

Academic Integrity in Higher Education: Challenges, Opportunities, and Policy Directions

¹Dr. Ayesha Butt, ²Sadia Rafique

¹Independent Researcher
Sydney, Australia

²Lecturer, Department of English, National University of Modern Languages, Lahore, Pakistan

Email: ayeshabuttlcwu@gmail.com

ORCID ID: <https://orcid.org/0000-0003-2788-8372>

Abstract

Academic integrity has become a pervasive and challenging concern in higher education with the introduction of Generative Artificial Intelligence (GenAI) tools such as ChatGPT, Gemini, Claude, and other large language models (LLMs). While this technological shift has created significant opportunities for personalised learning, accessibility, and intensive feedback, it has simultaneously disrupted traditional notions of authorship and the reliability and validity of assessment. This article examines academic integrity through three analytical lenses: (1) the pre-AI period (2017–2022), characterised by concerns related to plagiarism, contract cheating, and essay mills; (2) the post-AI period from 2022 onwards, marked by AI-generated content, unreliable detection technologies, and evolving regulatory uncertainty; and (3) a comparative international perspective drawing on policy and practice in Australia, the United Kingdom, the United States, Hong Kong, and Pakistan. Using a mixed-methods policy analysis of regulatory documents, empirical research, and national-level institutional case examples, the study identifies a convergence towards pedagogical approaches that embed AI literacy, promote transparency and disclosure, emphasise learning processes over final products, and reduce dependence on automated detection systems. The article concludes that sustaining academic integrity in the GenAI era requires a fundamental shift from rule-based compliance models to pedagogies of responsible, ethical, and critically informed AI engagement.

Keywords: generative artificial intelligence, academic integrity, higher education, AI literacy, assessment design, policy analysis

1. Introduction

Following the public release of ChatGPT in 2022, students and academic staff across the globe rapidly began experimenting with new ways of generating ideas, summarising academic texts, producing code, and drafting essays. This marked a transformative moment in the integration of artificial intelligence into higher education (Kasneci et al., 2023; Tlili et al., 2023). Within a year, major technology corporations such as Google and Microsoft had launched their own large-scale language models, explicitly targeting educational users (Dwivedi et al., 2023).

The accelerated adoption of GenAI introduced both unprecedented pedagogical opportunities and profound challenges. On the one hand, AI tools have enhanced productivity, supported personalised learning, and improved accessibility for students with diverse linguistic, cognitive, and physical needs (Zawacki-Richter et al., 2019; UNESCO, 2023). On the other hand, they have

destabilised long-standing assumptions about originality, authorship, and the effectiveness of plagiarism detection systems, raising serious concerns about the validity and fairness of take-home and unsupervised assessments (Cotton et al., 2023).

This technological disruption has required a fundamental re-examination of the core values of academic integrity, including honesty, trust, fairness, responsibility, and respect (International Center for Academic Integrity [ICAI], 2021). Importantly, integrity challenges pre-date GenAI. Prior to 2022, higher education systems worldwide were already grappling with contract cheating, essay mills, and collusion, which prompted regulatory, pedagogical, and legal responses across multiple national contexts (Bretag et al., 2019; Dawson, 2021). However, GenAI has altered the scale, speed, and subtlety of potential misconduct, rendering detection-focused approaches

increasingly unreliable and ethically problematic (Liang et al., 2023).

The discourse has consequently expanded beyond individual misconduct to encompass broader questions of policy, pedagogy, equity, and professional ethics. National quality assurance and regulatory bodies in Australia, the United Kingdom, the United States, Hong Kong, and Pakistan have issued guidance urging institutions to redesign assessment, integrate AI literacy, and avoid over-reliance on automated detection technologies (QAA, 2023; TEQSA, 2023; UNESCO, 2023). Across these jurisdictions, institutions face a complex balancing act: prohibiting AI risks disadvantaging graduates in AI-saturated labour markets, while unregulated use threatens the credibility of qualifications and learning outcomes (Selwyn et al., 2023).

This article addresses these tensions through a comparative, evidence-based analysis of academic integrity in the GenAI era. Specifically, it:

1. Reviews the pre-AI integrity landscape (2017–2022), focusing on systemic challenges such as contract cheating and the pedagogical lessons derived from earlier reform efforts.
2. Analyses post-2022 integrity challenges, including the limitations of AI-detection tools, shifting student practices, and emerging concerns around authorship and accountability.
3. Compares policy responses across Australia, the United Kingdom, the United States, Hong Kong, and Pakistan, drawing on regulatory frameworks and sector-wide initiatives.
4. Examines national-level case examples illustrating how integrity policies are operationalised in diverse educational systems.
5. Proposes an integrity-by-design framework to guide institutions in embedding ethical and transparent AI use within curriculum and assessment.

