Evaluating the Effectiveness of Institutional AI Ethics Policies in Promoting Academic Integrity: A Conceptual Framework

Yap Pek Jiuan[1], Haslina Abdullah[2], Nooraida Ayub[3]

Faculty of Business, UNITAR University College Kuala Lumpur, Malaysia

DOI: https://dx.doi.org/10.47772/IJRISS.2026.10100383

Received: 23 January 2026; Accepted: 29 January 2026; Published: 07 February 2026

ABSTRACT

The rapid adoption of generative artificial intelligence (AI) in higher education has transformed teaching, learning, and assessment practices, while intensifying concerns about academic integrity. Although many universities have introduced institutional AI ethics policies to promote responsible AI use, there remains limited evidence regarding the effectiveness of these policies in shaping students' academic integrity behaviours. Existing studies predominantly focus on students' ethical awareness and perceptions within single institutions and cross-sectional contexts, offering insufficient insight into the long-term behavioural impact of institutional governance mechanisms.

This paper develops a conceptual framework for evaluating the effectiveness of institutional AI ethics policies in promoting academic integrity in higher education. Drawing on the Theory of Planned Behaviour and Institutional Governance Theory, AI ethics policies are conceptualized as a multidimensional construct comprising policy clarity, AI ethics training, and enforcement mechanisms. Methodologically, the paper adopts a conceptual and analytical approach, integrating established theoretical perspectives with illustrative examples from Malaysian public and private universities, situated within national quality assurance expectations articulated by the Malaysian Qualifications Agency (MQA, 2023) and UNESCO (2023).

The proposed framework identifies how policy clarity and training enhance students' ethical awareness and normative beliefs, while enforcement mechanisms strengthen perceived behavioural control and accountability. Collectively, these dimensions mediate the relationship between institutional AI ethics policies and academic integrity outcomes, including ethical AI use and reduced AI-related misconduct. Illustrative cases demonstrate that institutions implementing coordinated and consistently enforced policy components are better positioned to influence student behaviour meaningfully. By examining this gap, the framework aims to support higher education institutions in enhancing ethical AI use.

Keywords: academic integrity, AI ethics policy, higher education governance, generative artificial intelligence, quality assurance

INTRODUCTION

Artificial intelligence (AI) is changing the way colleges and universities teach, learn, and grade (Evangelista, 2025; UNESCO, 2023). While AI brings advantages to the classroom, it also raises questions about academic honesty, ethics, and whether students complete their own work (Marín et al., 2025; Bittle & El-Gayar, 2025). To address these concerns, many institutions have implemented AI ethics policies to promote responsible AI use (Li et al., 2025; UNESCO, 2023).

Most research on AI and academic integrity focuses on students' perceptions and ethical awareness, including their attitudes toward ethical behaviour (bin Mohd Khidir et al., 2025; Bittle & El-Gayar, 2025). Typically, these studies are confined to single universities and rely on cross-sectional data (Bittle & El-Gayar, 2025). While these studies provide insight into student perspectives, they do not fully reveal whether institutional AI ethics policies influence student behaviour over time (Bittle & ElGayar, 2025). Furthermore, there is limited empirical evidence regarding which components of these policies—policy clarity, ethics training, or enforcement mechanisms—effectively guide responsible AI use across different universities (Marín et al., 2025; Bittle & El-Gayar, 2025).

Currently, longitudinal, multi-institutional research examining the effectiveness of AI ethics policies in higher education is lacking (Bittle & El-Gayar, 2025). In the absence of empirical evidence, institutions may implement policies that appear comprehensive on paper but do not produce meaningful improvements in student behaviour (UNESCO, 2023). Rigorous research is therefore required to identify policy components and practices that genuinely promote ethical AI use and uphold academic integrity (Bittle & El-Gayar, 2025).

Conceptual Framework Explanation

This paper lays out a framework to look at how university AI ethics policies shape the way students approach academic integrity. It is built on the Theory of Planned Behaviour (TPB) and Institutional Theory, which explain how official rules and policies influence behaviour, particularly in large organizations such as universities. In Malaysia, the Malaysian Qualifications Agency (MQA) takes the lead in ensuring that universities and colleges adhere to high academic standards (Malaysian Qualifications Agency [MQA], 2023). Although MQA has not introduced a dedicated AI policy, Advisory Note No. 2/2023 addresses the integration of generative AI in higher education and emphasizes responsible and ethical AI use aligned with existing quality assurance standards (MQA, 2023; UNESCO, 2023).

