistical Probability Analysis

These tools measure "perplexity" and "burstiness." If a text has consistently low perplexity, it is flagged as likely AI-generated. However, research indicates high false-positive rates, particularly against non-native English speakers, mirroring the "Annotation Bias" where certain groups are unfairly penalized.

B. Process Tracing and Version History

This method analyzes the metadata of document creation. A document that appears instantly suggests "copy-paste" behavior. This forensic approach privileges the *labor* of writing over the *result*, aligning with the need to assess the learning process rather than just the artifact.

C. Oral Defense

For high-stakes assessments, the text is treated as an entry ticket to an oral examination. If a student cannot define the terms used in their own title, authorship is compromised. This prioritizes the student's ability to articulate knowledge in real-time.

6. DATASET-DRIVEN SOURCES OF PEDAGOGICAL FRICTION

The disruption caused by GenAI is a reflection of the data used to train the models. We examine sources of friction through the lens of the dataset-centric approach.

A. The Bias of the Training Corpus

LLMs are trained on predominantly Western, English-centric corpora. When students use GenAI to research non-Western history, the model is likely to hallucinate or revert to stereotypes. The "dataset" of the AI often does not match the "dataset" of a diversified curriculum.

B. The "Average" as the Enemy of Excellence

LLMs function by predicting the most probable next token. In education, "probable" often equates to "consensus." If a student uses AI to generate ideas, they start from the statistical mean of human discourse, creating a "Semantic Leakage" of mediocrity.

C. Filtering and Sanitization Risks

Datasets are filtered to remove toxicity, which can inadvertently remove valid language varieties or controversial topics. An AI with strong safety filters may refuse to summarize a text containing hate speech that is being analyzed for its rhetorical strategies, acting as a censor.

7. MITIGATION STRATEGIES: PEDAGOGICAL INTERVENTIONS

Effective mitigation strategies must be fundamentally pedagogical in nature, requiring a comprehensive redesign of the learning environment. The ultimate objective is to restructure assessments and activities in such a way that the utilization of an 'Unauthorized Proxy' is rendered functionally ineffective, as the tasks would demand authentic, human-centric cognitive engagement that cannot be offloaded

A. Authentic and Contextual Assessment Design

Assignments should be anchored in contexts inaccessible to the LLM's training data, such as specific in-class discussions or local events. This forces the student to engage with the immediate learning environment.

B. The Flipped Assessment

This strategy moves low-level information gathering (potentially AI-assisted) outside the class, while high-level synthesis and writing occur inside the class without technology. This validates that the student possesses the internal cognitive structures necessary to write.

C. Critical AI Integration (Adversarial Prompting)

Assignments can require students to generate an essay using GenAI and then write a secondary reflection critiquing the output. The student must identify hallucinations and bias , moving them up Bloom's Taxonomy from "Create" to "Evaluate".

D. Comprehensive Assessment Frameworks

Recent frameworks, such as the Comprehensive AI Assessment Framework (CAIAF), suggest distinct levels of AI integration based on educational level and ethical guidelines. This ensures that AI adoption is structured rather than ad-hoc.

8. EVALUATION PROTOCOL

Effective mitigation depends on rigorous institutional evaluation. We propose a protocol structured around five key steps, adapted from fairness auditing frameworks.

1. **Scope Definition:** Define what constitutes "AI Misconduct" versus "AI Literacy."
2. **Infrastructure Audit:** Audit the "Allocational Fairness" of access. Do all students have access to advanced models?
3. **Literacy Probing:** Measure whether faculty and students can distinguish between AI and human writing and understand ethical implications.
4. **Attribution Analysis:** Analyze longitudinal data to determine if GenAI is affecting learning outcomes or if detection tools are exhibiting bias against specific demographics.
5. **Iterative Review:** Establish a governance committee to continuously refine policy as models evolve.

9. DISCUSSION

The dataset-centric perspective highlights several implications. First, "Academic Integrity" must be reframed not as a moral failing of the student, but as a structural failing of assessment. Second, "Data Literacy" becomes a core competency; students must understand "Sampling Bias" to avoid becoming passive consumers. Third, the scale of GenAI introduces an "Epistemic Crisis" regarding the verification of truth in an era of synthetic media. Finally, we emphasize "Human-in-the-Loop" education; AI should support the teacher, not replace the relational process of mentorship.

