Revue Internationale du Law & Education

Vol. 01 No.02,July 2026
DOI: 10.67590/ridle.v1i2.17

Academic Integrity in the Age of Generative Artificial Intelligence: Legal and Ethical Perspectives

Asha Maria
Research Scholar, Department of law, Vikrant University, Gwalior, Madhya Pradesh,474 006,
Email id: ashamarialawaim@gmail.com, ORCID id: 0009-0003-1070-7683

Received: 12/06/2026
Revised: 13/07/2026
Accepted: 20/07/2026
Published: 28/07/2026

Abstract

The generative artificial intelligence (AI) plays a significant role in fields of higher education and research. It has disturbed the established concepts of authorship, novelty and academic integrity. This paper examines the legal and ethical perspectives of use of generative AI in academic backgrounds through a doctrinal and thematic analysis of recent scholarship, institutional guidance and evolving governing instruments. It reflects that there is inadequacy of binding statutory regulations, governance currently focus on institutional policy, publisher and funder mandates and developing rules of disclosure rather than legislation. Ethically, there has a shift of the debate from individual misconduct towards general accounts that involve assessment design, institutional culture and equitable access to AI tools. The analysis identifies continuing tensions between detection-based enforcement and due process, and between innovation and fairness. The paper concludes that transparent disclosure frameworks, redesigned assessment practices and layered governance models offer a more secure path forward than total prohibition or continued reliance on untrustworthy technologies of detection.

Keywords: Academic Integrity, Generative Artificial Intelligence, Higher Education, Research Ethics, Disclosure Policy, Plagiarism

1. Introduction

The large-scale generative AI systems are now available publicly and it has facilitated students and researchers to generate eloquent, structurally competent text, code and analysis with minimum of effort. This resulted to the interruption of the foundation on which academic integrity systems were created. Traditional definitions of plagiarism and collusion presume that words or ideas of a human source are appropriated without acknowledgment. A generative AI system is not a person in that sense and its output is often original in form even where it reproduces patterns learned from a training corpus. Therefore, institutions have to decide whether undisclosed use of AI constitutes plagiarism as a new species of misconduct or just a tool that should be regulated rather than prohibited (Eke, 2023).

The problem is not confined to student writing. Generative AI has also entered the field of research itself. It raises serious questions about authorship credit, responsibility for factual errors and the integrity of peer review (Bittle & El-Gayar, 2025). This may be because of the fact that these tools are inexpensive, widely accessible and difficult to detect reliably. The scale of probable misuse is exceeding the earlier forms of academic dishonesty. The automated AI-detection softwares which are used to police it, carry significant error rates and it raises concerns when it is used as the basis for disciplinary action.

This paper raises two related questions. First, about legal frameworks that are currently govern the use of generative AI in academic contexts and adequacy of it to address the problem. Second, about principles of ethics which should guide the dealing of generative AI in teaching, learning and research and the connection of these principles to the existing or emerging legal frameworks. The paper progresses by reviewing recent literature and institutional guidance by describing the methodology used to select and synthesize that material, presenting the resulting findings, analysing the underlying tensions, identifying the principal challenges facing institutions and regulators and proposing a path forward grounded in transparency, proportionality and assessment redesign rather than prohibition.

Literature Review

Intellectual attention to generative AI and academic integrity has witnessed rapid growth since 2023. The literature can be grouped into three broad parts: empirical and conceptual studies of student and faculty behaviour, institutional and regulatory responses and normative or ethical frameworks.

The first part is how generative AI is really being used and perceived in academic settings. Bittle and El-Gayar (2025) made a systematic review of synthesis of studies published between 2021 and 2024. It was concluded that generative AI at once improves educational engagement and efficiency. At the same time, it can be seen that it creates serious risks of academic dishonesty with the literature still limited in its geographic and disciplinary diversity. Eke (2023) considered ChatGPT explicitly as a disturbing threat to conventional integrity safeguards. He argues that existing plagiarism-detection infrastructure was not designed to address AI-generated text. Evangelista (2025) took a more practical approach, considering that detection, exam design and assessment strategy needed to be reorganised to preserve the integrity in an AI-saturated environment.