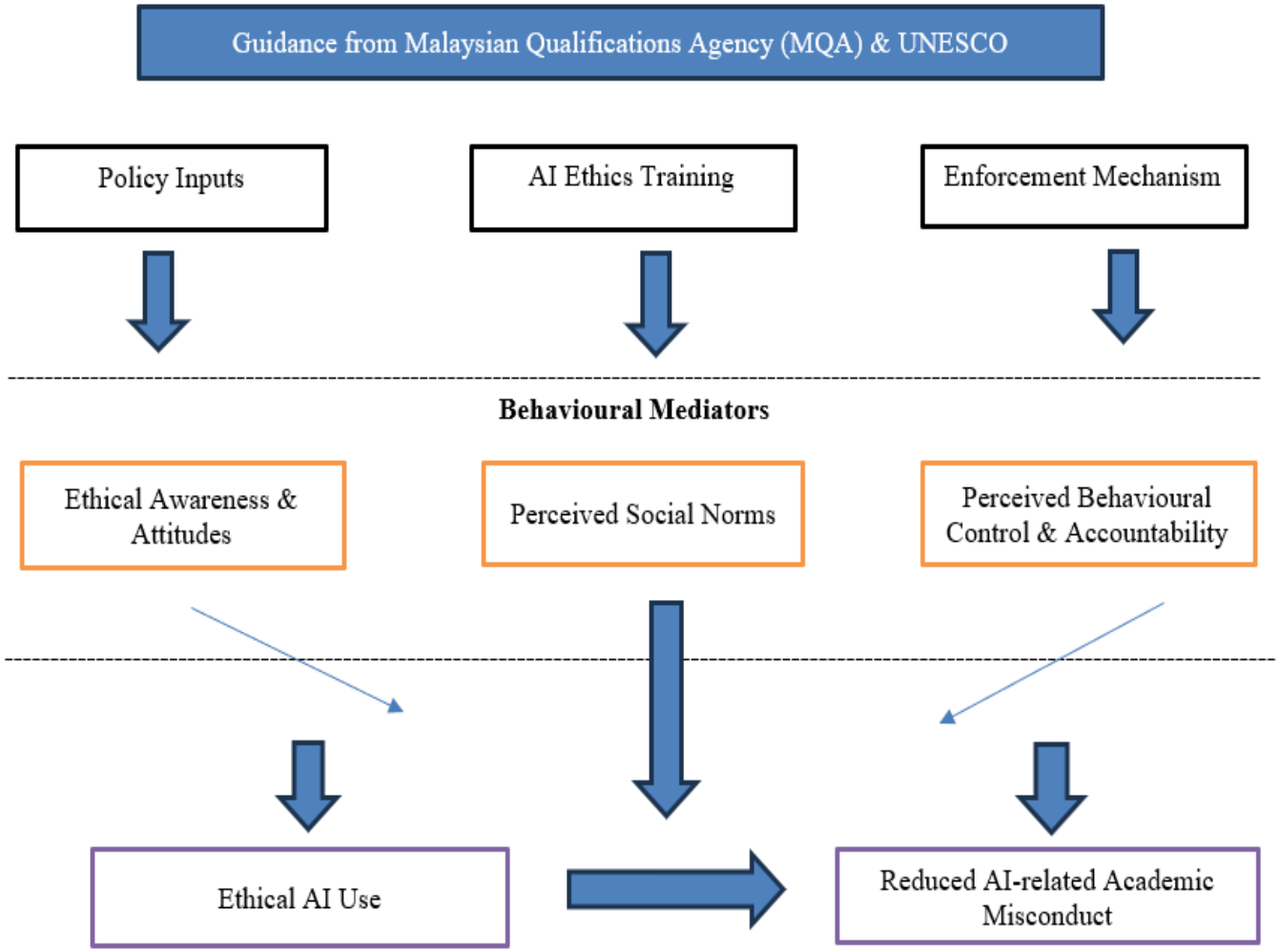
Malaysian universities are beginning to implement formal AI ethics policies, largely guided by MQA recommendations and national developments in AI governance (Malaysian Qualifications Agency [MQA], 2023). Effective policies are not merely documented but are clear, supported by training, and consistently enforced (Li et al., 2025; Bittle & El-Gayar, 2025). When embedded within teaching, research, and assessment practices, these policies are more likely to achieve their intended outcomes (Li et al., 2025).