10. LIMITATIONS

A critical limitation of this work is that our framework assumes institutions have the organizational agility to rapidly update curricula, a capability which is often not the case due to rigid administrative structures that lag behind the speed of AI evolution. Additionally, the persistent digital divide remains a significant barrier to achieving 'Allocational Fairness,' ensuring that access to these powerful tools is not equitably distributed. Finally, we observe that technical detection metrics are currently failing to provide reliable results, suggesting a dangerous move toward invasive surveillance and proctoring technologies that introduces its own set of profound ethical harms regarding student privacy

11. CONCLUSION

The pervasive issue of bias inherent in large language models, coupled with the profound disruption caused by Generative AI within higher education, should not be viewed as accidental anomalies; rather, they are systemic outcomes resulting directly from specific data curation and architectural design choices. To address this, this paper has presented a comprehensive framework specifically engineered for detecting and mitigating these inherent risks before they manifest in the classroom. By rigorously classifying distinct interaction types between users and agents and proposing a structured,

multi-layered evaluation protocol, we provide educators and developers with critical tools for responsible AI development and deployment. Ultimately, we argue that true fairness and academic integrity must begin upstream—scrutinizing the datasets that shape the foundational models, as well as reimagining the assessments that shape student learning.

References

[1] OpenAI, "GPT-4 Technical Report," *arXiv preprint arXiv:2303.08774*, 2023.

[2] H. Touvron *et al.*, "Llama 2: Open Foundation and Fine-Tuned Chat Models," *arXiv preprint arXiv:2307.09288*, 2023.

[3] S. Bachhar and S. Choudhary, "Bias Detection and Mitigation in Large Language Models: A Dataset-Centric Study," Quantum University, Roorkee, India, 2024.

[4] Digital Education Council, "Global AI Student Survey 2024," *Campus Technology*, Aug. 2024.

[5] Higher Education Policy Institute (HEPI), "Student Generative AI Survey 2025," *HEPI Reports*, Feb. 2025.

[6] UNESCO, "Guidance for generative AI in education and research," UNESCO, Paris, 2023.

[7] E. M. Bender, T. Gebru, A. McMillan-Major, and S. Shmitchell, "On the Dangers of Stochastic Parrots: Can Language Models Be Too Big?" in *Proc. 2021 ACM Conf. on Fairness, Accountability, and Transparency*, 2021, pp. 610–623.

[8] Y. Bai *et al.*, "Constitutional AI: Harmlessness from AI Feedback," *arXiv preprint arXiv:2212.08073*, 2022.

[9] EDUCAUSE, "2025 EDUCAUSE Horizon Report | Teaching and Learning Edition," EDUCAUSE, Boulder, CO, 2025.

[10] J. Wei *et al.*, "Chain-of-Thought Prompting Elicits Reasoning in Large Language Models," *arXiv preprint arXiv:2201.11903*, 2022.

[11] "Thousands of UK university students caught cheating using AI," *The Guardian*, June 2025.

[12] Engageli, "AI in Education Statistics: Impact on Learning Outcomes," *Engageli Blog*, 2025.

[13] T. Yankova, "A Framework for Generative AI-Driven Assessment in Higher Education," *ResearchGate*, 2025.

[14] P. Liang *et al.*, "Holistic Evaluation of Language Models (HELM)," *arXiv preprint arXiv:2211.09110*, 2022.

[15] J. Jin *et al.*, "The Generative AI Literacy Assessment Test (GLAT)," *arXiv preprint arXiv:2504.19038*, 2025.

[16] A. Vaswani *et al.*, "Attention Is All You Need," in *Advances in Neural Information Processing Systems*, 2017, pp. 5998–6008.

[17] Anthropic, "The Claude 3 Model Family: Opus, Sonnet, Haiku," Anthropic Technical Report, 2024.

[18] Google DeepMind, "Gemini 1.5: Unlocking Multimodal Understanding Across Millions of Tokens of Context," *arXiv preprint arXiv:2403.05530*, 2024.

[19] T. Brown *et al.*, "Language Models are Few-Shot Learners," in *Advances in Neural Information Processing Systems*, vol. 33, 2020, pp. 1877–1901.