The second part is in relation to institutional and regulatory responses. Cabrera Vélez et al. (2026) considered the legislative landscape across South American jurisdictions and found no country-specific statute addressing the assurance of academic integrity in relation to generative AI. The drawing of the boundary between educational support and academic dishonesty has been left to be drawn by individual institutions rather than by law. Where formal regulatory instrument exists, they tend to originate from research-funding or standard-setting bodies rather than from education ministries or legislatures. Brazil's Conselho Nacional de Desenvolvimento Científico e Tecnológico, for example, issued Portaria CNPq No. 2.664 in March 2026, establishing a Policy of Integrity in Scientific Activity that does not prohibit generative AI use but requires researchers to declare the specific tool and purpose of any AI assistance. It forbids presenting AI-generated content as if it were unaided human work and holds authors fully responsible for the accuracy and originality of the final output regardless of AI involvement (Conselho Nacional de Desenvolvimento Científico e Tecnológico, 2026). Universities have moved ahead in a similar direction through internal guidance typically requiring disclosure of AI assistance and alerting researchers about bias, misinformation, intellectual-property compliance and overreliance (as cited in Xexéo, 2026).

The third part is normative and ethical. García-López and Trujillo-Liñán (2025) conducted a systematic review of the ethical and regulatory challenges raised by generative AI in education. It was concluded that existing integrity frameworks inadequately address algorithmic bias, data protection and the unequal distribution of access to AI tools across institutions and student populations. Gallent Torres et al. (2023) likewise centred their analysis on the ethical dimension. It was argued that academic integrity in the generative AI era cannot be separated from broader questions about the ethics of the technology itself. Maleki (2026) proposed a Higher Education Integrity Ecosystem Model in which integrity is considered as an emergent property of four interacting layers, namely governance and policy, assessment design, institutional and cultural norms and individual moral agency, rather than as a matter of individual compliance alone. This model specifically reframes generative AI as a force that exposes the prior weaknesses of institutions in design assessment and policy but not as a reason of increased individual misconduct.

Another significant field of academic attention is of attribution and disclosure design. Xexéo (2026) remarked that present disclosure practice normally records whether AI was used only, not how, where, or with what degree of human supervision. An attribution model was proposed to cover the categories of generation, evaluation, control and traceability, to make hybrid human-AI authorship legible. It can be better rather than forcing it into a binary human-or-AI classification. This line of work bridges the legal necessity of disclosure with the ethical requirement of honest authorship attribution and it is taken up for further analysis.

The literature signifies an evolution that empirical work approves the use of generative AI as widespread. In addition, it is difficult to regulate through detection alone. Regulatory frameworks are largely non-statutory and not unified. The ethical scholarship is moving away from individual-blame models towards general and equity-oriented frameworks. Therefore, this paper treats the legal and ethical magnitudes of the problem as interdependent rather than separate tracks of inquiry.

Methodology

This paper adopts a doctrinal and thematic literature review methodology suitable to a legal-ethical inquiry, rather than an empirical study design. Sources were identified through targeted searches of academic databases and preprint repositories for material published between 2023 and 2026 addressing generative AI, academic integrity, research ethics and higher education regulation. Preference was given to peer-reviewed systematic reviews, government policy instruments and institutional guidance documents, supplemented by recent preprints addressing attribution and disclosure design.

An inductive coding approach is used for analysis of sources. The material was sorted into three categories: empirical and conceptual accounts of AI use and misuse, legal and institutional regulatory responses and normative or ethical frameworks. Within each category, sources were compared for points of convergence and divergence, with particular attention to whether they treated academic integrity as a matter of individual conduct, institutional design or systemic governance. The doctrinal element of the analysis focuses on identifying the current legal instruments in force, their scope and enforceability. The ethical element draws on principles articulated in the reviewed normative literature, including honesty, transparency, proportionality and equity.

This approach is not free from limitations. It is not a systematic review, as it does not apply predefined inclusion and exclusion criteria across an exhaustive database search. Further, it draws on a rapidly evolving literature which will be obsolete soon. But it is suitable to the exploratory, cross-disciplinary purpose of this paper. It helps to signify the relationship between legal regulation and ethical analysis on a problem for which there is non-existence of widespread empirical data and settled legal doctrine.