Universiti Teknologi Malaysia (UTM) has introduced clear guidelines on generative AI use in teaching and assessment, specifying permitted and prohibited uses and requiring disclosure of AI assistance (MQA, 2023). UTM also conducts workshops and seminars to support students' understanding of responsible AI use. For research and publication, institutions such as Universiti Teknologi MARA (UiTM) and Universiti Malaya (UM) have implemented journal policies requiring disclosure of AI usage and limiting certain AI-assisted practices (Marín et al., 2025). Universiti Malaysia Pahang (UMP) has introduced editorial restrictions on AI-assisted manuscript preparation, while private institutions such as UNITAR International University have established internal policies defining acceptable AI use in assessments (MQA, 2023).

These institutional practices illustrate how Malaysian universities are operationalizing AI ethics through a combination of policy clarity, training, and enforcement (MQA, 2023; Li et al., 2025; Bittle & El-Gayar, 2025). The framework proposes that these components shape students' ethical awareness, perceived social norms, and perceived behavioural control, mediating the relationship between institutional AI ethics policies and academic integrity outcomes, including ethical AI use and reduced AI-related misconduct (Bittle & El-Gayar, 2025; Marín et al., 2025; UNESCO, 2023). By integrating institutional and behavioural perspectives, the framework provides a foundation for future quantitative, longitudinal, and multi-institutional research and shifts the focus from ethical perceptions alone to evaluating policy effectiveness in promoting ethical AI use in higher education (Li et al., 2025; Bittle & El-Gayar, 2025).

Conceptual Framework Figure

Figure 1. Conceptual framework for evaluating the effectiveness of institutional AI ethics policies in promoting academic integrity.



Conceptual Framework for Evaluating the Effectiveness of Institutional AI Ethics Policies in Promoting Academic Integrity

Behavioural Mediators

Conceptual Framework for Evaluating the Effectiveness of Institutional AI Ethics Policies in Promoting Academic Integrity

This proposed conceptual framework in Figure 1 illustrates how institutional AI ethics policy dimensions—policy clarity, AI ethics training, and enforcement mechanisms—influence students' academic integrity outcomes through key behavioural mediators derived from the Theory of Planned Behaviour. The model is situated within a broader institutional governance and quality assurance context and is intended to guide evaluation and future empirical inquiry rather than represent a tested causal model.

Explanatory paragraph

Figure 1 presents the proposed conceptual framework underpinning this study.

The framework conceptualizes institutional AI ethics policies as consisting of clear policies, training, and enforcement. It is based on the Theory of Planned Behaviour, which explains how AI ethics policies influence student behaviour. According to this theory, such policies only influence student behaviour indirectly through three key psychological factors that determine whether students behave ethically when utilizing Artificial Intelligence: students' ethical awareness and attitudes towards using AI, their perceptions of social norms regarding the ethical AI usage, their confidence in their ability to utilize AI responsibly within academic boundaries. It is imperative that researchers to examine whether clear and well-implemented AI ethics policies support responsible AI use and reduce AI-related academic misconduct, such as plagiarism or misuse of generative tools. This framework serves as a starting point for further studies investigating the effectiveness of institutional AI ethics policies across different types of higher education institutions.

Institutional Approaches to Regulating AI Use in Higher Education

Higher education institutions adopt differing approaches to regulating the use of artificial intelligence, reflecting variations in governance structures, institutional culture, and educational priorities (Li et al., 2025; UNESCO, 2023). In Malaysia, public universities, operating under stronger state oversight and centralized governance, often favour strict enforcement approaches. These institutions emphasise formal rules, monitoring mechanisms, and clearly defined sanctions for policy violations (Malaysian Qualifications Agency [MQA], 2023; Evangelista, 2025). In contrast, private universities, which operate with greater autonomy and market-driven flexibility, may place more responsibility for regulating AI use in the hands of academic staff, allowing lecturers and faculties to interpret institutional guidelines in discipline-specific and pedagogically appropriate ways (Li et al., 2025; Bittle & El-Gayar, 2025).

These governance strategies have distinct implications for motivating ethical behaviour. Enforcement centred approaches strengthen perceived behavioural control by highlighting consequences for noncompliance and reinforcing accountability structures (Bittle & El-Gayar, 2025; Evangelista, 2025).

Ethical AI use is encouraged primarily through deterrence and compliance, particularly in high-stakes assessment or research contexts. In academia-led models, students and staff are encouraged to internalise ethical principles through professional judgement and contextual decision-making, shaping ethical awareness and normative beliefs rather than relying on formal sanctions (Johnston et al., 2024; bin Mohd Khidir et al., 2025).