[20] L. Ouyang *et al.*, "Training language models to follow instructions with human feedback," in *Advances in Neural Information Processing Systems*, vol. 35, 2022, pp. 27730–27744.

[21] S. Bubeck *et al.*, "Sparks of Artificial General Intelligence: Early experiments with GPT-4," *arXiv preprint arXiv:2303.12712*, 2023.

[22] A. Q. Jiang *et al.*, "Mistral 7B," *arXiv preprint arXiv:2310.06825*, 2023.

[23] E. Hu *et al.*, "LoRA: Low-Rank Adaptation of Large Language Models," in *Int. Conf. on Learning Representations (ICLR)*, 2022.

[24] R. Rafailov *et al.*, "Direct Preference Optimization: Your Language Model is Secretly a Reward Model," in *Advances in Neural Information Processing Systems*, vol. 36, 2024.

[25] Stanford HAI, "Artificial Intelligence Index Report 2024," Stanford Institute for Human-Centered AI, Stanford University, 2024.

[26] World Economic Forum, "The Future of Jobs Report 2025," WEF, Geneva, Switzerland, May 2025.

[27] U.S. Department of Education, Office of Educational Technology, "Artificial Intelligence and the Future of Teaching and Learning: Insights and Recommendations," Washington, DC, 2023.

[28] European Commission, "Ethical guidelines on the use of artificial intelligence (AI) and data in teaching and learning for educators," Publications Office of the European Union, Luxembourg, 2022.

[29] C. K. Y. Chan, "A comprehensive AI literacy framework for scientific education," in *Proc. Int. Conf. on Education and Artificial Intelligence*, 2023.

[30] A. Bozkurt *et al.*, "Generative Artificial Intelligence (AI) Powered Trends and Issues in Education: A Review of the Literature," *Open Praxis*, vol. 15, no. 2, 2023.

[31] Y. K. Dwivedi *et al.*, "So what if ChatGPT wrote it? Multidisciplinary perspectives on opportunities, challenges and implications of generative conversational AI for research, practice and policy," *Int. J. Inf. Manage.*, vol. 71, 2023.

[32] E. Mollick, *Co-Intelligence: Living and Working with AI*. Portfolio/Penguin, 2024.

[33] M. Bearman and R. Ajjawi, "Learning to work with the black box: Pedagogy for a world with artificial intelligence," *British J. Educ. Technol.*, vol. 54, no. 5, 2023.

[34] W. Holmes and K. Porayska-Pomsta, *The Ethics of Artificial Intelligence in Education: Practices, Challenges, and Debates*. Routledge, 2023.

[35] N. Selwyn, *Should Robots Replace Teachers? AI and the Future of Education*. Polity Press, 2019.

[36] L. Floridi *et al.*, "AI4People—An Ethical Framework for a Good AI Society: Opportunities, Risks, Principles, and Recommendations," *Minds and Machines*, vol. 28, pp. 689–707, 2018.

[37] K. Crawford, *The Atlas of AI: Power, Politics, and the Planetary Costs of Artificial Intelligence*. Yale University Press, 2021.

[38] Microsoft, "The Future of Work: Reinventing Productivity with AI," Microsoft WorkLab, 2024.

[39] J. Manyika *et al.*, "Generative AI and the future of work in America," McKinsey Global Institute, 2023.

[40] P. Lewis *et al.*, "Retrieval-Augmented Generation for Knowledge-Intensive NLP Tasks," in *Advances in Neural Information Processing Systems*, vol. 33, 2020.

[41] T. Kojima *et al.*, "Large Language Models are Zero-Shot Reasoners," in *Advances in Neural Information Processing Systems*, vol. 35, 2022.

[42] O. Zawacki-Richter *et al.*, "Systematic review of research on artificial intelligence applications in higher education – where are the educators?" *Int. J. Educ. Technol. High. Educ.*, vol. 16, no. 39, 2019.

[43] G. Siemens *et al.*, "Human and Artificial Cognition," in *Proc. 13th Int. Learning Analytics and Knowledge Conf. (LAK23)*, 2023.

[44] B. Williamson, *Big Data in Education: The Digital Future of Learning, Policy and Practice*. SAGE Publications, 2017.

[45] UNESCO, "AI Competency Framework for Teachers," UNESCO, Paris, Sept. 2024.