Results and Discussion

Three principal findings are the result of this review. Firstly, the legal regulation of generative AI in academic contexts is scarce, fragmented and directed at research-integrity policy, not education law. No jurisdiction reviewed in this analysis has enacted a statute directly defining the boundary between permissible AI assistance and academic misconduct. The policy of the CNPq in Brazil, while binding on CNPq-funded research activity, is instructive precisely because it is the exception, not the rule (Cabrera Vélez et al., 2026; Conselho Nacional de Desenvolvimento Científico e Tecnológico, 2026). In the absence of such, the operational norms for most students and researchers are institutional codes of conduct, which vary greatly in specificity.

Second, the disclosure model is emerging as the preferred regulatory mechanism where formal rules exist. Instead of an outright ban on the use of generative AI, the CNPq policy and institutional guidance reviewed require researchers and students to declare the use of the tool, the purpose and extent of reliance on it and hold the human author fully responsible for the accuracy, originality and propriety of the final work product (Conselho Nacional de Desenvolvimento Científico e Tecnológico, 2026). This is important because it changes the legal question from whether AI was used, to whether its use was disclosed and the resultant content was verified. Xexéo's (2026) faceted attribution model directly tackles a weakness of this approach, namely that a plain yes-or-no disclosure statement cannot capture gradations of human oversight, and suggests a structured vocabulary for describing how AI contributed to a text at the level of the document, chapter, section or paragraph. Such models show that the legal and technical framework for meaningful disclosure is still playing catch-up.

Third, the ethical literature has shifted from an individual behavior framing to a systemic frame, and this has legal implications. Maleki's (2026) ecosystem model of academic integrity considers academic integrity as an emergent attribute of governance, assessment design, institutional culture and individual activity. This suggests that punishing individual students or researchers for AI related misconduct addresses only one of four interacting layers and may leave the root causes, such as assessment tasks that generative AI can complete without engaging the skills being tested, unaddressed (Evangelista, 2025; Maleki, 2026). García-López and Trujillo-Linan (2025) add an equity dimension to this systemic account, noting that reliance on AI-detection tools risks disproportionately burdening students who lack access to premium AI tools or institutional support, or whose writing style is flagged by detection software due to bias in its training data rather than actual misuse.

These findings are mutually reinforcing. The absence of statutory regulation places most of the burden on institutional policy and the systemic ethical critique explains why disclosure and assessment redesign, rather than detection and punishment, have become the preferred institutional response. The next section analyses the tensions this creates and their implications for practice.

A further relevant finding to be noted is the asymmetry between teaching-focused and research-focused governance. Institutional codes of conduct aimed at students tend to be phrased in prohibitive or permission-based terms, specifying what a student may or may not do with AI assistance on a given task. But the research-integrity instruments reviewed above, including the CNPq policy, are phrased in disclosure and accountability terms, specifying what must be declared and who remains responsible for the result. This difference in regulatory style reflects a difference in what each context is trying to protect: student assessment aims to verify a learner's own competency, while research integrity aims to protect the reliability of a public record that others will rely upon. The literature reviewed here has not yet produced a framework that reconciles these two governance logics.

Analysis

Three tensions recur across the reviewed material and analysis: the tension between detection-based enforcement and due process, the tension between disclosure requirements and the practical difficulty of verifying them, and the tension between individual responsibility and systemic causation.

The due process tension arises because AI-detection tools, that are used by institutions act as a first line of enforcement. That are probabilistic classifiers rather than forensic instruments. They estimate a likelihood that a text was AI-generated based on statistical patterns and misclassify text written by non-native English speakers, neurodivergent students and writers who use structured or formulaic prose, all of whom may be flagged despite having written the text themselves. Bittle and El-Gayar (2025) noted this as an unresolved risk within their systematic review, and García-López and Trujillo-Linan (2025) treat it as a fairness and equity concern in its own right. Because academic sanctions can carry consequences comparable to those of formal legal proceedings, including expulsion, degree revocation or loss of funding, the evidentiary weakness of detection tools raises the same concerns that arise in any adjudicative process relying on unreliable forensic evidence. Institutions risk sanctioning while at the same time under-detecting more sophisticated misuse.