Although global in scope, the analysis pays particular attention to comparative dynamics between developed and developing higher education systems, enabling insights into how regulatory capacity, digital infrastructure, and

linguistic diversity shape integrity responses. The inclusion of Pakistan alongside Australia, the United Kingdom, the United States, and Hong Kong allows for examination of how GenAI challenges intersect with issues of access, language justice, and academic capacity building in emerging knowledge economies. Collectively, the findings contribute to a growing body of scholarship advocating context-sensitive, pedagogically grounded approaches that preserve academic integrity while harnessing the educational potential of generative artificial intelligence.

2. Literature Review

2.1 Academic Integrity in the Pre-AI Context (2017–2022)

2.1.1 Core values and frameworks

Between 2017 and 2022, academic integrity in higher education was primarily conceptualised through shared ethical values and preventative institutional cultures. The International Center for Academic Integrity (ICAI, 2021) articulated six foundational values; honesty, trust, fairness, respect, responsibility, and courage, which became global reference points for policy development and educational practice. These values informed orientation programs, academic skills curricula, and misconduct procedures across diverse national systems.

Regulatory and quality assurance bodies reinforced these principles. In Australia, integrity requirements were embedded within the Higher Education Standards Framework, obligating providers to ensure responsible academic conduct and robust assessment design (TEQSA, 2021). In the United Kingdom, national quality agencies issued sector-wide guidance on preventing contract cheating and the use of third-party services (QAA, 2020). In the United States, where higher education is decentralised and lacks a single regulator, professional associations and accreditation bodies played a central role in promoting integrity norms and institutional self-regulation. In Hong Kong, quality assurance mechanisms emphasised academic honesty as a core graduate attribute aligned with international benchmarks. In Pakistan, national higher education authorities similarly highlighted plagiarism prevention, ethical research practices, and the development of academic writing

competencies as foundational to quality assurance and international recognition.

Across these contexts, integrity was increasingly framed as a cultural and educational issue rather than solely a disciplinary one, with emphasis placed on prevention, capacity building, and student engagement rather than surveillance-based enforcement.

2.1.2 Contract cheating and essay mills

Prior to the emergence of GenAI, contract cheating represented one of the most significant threats to academic integrity globally. Empirical studies in Australia indicated that a measurable proportion of students admitted to outsourcing assignments to commercial providers (Bretag et al., 2019), while similar patterns were reported in the United Kingdom, the United States, Hong Kong, and Pakistan through institutional investigations and sector reports. Essay mills operated transnationally via digital platforms, exploiting regulatory gaps, linguistic insecurity, and high-stakes assessment environments (Harper et al., 2020).

Research identified opportunity structures that facilitated misconduct, including large class sizes, limited formative feedback, generic assessment tasks, and a focus on product rather than process (Dawson, 2021). In multilingual and developing-country contexts such as Pakistan, additional pressures related to academic English proficiency, limited access to writing support, and uneven digital infrastructure further heightened vulnerability to third-party assistance. These findings shifted the analytical focus from individual moral failure to systemic design flaws, establishing the foundation for later critiques of assessment vulnerability in the GenAI era.

2.1.3 Assessment design and prevention strategies

In response, the literature increasingly advocated for authentic, process-oriented assessment. Ellis et al. (2020) demonstrated that tasks requiring contextualisation, reflection, and staged development significantly reduced the likelihood of outsourcing. Recommended strategies included iterative drafting, reflective commentaries, oral presentations, in-class problem solving, and personalised case-based assignments.

The rapid digitalisation of higher education during the COVID-19 pandemic further intensified these debates. Remote assessment, online proctoring, and algorithmic plagiarism detection became widespread, particularly in Australia, the United Kingdom, Hong Kong, and parts of the United States, while in Pakistan emergency online provision exposed significant inequalities in access and digital literacy (Lancaster & Cotarlan, 2021). Concerns emerged regarding privacy, data security, cultural bias, and the ethics of surveillance-based integrity systems (Selwyn et al., 2020).

Although artificial intelligence featured in pre-2022 discussions, its role was largely limited to narrow applications such as plagiarism detection, adaptive learning platforms, and learning analytics. The possibility that AI systems could generate original, discipline-appropriate academic text at scale had not yet been fully anticipated in policy or pedagogical design, leaving higher education systems globally underprepared for the disruption that followed.