Policy clarity remains critical across both approaches. In enforcement-focused institutions, clear rules support consistent application of policies and sanctions (Li et al., 2025; Marín et al., 2025). In academia-led models, clarity provides a shared ethical baseline that enables lecturers to exercise discretion without creating contradictory practices. Many institutions adopt hybrid strategies, combining centralised policy enforcement with faculty-level autonomy, highlighting the need for flexible yet coherent policy design (MQA, 2023; UNESCO, 2023).

The proposed conceptual framework accommodates both regulatory models by focusing on how policy clarity, training, and enforcement interact with behavioural mediators, rather than prescribing a single governance approach (Li et al., 2025; Bittle & El-Gayar, 2025). This ensures the framework is actionable and adaptable, capable of guiding evaluation across diverse institutional contexts while illuminating how different strategies may shape motivation and responsibility in the ethical use of AI.

Theoretical Rationale and Literature Integration

This study builds its framework on organizational governance theory and higher education quality assurance, aiming to show how AI ethics policies within universities can shape academic integrity (Malaysian Qualifications Agency [MQA], 2023; Li et al., 2025). Governance theory suggests that when formal rules and clear structures are established and enforced, institutional actors are more likely to align with shared values and goals (Li et al., 2025; Bittle & El-Gayar, 2025). Within higher education, AI ethics policies therefore function not merely as guidelines but as institutional mechanisms that shape how students engage with emerging technologies while maintaining academic standards (Marín et al., 2025; UNESCO, 2023).

This framework aligns with existing academic integrity literature emphasizing the role of policy clarity, training, and enforcement in shaping ethical behaviour in higher education (Li et al., 2025; Bittle & ElGayar, 2025). Policy clarity reduces ambiguity surrounding acceptable and unacceptable AI use, while AI ethics training enhances students' knowledge and ethical awareness in complex digital learning environments (Bittle & El-Gayar, 2025; Marín et al., 2025). Enforcement mechanisms, including monitoring, reporting, and accountability structures, reinforce compliance and signal institutional commitment to academic standards (Li et al., 2025; Bittle & El-Gayar, 2025).

Illustrative cases from Malaysian universities demonstrate how these principles are applied in practice. At Universiti Teknologi Malaysia (UTM), clear guidelines on generative AI use are complemented by workshops that support students' ethical awareness (MQA, 2023). At Universiti Teknologi MARA (UiTM) and Universiti Malaya (UM), journal policies require disclosure of AI usage and restrict certain AI-assisted practices,

reflecting enforcement in research contexts (Marín et al., 2025). Universiti Malaysia Pahang (UMP) enforces editorial rules to uphold academic integrity, while private institutions such as UNITAR International University implement internal policies clarifying acceptable AI use in assessments (MQA, 2023). Collectively, these practices illustrate how Malaysian institutions are operationalizing clarity, training, and enforcement in line with governance theory and MQA quality standards (MQA, 2023).

The underlying premise of the framework is that clearly articulated and consistently implemented AI ethics policies can shape institutional and student behaviour, support academic integrity, and provide a basis for evaluating policy effectiveness and guiding future improvements (Li et al., 2025; Bittle & ElGayar, 2025; UNESCO, 2023).

Illustrative Application of the Framework Using Institutional Examples

This framework is intended to assess the capacity of different institutions to manage AI ethics (Malaysian Qualifications Agency [MQA], 2023; Li et al., 2025). Higher education institutions seek to establish clear policies and provide AI ethics training for both academic staff and students (Li et al., 2025; Bittle & El-Gayar, 2025). For example, Universiti Teknologi Malaysia (UTM) has developed Guidelines for the Use of Generative Artificial Intelligence in Teaching and Learning, supported by workshops and seminars that enhance awareness of responsible AI use and require students to acknowledge AI contributions in academic work, thereby promoting accountability (MQA, 2023; Marín et al., 2025).

At Universiti Teknologi MARA (UiTM) and Universiti Malaya (UM), journal-level policies require authors to disclose AI use and restrict certain AI-assisted practices to uphold research integrity (Marín et al., 2025; UNESCO, 2023). Universiti Malaysia Pahang (UMP) enforces stricter editorial rules to ensure ethical research practices (Marín et al., 2025). Private institutions, such as UNITAR International University, adopt differentiated approaches in which some assessments prohibit unapproved AI use, while others permit AI assistance only when explicitly allowed and properly disclosed (MQA, 2023; Li et al., 2025). Disclosure requirements and the classification of unapproved AI use as academic misconduct reinforce accountability through enforcement mechanisms, with sanctions ranging from remedial education to formal penalties (Bittle & El-Gayar, 2025; UNESCO, 2023).