The disclosure is legally and ethically coherent in principle but difficult to be in practice. A student or researcher who is required to declare AI use has an incentive and an institution that relies solely on self-report has limited means of verifying compliance. Xexéo's (2026) faceted attribution framework partially addresses this by making disclosure more granular and therefore more informative. But granularity alone does not solve the verification problem. It only makes a false or incomplete disclosure easier to identify as such after the fact, not to prevent in advance.

The responsibility tension is the most consequential for institutional design. If, as Maleki (2026) argues, integrity failures are of individual choice. Hence, punitive response to individual cases of undisclosed AI use treats a general problem as though it were solely a matter of personal ethics. This does not mean that individual responsibility disappears. The CNPq policy is explicit that authors remain fully responsible for their final work regardless of AI involvement (Conselho Nacional de Desenvolvimento Científico e Tecnológico, 2026). The allocation of responsibility is both legally and ethically sound, because a generative AI system cannot bear legal or moral responsibility for its output. The more secure policy is that individual responsibility for disclosure and accuracy should be combined with institutional responsibility for designing assessments and policies that do not depend on AI-inaccessibility to function.

A fourth, cross-cutting tension deserves mention: the mismatch in pace between the technology and the governance instruments meant to regulate it. Generative AI capabilities have changed substantially within the same period that the reviewed policies and frameworks were drafted. It means that a disclosure category, an assessment redesign or a detection tool regulated to one generation of AI systems can be a poor fit for the next. This is a practical inconvenience. It has legal significance, since institution rules of current limitations of a technology, such as its inability to perform a certain type of reasoning or citation, risk becoming under-inclusive or over-inclusive within a short time, exposing them to challenge on grounds of vagueness or inconsistent application. Any durable governance framework therefore needs to be evaluated not only on whether it addresses today's technology but on how it degrades as that technology changes.

Challenges and Suggestions

There are several practical challenges for this analysis. The first is the absence of legal certainty. Because, most of the governing rules sit at the institutional rather than statutory level. Students and researchers moving between institutions or jurisdictions face inconsistent standards and institutions themselves lack authoritative guidance on questions such as whether AI-assisted text can be copyrighted, who bears liability for AI-introduced factual errors in published research and what evidentiary standard should apply to a misconduct finding based on detection software. The first suggestion is for higher-education regulators and professional and disciplinary bodies to develop model disclosure and integrity standards that institutions can adopt or adapt, following the precedent set by research-funding bodies such as CNPq.

The second challenge is the unreliability of detection technology relative to the weight placed on it in disciplinary proceedings. A corresponding suggestion is that institutions should treat AI-detection output as, at most, a prompt for further inquiry rather than as dispositive evidence of misconduct. Additionally it should require corroborating evidence, such as inconsistency with a student's earlier work, drafting history or process documentation, before imposing serious sanctions. Institutions that offer version-history or process-tracking tools as part of coursework submission can generate this kind of corroborating evidence.

The third challenge is assessment design that remains vulnerable to AI substitution regardless of disclosure policy. Evangelista (2025) and Maleki (2026) both point toward the same suggestion from different directions. Institutions should shift a meaningful share of assessment towards formats that are difficult for generative AI to complete on a student's behalf without also demonstrating the underlying competency, such as oral defenses, in-class or supervised work, iterative process portfolios and tasks that require engagement with unpublished or highly localized material. This does not eliminate the need for integrity policy, but it reduces the stakes riding on detection and disclosure enforcement alone.

Equity is the fourth challenge. Moreover, if access to capable AI tools, training in their effective and ethical use and exposure to institutional guidance is unevenly distributed, then integrity enforcement risks becoming another axis of educational inequality. It will result in penalizing under-resourced students and institutions more heavily than well-resourced ones (García-López & Trujillo-Linan, 2025). Another suggestion is that institutions can incorporate AI literacy which is inclusive of

instruction on appropriate disclosure and citation of AI assistance into the curriculum itself. Hence students are prepared to comply with the disclosure norms and not merely punished for failure to anticipate them.

The fifth challenge is about research integrity and the reliability of the published scientific record. The implication from the CNPq model and from Xexéo's (2026) attribution framework is that disclosure requirements for research outputs need to be more granular than a single statement that AI was used somewhere in the process. Journals, funders and institutions should adopt structured disclosure formats that specify which parts of a manuscript involved AI assistance, of what kind and with what degree of human verification, so that reviewers and readers can weigh the reliability of specific claims rather than the manuscript as an undifferentiated whole.