2.2 Academic Integrity in the GenAI Era (2022–2025)

2.2.1 The rise of generative AI in education

Following the public release of large language models in late 2022, rapid adoption was documented across higher education systems in Australia, the United Kingdom, the United States, Hong Kong, and Pakistan (Educause, 2023; Montacute, 2023; Selwyn et al., 2023). Unlike contract cheating, GenAI does not rely on human intermediaries, complicating established definitions of plagiarism, collusion, and unauthorised assistance. Its capacity to produce fluent academic prose, discipline-specific argumentation, and executable code within seconds challenged the validity of assessment tasks designed under pre-AI assumptions.

Key affordances included automated drafting, disciplinary style emulation, coding support, data summarisation, and multilingual translation. While these functions offered substantial learning support, particularly for second-language writers and students in resource-constrained contexts such as Pakistan, they simultaneously destabilised traditional markers of individual authorship and independent work.

2.2.2 AI detection: reliability and bias issues

Initial policy responses across many systems involved the deployment of AI-text detection tools. However, empirical studies demonstrated that these systems produced high rates of false positives and exhibited systematic bias against multilingual writers (Liang et al., 2023). The withdrawal of OpenAI's own classifier due to low reliability further undermined confidence in automated detection (OpenAI, 2023).

These findings were particularly consequential for international students in Anglophone systems and for English-medium instruction contexts in Hong Kong and Pakistan, where linguistic diversity is high. Consequently, institutions across the five national contexts increasingly cautioned against the use of detection scores as sole evidence of misconduct and emphasised triangulation through draft analysis, oral verification, and reflective explanation of learning processes.

2.2.3 Policy uncertainty and variation

Between 2023 and 2025, national and institutional policies evolved rapidly but unevenly. Two key dimensions characterised this evolution: degrees of permissibility and expectations of disclosure. While early responses in many countries involved precautionary restrictions or temporary bans, later policies increasingly permitted conditional use for formative purposes, language support, and ideation, accompanied by requirements for transparency and ethical justification.

The extent of regulatory coherence varied. Systems with strong central quality agencies, such as Australia, the United Kingdom, and Hong Kong, moved more quickly towards sector-wide guidance, whereas the decentralised United States context and the developing regulatory environment in Pakistan exhibited greater institutional variation. Nonetheless, a shared trajectory emerged towards integrating AI literacy, clarifying acceptable use, and reframing integrity as a pedagogical rather than purely disciplinary issue.

2.2.4 Ethical and equity considerations

Across all five contexts, GenAI raised critical ethical questions related to algorithmic bias, unequal access to advanced tools, language justice, and intellectual property. International bodies such as UNESCO (2023) and the OECD (2024)

emphasised the importance of transparency, human oversight, and equitable access, warning that uncritical adoption could entrench existing educational inequalities. These concerns were particularly salient in Pakistan and other developing systems, where disparities in infrastructure, subscription access, and academic support risk creating a two-tier AI-enabled learning environment.

3. Comparative Policy Analysis

3.1 Australia

Australia's higher education sector has responded proactively to the challenges posed by GenAI, supported by a strong national regulatory framework and coordinated sector action. National guidance has urged institutions to redesign assessment, embed AI literacy as a graduate capability, and avoid over-reliance on automated detection tools, particularly in high-stakes decision-making (TEQSA, 2023; Selwyn et al., 2023; Cotton et al., 2023). This regulatory stance reflects earlier integrity reforms targeting contract cheating and assessment vulnerability, where emphasis was placed on process-oriented assessment and educational rather than punitive approaches (Bretag et al., 2019; Dawson, 2021).

At a sector level, Australian universities have engaged in collaborative working groups to share policy models, staff development resources, and assessment exemplars aligned with ethical AI use and procedural fairness (Kasneci et al., 2023; UNESCO, 2023). These initiatives aim to harmonise expectations across institutions while supporting disciplinary variation and academic autonomy. Collectively, the Australian response reflects a shift from surveillance-based integrity enforcement towards pedagogically grounded, literacy-focused approaches to GenAI integration (TEQSA, 2023; Zawacki-Richter et al., 2019).

3.2 United Kingdom

In the United Kingdom, national policy direction has been shaped by quality assurance guidance and sector-wide collaboration. Regulatory advisories issued since 2023 have focused on assessment validity, transparency of AI use, and fairness in misconduct investigation processes, particularly in relation to the known limitations and biases of detection technologies (QAA, 2023; Liang et al.,

2023; Selwyn et al., 2023). These policies build on earlier work addressing contract cheating and third-party services, which emphasised prevention through assessment redesign and academic skills development (QAA, 2020; Ellis et al., 2020).