These institutional practices illustrate the three pillars of the framework—policy clarity, training, and enforcement—while also highlighting variation in how institutions prioritize and implement these components (Li et al., 2025; Bittle & El-Gayar, 2025; Marín et al., 2025). Some institutions emphasize training and awareness-building, others prioritize enforcement, and a smaller number adopt more balanced approaches. The framework therefore serves both as an analytical tool for evaluating existing institutional practices and as a guide for developing more integrated AI ethics policies. It also identifies directions for future research, including empirical assessment of policy effectiveness and the identification of best practices for aligning institutional AI governance with academic integrity objectives (UNESCO, 2023; Bittle & El-Gayar, 2025).

Implications for Theory and Practice

The conceptual framework and illustrative examples presented in this paper have several important implications for both theory and practice in the context of higher education governance and academic integrity (Malaysian Qualifications Agency [MQA], 2023; Li et al., 2025). From a theoretical perspective, the framework reinforces the utility of organizational governance principles in understanding how institutional policies shape stakeholder behaviour (Li et al., 2025; Bittle and ElGayar, 2025). Practically, the framework provides guidance for university administrators and policymakers seeking to implement AI ethics policies that effectively promote ethical behaviour, offering a structured approach to integrating clarity, training, and enforcement in institutional settings (Li et al., 2025; Bittle and El-Gayar, 2025).

By treating AI ethics policies as a mix of clear rules, training, and enforcement, the framework gives researchers a way to judge if these policies encourage ethical behavior and academic honesty (Li et al., 2025; Bittle and El-Gayar, 2025; Marín et al., 2025). The differences seen across Malaysian universities show the framework is not one-size-fits-all and can flex to fit different cultures, available resources, and governance styles (MQA, 2023; Marín et al., 2025).

This framework does not merely serve as a theory but is intended to be useful for policymakers shaping AI ethics in universities and beyond (UNESCO, 2023; Li et al., 2025). Policymakers, quality assurance teams, and university leaders can use it to build or refine AI ethics policies (MQA, 2023; Li et al., 2025). Awareness of the existence of such guidelines and policies among students and academia contributes to the ethical use of AI technology (Marín et al., 2025; UNESCO, 2023). Training regarding ethical AI use equips students and academia with confidence in using AI technology responsibly (Li et al., 2025; Bittle and El-Gayar, 2025), while enforcement of such policies reinforces accountability in the use of AI technology (Bittle and El-Gayar, 2025; Marín et al., 2025).

The framework is relevant for national quality assurance bodies such as the Malaysian Qualifications Agency (MQA) because it provides a clear way to check whether institutions are following AI guidelines (MQA, 2023; Li et al., 2025). When universities align their AI ethics policies with broader recommendations, they enhance academic integrity while meeting quality assurance expectations (MQA, 2023; UNESCO, 2023). The framework also highlights areas for further research, including examining how different policy components affect academic integrity and comparing best practices for AI governance across universities (Bittle and El-Gayar, 2025; Li et al., 2025).

Framework Limitations

Although the proposed framework of this paper can be used to guide educational institutions in making policies, however it has its limitations. First, institutional diversity across higher education settings may influence how AI ethics policies are designed and implemented. For example, the Malaysian education institutions mentioned in this paper consists of both public and private universities. Malaysian Public and private universities differ significantly in governance structures, funding models, and educational priorities. Public universities operate under stronger state oversight and centralized governance, while private universities function within more market-driven and flexible institutional arrangements. These differences mean that a single framework may not fully capture variation in policy design and implementation across universities. (Li et al., 2025; UNESCO, 2023).

Second, policy enforcement variability represents a key limitation. Enforcement mechanisms may differ between public and private institutions due to variations in institutional autonomy, administrative capacity, and accountability structures. While some institutions may apply AI ethics rules consistently, others may enforce them unevenly or symbolically. Such variability can affect the relationship between policy presence and actual academic integrity outcomes, limiting the framework's ability to account for inconsistent enforcement practices (Bittle & El-Gayar, 2025; Evangelista, 2025).