Conclusion

Generative AI has not created academic dishonesty; it has changed its scale, its detectability and the conceptual categories through which institutions try to regulate it. The legal response to date has been limited and largely non-statutory, concentrated in research-funder policy and institutional codes of conduct rather than legislation. But the ethical response has matured from a narrow focus on individual misconduct toward systemic frameworks that implicate governance, assessment design and equity of access. The tensions this creates, between detection and due process, between disclosure and verifiability and between individual and institutional responsibility, cannot be resolved by prohibition. This is because generative AI tools are already too embedded in ordinary academic work to be excluded by policy alone, nor by unregulated permissiveness, since unmanaged use threatens both the validity of assessment and the reliability of the research record.

The more defensible position to emerge from this review is one of managed transparency. Clear disclosure obligations regulated to the stakes involved, corroborated rather than purely algorithmic enforcement, assessment redesigned to remain meaningful in an AI-saturated environment and attention to the equity implications of every element of that framework. Neither the legal nor the ethical literature has fully worked out what such a framework should look like in operational detail, which is itself an indication of how early institutions and regulators remain in responding to this technology.

Path Forward

Three developments would most improve the current state of governance. First, model disclosure standards, comparable to the CNPq policy but developed collaboratively across institutions, disciplines and jurisdictions, would reduce the current inconsistency between institutions and give students and researchers a stable set of expectations regardless of where they study or publish. Second, investment in structured attribution frameworks of the kind proposed by Xexéo (2026), rather than binary disclosure statements, would make compliance more meaningful and more verifiable and would give journals, funders, and disciplinary bodies a shared vocabulary for describing hybrid human-AI authorship. Third, sustained empirical research is needed on the actual reliability of AI-detection tools across student populations, since the fairness of any enforcement regime that continues to rely on them depends on evidence that is still comparatively thin (Bittle & El-Gayar, 2025).

More broadly, the field would benefit from treating academic integrity and generative AI governance as a shared project between legal scholars, ethicists, and educational-assessment specialists rather than as a matter for any one discipline alone. The systemic models proposed by Maleki (2026) and the equity concerns raised by García-López and Trujillo-Linan (2025) both suggest that durable solutions will come from redesigning the systems in which students and researchers work, not merely from refining the rules that punish them when those systems fail.

References

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

2. Cabrera Vélez, J. P., Llanos García, R. V., & Bonilla Fierro, L. F. (2026). Generative AI and the assurance of academic integrity in the context of South American higher education. Frontiers in Education, 11, Article 1796556. https://doi.org/10.3389/feduc.2026.1796556

3. Conselho Nacional de Desenvolvimento Científico e Tecnológico. (2026). Portaria CNPq No. 2.664, de 6 de março de 2026: Política de Integridade na Atividade Científica.

4. Eke, D. O. (2023). ChatGPT and the rise of generative AI: Threat to academic integrity? Journal of Responsible Technology, 13, 100060. https://doi.org/10.1016/j.jrt.2023.100060

5. Evangelista, E. D. L. (2025). Ensuring academic integrity in the age of ChatGPT: Rethinking exam design, assessment strategies, and ethical AI policies in higher education. Contemporary Educational Technology, 17(1), Article ep559.

6. Gallent Torres, C., Zapata-González, A., & Ortego-Hernando, J. L. (2023). The impact of Generative Artificial Intelligence in higher education: A focus on ethics and academic integrity. RELIEVE, 29(2), 1–19.

7. García-López, I. M., & Trujillo-Linan, L. (2025). Ethical and regulatory challenges of Generative AI in education: A systematic review. Frontiers in Education, 10, Article 1681252. https://doi.org/10.3389/feduc.2025.1681252

8. Maleki, A. (2026). Rethinking ethical academic integrity ecosystems in higher education in the age of artificial intelligence. Discover Artificial Intelligence, 6, Article 145. https://doi.org/10.1007/s44163-026-00887-z

9. Xexéo, G. (2026). A faceted proposal for transparent attribution of AI-assisted text production. arXiv. https://arxiv.org/abs/2604.25346