Sector-level principles have promoted the integration of AI literacy into curricula, the articulation of acceptable and unacceptable uses, and regular review of assessment practices to ensure alignment with learning outcomes rather than tool capabilities (Kasneci et al., 2023; UNESCO, 2023; Cotton et al., 2023). The UK approach thus combines regulatory oversight with collective norm-setting, reinforcing integrity as a shared pedagogical responsibility rather than solely a compliance issue (Dawson, 2021; ICAI, 2021).

3.3 United States

The policy landscape in the United States is characterised by institutional autonomy and regulatory decentralisation, resulting in diverse approaches to GenAI governance. While no single national framework exists, professional associations and leading institutions have issued guidance cautioning against the sole use of AI-detection tools in misconduct cases due to reliability and equity concerns (Liang et al., 2023; OpenAI, 2023; Educause, 2023). Instead, emphasis has been placed on triangulation of evidence, educator judgment, and due process (Dawson, 2021; Selwyn et al., 2023).

Accreditation bodies in professionally regulated disciplines have also influenced policy by requiring assessment practices that ensure independent competence and ethical accountability (Kasneci et al., 2023; UNESCO, 2023). Faculty development initiatives and interdisciplinary task forces have further supported pedagogical innovation, including AI-inclusive assessment design and student disclosure practices (Cotton et al., 2023; Zawacki-Richter et al., 2019; ICAI, 2021). Despite variability, a common trend is emerging towards process-based assessment and critical AI literacy as foundations for academic integrity in the GenAI era.

3.4 Hong Kong

Hong Kong's higher education system, operating within a centrally coordinated quality assurance environment, has positioned AI literacy as a strategic graduate attribute and an essential

component of academic integrity policy (UNESCO, 2023; Kasneci et al., 2023; Selwyn et al., 2023). Sector guidance emphasises ethical use, transparency, and student capability to critically evaluate AI outputs, alongside formal requirements for disclosure and attribution (ICAI, 2021; Liang et al., 2023).

Assessment policies increasingly prioritise reflective engagement with AI, staged submission, and oral or applied verification of learning, reflecting international concerns about detection reliability and linguistic bias (OpenAI, 2023; Cotton et al., 2023; Ellis et al., 2020). Hong Kong's regulatory coherence has enabled relatively rapid alignment between policy, curriculum, and staff development, supporting a systemic shift from control-oriented responses to literacy- and design-based integrity frameworks (Zawacki-Richter et al., 2019; UNESCO, 2023).

3.5 Pakistan

In Pakistan, the emergence of GenAI has intersected with longstanding priorities related to plagiarism prevention, academic writing development, and quality assurance in English-medium higher education. National policy has traditionally focused on originality verification and ethical research practice, but since 2023 discourse has expanded to include responsible AI use, academic honesty, and staff capacity building (UNESCO, 2023; Kasneci et al., 2023; Selwyn et al., 2023).

Institutional responses have begun to incorporate AI literacy into academic skills curricula, revise misconduct regulations to clarify AI-related expectations, and promote assessment designs that emphasise drafting, reflection, and oral explanation (Ellis et al., 2020; Dawson, 2021; ICAI, 2021). Equity considerations are particularly salient, as unequal access to advanced AI tools and digital infrastructure risks creating a two-tier learning environment (Liang et al., 2023; Zawacki-Richter et al., 2019; Cotton et al., 2023). Consequently, emerging policy debates in Pakistan increasingly frame GenAI not only as an integrity challenge but also as a question of educational justice, access, and capacity development within a globalised academic system.

4. Case Studies

4.1 Australia – From Prohibition to Conditional Integration

In Australia, early institutional reactions to GenAI in 2023 were largely precautionary, with many providers discouraging or temporarily restricting AI use due to concerns about assessment validity and the unreliability of detection technologies (TEQSA, 2023; Cotton et al., 2023). These initial responses mirrored earlier approaches to contract cheating, where control-oriented strategies dominated before pedagogical redesign gained prominence (Bretag et al., 2019; Dawson, 2021). By 2024, policy direction shifted towards conditional integration. National guidance encouraged transparent disclosure of AI use, integration of AI literacy into first-year curricula, and redesign of assessment to foreground learning processes, reflective engagement, and oral verification (Selwyn et al., 2023; Kasneci et al., 2023; UNESCO, 2023). Professional development programs for academic staff focused on ethical AI use, assessment scaffolding, and critical evaluation of AI-assisted work, reflecting a transition from detection-based enforcement to trust-plus-accountability models (ICAI, 2021; Zawacki-Richter et al., 2019).