Third, the framework does not fully account for student agency and resistance. Students are active decision-makers who may interpret, negotiate, or resist institutional rules, particularly in relation to rapidly evolving AI technologies. As a result, the existence of policies alone may not guarantee ethical behaviour (Johnston et al., 2024; bin Mohd Khidir et al., 2025), especially in institutional environments where flexibility or market responsiveness shapes educational delivery.

Finally, cultural differences in ethical norms may shape how students and institutions understand academic integrity and acceptable AI use. Ethical expectations are influenced by broader cultural, social, and educational contexts, which may differ across institutions and sectors. These variations may limit the transferability of the framework across different national, institutional, or cultural environments. (Marín et al., 2025; UNESCO, 2023).

CONCLUSION

This paper has proposed a conceptual framework for evaluating the effectiveness of institutional AI ethics policies in promoting academic integrity (Li et al., 2025; Bittle and El-Gayar, 2025). By conceptualizing AI ethics policies as a multidimensional construct comprising policy clarity, AI ethics training, and enforcement mechanisms, the framework provides both a theoretical and practical lens for understanding how higher education institutions can guide ethical behaviour in the use of AI (Li et al., 2025; Marín et al., 2025; Bittle and El-Gayar, 2025). Illustrative examples from Malaysian universities, including UTM, UiTM, UM, UPM,

and private institutions such as UNITAR International University, show that institutions are beginning to operationalize these dimensions in varied ways, highlighting both opportunities and challenges in policy design and implementation (Malaysian Qualifications Agency [MQA], 2023; Marín et al., 2025; UNESCO, 2023).

The framework is relevant for scholars, policymakers, and practitioners (Li et al., 2025; Bittle and ElGayar, 2025). Theoretically, it demonstrates how governance and quality assurance shape stakeholder behaviour (Li et al., 2025; Malaysian Qualifications Agency [MQA], 2023). Practically, it provides guidance for building, assessing, and improving AI ethics policies, emphasizing the importance of aligning institutional policies with national guidance, such as that provided by the Malaysian Qualifications Agency (MQA, 2023; UNESCO, 2023). This alignment supports academic integrity and reinforces institutional credibility (Marín et al., 2025; UNESCO, 2023).

Finally, the framework offers a foundation for future empirical research, including studies that assess the relative effectiveness of each policy dimension, examine cross-institutional variations, and identify best practices in AI governance in higher education (Li et al., 2025; Bittle and El-Gayar, 2025; Marín et al., 2025). By systematically linking policy structure to ethical outcomes, this conceptual paper contributes to the emerging discourse on responsible AI integration in universities and offers guidance for institutions navigating the complex ethical landscape of AI-assisted teaching, learning, and research (UNESCO, 2023; Li et al., 2025). Minor writing and formatting assistance for this manuscript was provided using ChatGPT (OpenAI, 2026).

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. bin Mohd Khidir, M. L., bin Sa'ari, S. N., & bin Mohd Mokhtar, M. F. (2025). Academic integrity in the age of generative AI: Ethical challenges and perceptions among TVET students at POLIMAS. Journal on Technical and Vocational Education, 10(2), 398–409.
3. 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), ep559.
4. Johnston, H., Wells, R. F., Shanks, E. M., Boey, T., & Parsons, B. N. (2024). Student perspectives on the use of generative artificial intelligence technologies in higher education. International Journal for Educational Integrity, 20(1), 2.
5. Li, M., Xie, Q., Enkhtur, A., Meng, S., Chen, L., Yamamoto, B. A., & Murakami, M. (2025). A framework for developing university policies on generative AI governance: A cross-national comparative study. arXiv. https://arxiv.org/abs/2504.02636
6. Malaysian Qualifications Agency. (2023). Advisory note no. 2/2023: Penggunaan Teknologi Kecerdasan Buatan Generatif (Generative Artificial Intelligence) Dalam Pendidikan Tinggi. Malaysian Qualifications Agency. https://www.mqa.gov.my/new/pubs_adv_notes_2023.cfm
7. Marín, Y. R., Caro, O. C., Rituay, A. M. C., Llanos, K. A. G., Perez, D. T., Bardales, E. S., ... & Santos, R. C. (2025). Ethical challenges associated with the use of artificial intelligence in university education. Journal of Academic Ethics, 1–25.
8. UNESCO. (2023). Guidance for generative AI in education and research. UNESCO. https://www.unesco.org/en/articles/guidance-generative-ai-education-and-research