4.2 United Kingdom – Embedding Sector Principles in Curriculum and Assessment

Across the United Kingdom, sector-wide principles on generative AI have been operationalised through curriculum-embedded AI literacy, discipline-specific guidance on acceptable use, and iterative policy review mechanisms (QAA, 2023; Kasneci et al., 2023; Selwyn et al., 2023). Institutions have incorporated mandatory instruction on ethical AI use, bias, and authorship into first-year programs, positioning AI literacy alongside academic writing and information literacy as a foundational skill (UNESCO, 2023; ICAI, 2021).

Assessment practices increasingly require reflective commentary on AI use, staged drafting, and contextualised tasks that emphasise critical thinking and disciplinary reasoning rather than surface-level text production (Ellis et al., 2020; Cotton et al., 2023; Dawson, 2021). These reforms align with earlier anti-contract-cheating strategies and respond to evidence on the limitations and inequities of AI detection tools, particularly for

multilingual students (Liang et al., 2023; OpenAI, 2023).

4.3 United States – Rejecting Over-Reliance on Detection Technologies

In the United States, several institutions have publicly moved away from automated AI-detection systems due to concerns about false positives, algorithmic bias, and due-process implications (Liang et al., 2023; OpenAI, 2023; Educause, 2023). Instead, integrity investigations increasingly rely on triangulated evidence, including draft histories, oral explanations, learning analytics, and reflective justification of AI use (Dawson, 2021; Selwyn et al., 2023).

Faculty development initiatives and interdisciplinary AI task forces have supported the redesign of assessment to include process documentation, oral defences, and authentic problem-based tasks, consistent with broader pedagogical calls for integrity-by-design (Kasneci et al., 2023; Cotton et al., 2023; Zawacki-Richter et al., 2019). Professional accreditation requirements in fields such as engineering, medicine, and law have further reinforced the need for demonstrable independent competence, shaping national discourse on ethical AI integration (ICAI, 2021; UNESCO, 2023).

4.4 Hong Kong – AI as a Graduate Literacy

Hong Kong's higher education system has framed AI literacy as a core graduate attribute, integrating technical understanding, ethical reasoning, and critical evaluation of algorithmic outputs into compulsory undergraduate curricula (UNESCO, 2023; Kasneci et al., 2023; Selwyn et al., 2023). National policies require explicit attribution and disclosure of AI use, supported by staff training and centrally coordinated learning resources (ICAI, 2021; Liang et al., 2023).

Assessment reforms emphasise reflective engagement with AI, bias identification, and oral or applied verification of learning, responding to international evidence on detection unreliability and linguistic discrimination (Ellis et al., 2020; Cotton et al., 2023; OpenAI, 2023). Continuous policy review cycles and systematic feedback from students and staff enable rapid adaptation to technological change, aligning integrity

governance with global ethical standards (Zawacki-Richter et al., 2019; UNESCO, 2023).

4.5 Pakistan – Capacity Building and Equity-Focused Integration

In Pakistan, the GenAI transition has intersected with long-standing priorities in plagiarism control, academic writing development, and quality assurance within English-medium higher education (UNESCO, 2023; Kasneci et al., 2023; Selwyn et al., 2023). Emerging national and institutional policies emphasise responsible AI use, transparency, and the integration of AI literacy into academic skills courses, particularly to support second-language writers and first-generation university students (ICAI, 2021; Ellis et al., 2020). Assessment reforms increasingly promote staged drafting, reflective explanation, and oral defence to verify authorship and conceptual understanding, aligning with international integrity-by-design principles (Dawson, 2021; Cotton et al., 2023; Zawacki-Richter et al., 2019). Equity concerns are central, as unequal access to premium AI tools and digital infrastructure risks deepening existing educational disparities (Liang et al., 2023; UNESCO, 2023; Kasneci et al., 2023). Consequently, GenAI policy discourse in Pakistan frames academic integrity not only as a compliance issue but also as a matter of access and inclusion.

5. Discussion

5.1 Shifting from a Policing Paradigm to a Pedagogy Paradigm

Comparative evidence from Australia, the United Kingdom, the United States, Hong Kong, and Pakistan indicates a clear transition from surveillance-oriented integrity enforcement to pedagogically grounded approaches that acknowledge the permanence of GenAI in academic practice. Early institutional reactions in 2023 were largely precautionary and control-focused, characterised by temporary bans, strict prohibition language in assessment briefs, and reliance on automated detection technologies (Cotton et al., 2023; Selwyn et al., 2023; TEQSA, 2023). These responses, while understandable, reproduced limitations previously observed in the contract cheating literature, where detection and punishment alone proved insufficient to sustain

academic honesty (Bretag et al., 2019; Dawson, 2021).

Subsequent policy evolution across the five national contexts has increasingly emphasised pedagogical strategies that integrate AI literacy, transparency, and assessment redesign. The emerging pedagogy paradigm frames GenAI as a legitimate learning tool whose ethical use must be explicitly taught, critically examined, and reflected upon (Kasneci et al., 2023; UNESCO, 2023; ICAI, 2021). This shift is evident in curriculum-embedded AI literacy, the requirement for disclosure and reflective justification of AI use, and the move towards process-oriented assessment that makes learning visible (Ellis et al., 2020; Cotton et al., 2023; Selwyn et al., 2023). Across both developed and developing systems, this paradigm supports a trust-plus-accountability model in which integrity is cultivated through design, dialogue, and education rather than surveillance alone.

5.2 Rethinking Assessment Validity in the GenAI Era

Assessment validity in the GenAI context depends on ensuring that tasks measure student learning rather than the output capabilities of AI systems. Across Australia, the United Kingdom, the United States, Hong Kong, and Pakistan, reforms increasingly emphasise multi-stage submissions, reflective commentaries, oral defences, and in-class applications of learning to verify conceptual understanding and authorship (Ellis et al., 2020; Dawson, 2021; Kasneci et al., 2023). These strategies align with pre-AI integrity research demonstrating that authentic, personalised, and process-focused assessment reduces opportunities for outsourcing and superficial performance (Bretag et al., 2019; Cotton et al., 2023).

The limitations and biases of AI-detection tools further reinforce the need for design-based solutions. Empirical evidence shows that automated classifiers produce high rates of false positives and disproportionately flag multilingual writers, raising concerns of procedural injustice, particularly in linguistically diverse systems such as Hong Kong and Pakistan (Liang et al., 2023; OpenAI, 2023; Selwyn et al., 2023). Consequently, assessment practices that require students to explain, defend, and contextualise their work—rather than merely submit polished products—are

increasingly recognised as central to maintaining both validity and fairness (ICAI, 2021; Zawacki-Richter et al., 2019; UNESCO, 2023).

5.3 Equity and Ethical Protection

GenAI presents both equity-enhancing opportunities and significant ethical risks. Across all five national contexts, AI tools have demonstrated potential to support multilingual students, learners with disabilities, and those with limited access to academic skills support by providing language scaffolding, summarisation, and formative feedback (Zawacki-Richter et al., 2019; Kasneci et al., 2023; UNESCO, 2023). In developing systems such as Pakistan, these affordances may help mitigate resource constraints and expand access to academic discourse.

However, risks of algorithmic bias, unequal access to premium tools, and linguistic discrimination remain substantial (Liang et al., 2023; Cotton et al., 2023; Selwyn et al., 2023). Without institutional licensing, open-access alternatives, and explicit instruction in critical AI use, GenAI may reinforce socio-economic and cultural inequalities. Ethical governance therefore requires systematic auditing of tools, equitable access policies, and curriculum-embedded discussions of bias, data provenance, and intellectual property (ICAI, 2021; UNESCO, 2023; OECD, 2024). These safeguards are particularly crucial in multilingual and postcolonial contexts, where Western-trained models may marginalise local epistemologies and linguistic identities.

5.4 The Role of Sector Bodies and Regulatory Coherence

The presence of strong national quality assurance agencies in Australia, the United Kingdom, and Hong Kong has facilitated more coherent and timely policy responses to GenAI, enabling sector-wide guidance, shared principles, and coordinated professional development (TEQSA, 2023; QAA, 2023; UNESCO, 2023). In contrast, the decentralised regulatory structure of the United States and the evolving quality assurance framework in Pakistan have resulted in greater institutional variability, though also allowing for local experimentation and innovation (Educause, 2023; Selwyn et al., 2023).

Across all contexts, professional associations, accreditation bodies, and international organisations have played a critical convening role, disseminating best practice, developing ethical frameworks, and promoting alignment with global standards of transparency, fairness, and human oversight (ICAI, 2021; OECD, 2024; UNESCO, 2023). These bodies help bridge gaps between rapid technological change and slower institutional policy cycles, supporting the development of integrity-by-design approaches that integrate assessment reform, AI literacy, and equity considerations at system level.

6. Recommendations: An Integrity-by-Design Framework for the GenAI Era

Drawing on the literature, comparative policy analysis, and country-level case studies from Australia, the United Kingdom, the United States, Hong Kong, and Pakistan, this section proposes an Integrity-by-Design framework to guide higher education institutions in embedding academic integrity within teaching, learning, and assessment in the GenAI era. Rather than positioning AI as an external threat to be controlled, the framework emphasises pedagogical integration, transparency, and ethical capacity building (Bretag et al., 2019; Dawson, 2021; UNESCO, 2023).

6.1 Transparent AI-Use Policies

Institutions should articulate clear, accessible policies that define acceptable, conditional, and prohibited uses of GenAI in assessment, supported by discipline-specific exemplars (ICAI, 2021; QAA, 2023; TEQSA, 2023). Ambiguity has been consistently identified as a driver of unintentional misconduct, particularly for international and first-generation students (Selwyn et al., 2023; Kasneci et al., 2023). Policies should be reviewed regularly

to reflect technological developments, emerging ethical concerns, and stakeholder feedback (UNESCO, 2023; OECD, 2024).

6.2 Mandatory AI-Use Disclosure

Requiring students to declare how, when, and for what purposes AI tools were used promotes transparency and metacognitive awareness, and normalises ethical reflection rather than concealment (Cotton et al., 2023; Kasneci et al., 2023; ICAI, 2021). Disclosure statements should be accompanied by guidance and exemplars, enabling students to distinguish between supportive and substitutive AI use and to justify their evaluative decisions regarding AI-generated outputs (Ellis et al., 2020; Dawson, 2021; Selwyn et al., 2023).

6.3 Process-Oriented Assessment Design

Assessment should be redesigned to foreground learning processes through staged submissions, reflective commentaries, oral defences, and in-class applications (Bretag et al., 2019; Ellis et al., 2020; Kasneci et al., 2023). Such designs reduce the viability of unacknowledged AI substitution and provide richer evidence of conceptual understanding, critical thinking, and disciplinary reasoning (Dawson, 2021; Cotton et al., 2023; Zawacki-Richter et al., 2019).

6.4 Reduced Reliance on Automated Detection

Given the documented inaccuracy and linguistic bias of AI-detection tools, their use should be limited to supplementary indicators rather than primary evidence in misconduct determinations (Liang et al., 2023; OpenAI, 2023; Selwyn et al., 2023). Institutions should train staff in evidence triangulation, including draft analysis, oral verification, and reflective explanation, to ensure procedural fairness and equity (ICAI, 2021; UNESCO, 2023; Dawson, 2021).

6.5 Curriculum-Embedded AI Literacy

Critical AI literacy should be integrated across all disciplines, addressing technical affordances, limitations, bias, data ethics, and intellectual property (Kasneci et al., 2023; UNESCO, 2023; OECD, 2024). Professional development for academic and professional staff is equally essential to support ethical assessment design, transparent

communication, and inclusive pedagogical practice (Zawacki-Richter et al., 2019; Selwyn et al., 2023; QAA, 2023).

6.6 Equity and Access Safeguards

Institutions must monitor differential impacts of GenAI on multilingual students, students with disabilities, and those from low-income or digitally marginalised backgrounds (Liang et al., 2023; Cotton et al., 2023; UNESCO, 2023). Providing institutionally licensed tools, open-access alternatives, and Universal Design for Learning (UDL)-aligned practices can mitigate the risk of a two-tier AI-enabled learning environment, particularly in developing systems such as Pakistan (Zawacki-Richter et al., 2019; Kasneci et al., 2023; OECD, 2024).

6.7 Sector Collaboration and International Alignment

Cross-institutional communities of practice, supported by national quality agencies and international bodies, can facilitate the sharing of policy templates, assessment exemplars, and staff development resources (TEQSA, 2023; QAA, 2023; UNESCO, 2023). Engagement with global frameworks ensures alignment with emerging ethical standards and supports smaller or resource-constrained institutions in building coherent and future-oriented integrity strategies (ICAI, 2021; OECD, 2024; Selwyn et al., 2023).

Collectively, these recommendations promote a trust-plus-accountability model in which academic integrity is embedded through curriculum, assessment design, and institutional culture. By integrating transparency, literacy, equity, and pedagogical innovation, higher education systems across diverse national contexts can ensure that GenAI strengthens rather than undermines the credibility, fairness, and educational purpose of assessment in the digital age.

7. Conclusion

The rapid emergence of generative artificial intelligence has compelled higher education systems in Australia, the United Kingdom, the United States, Hong Kong, and Pakistan to re-examine long-established understandings of academic integrity. The challenges associated with unreliable detection technologies, shifting notions

of authorship, and widening equity gaps are substantial. However, the comparative evidence presented in this article demonstrates that these challenges also create opportunities to strengthen pedagogy, redesign assessment, and embed ethical and critical AI engagement within the core purposes of higher education (Kasneci et al., 2023; Selwyn et al., 2023; UNESCO, 2023).

Lessons from the pre-AI era highlight that integrity cannot be sustained through surveillance and prohibition alone. Research on contract cheating and assessment vulnerability consistently showed that durable integrity is cultivated through transparent expectations, authentic assessment, and educational cultures that emphasise responsibility and trust (Bretag et al., 2019; Dawson, 2021; ICAI, 2021). The GenAI era has reinforced this insight. Automated detection tools, now known to be unreliable and linguistically biased, are insufficient as primary integrity mechanisms and risk undermining procedural fairness, particularly for multilingual and marginalised students (Liang et al., 2023; OpenAI, 2023; Cotton et al., 2023).

Across the five national contexts examined, a clear shift is evident from a policing paradigm towards a pedagogy paradigm. This transition is characterised by the integration of AI literacy into curricula, explicit guidance on acceptable and ethical AI use, disclosure-based transparency, and process-oriented assessment design that prioritises critical thinking, reflection, and demonstrable understanding (Ellis et al., 2020; Kasneci et al., 2023; TEQSA, 2023). Such approaches recognise GenAI not as a temporary disruption but as a permanent feature of academic and professional life, requiring graduates to develop the capacity to work critically, responsibly, and ethically with intelligent systems.

Equity considerations are central to this transformation. While GenAI can enhance access and support for students in linguistically and resource-diverse contexts, including those in developing systems such as Pakistan, it also risks reinforcing socio-economic and epistemic inequalities if access, bias, and digital capability are not systematically addressed (Zawacki-Richter et al., 2019; UNESCO, 2023; OECD, 2024). Ethical AI governance in higher education must therefore be inseparable from broader commitments to

inclusion, language justice, and universal design for learning.

Ultimately, sustaining academic integrity in the age of generative AI requires a move beyond rule-based compliance towards integrity-by-design. This involves aligning policy, curriculum, assessment, and professional development around shared principles of transparency, critical literacy, fairness, and educational purpose. Institutions that adopt such integrated, context-sensitive approaches will be better positioned not only to protect the credibility of their qualifications but also to prepare graduates for ethical participation in an AI-saturated world. In this way, GenAI can become a catalyst for renewing, rather than eroding, the foundational values of higher education.

References

- Bretag, T., Harper, R., Burton, M., Ellis, C., Newton, P., Rozenberg, P., Saddiqui, S., & van Haeringen, K. (2019). Contract cheating: A survey of Australian university students. *Studies in Higher Education*, 44(11), 1837–1856.
<https://doi.org/10.1080/03075079.2018.1462788>
- Cotton, D. R. E., Cotton, P. A., & Shipway, J. R. (2023). Chatting and cheating: Ensuring academic integrity in the era of ChatGPT. *Innovations in Education and Teaching International*, 60(2), 195–207.
<https://doi.org/10.1080/14703297.2023.2190148>
- Dawson, P. (2021). *Defending assessment security in a digital world: Preventing e-cheating and supporting academic integrity in higher education*. Routledge.
<https://doi.org/10.4324/9780429324178>
- Dwivedi, Y. K., Hughes, L., Baabdullah, A. M., Ribeiro-Navarrete, S., Giannakis, M., Al-Debei, M. M., Dennehy, D., Metri, B., Buhalis, D., Cheung, C. M. K., Conboy, K., Doyle, R., Dubey, R., Dwivedi, R., Edwards, J., Gupta, B., Gutierrez, A., Ilavarasan, P. V., Janssen, M., ... Wamba, S. F. (2023). Metaverse beyond the hype: Multidisciplinary perspectives on emerging challenges, opportunities, and agenda for research, practice and policy. *International Journal of Information*

- Management*, 71, 102642.
<https://doi.org/10.1016/j.ijinfomgt.2023.102642>
- Educause. (2023). *7 things you should know about generative AI*. <https://www.educause.edu>
- Ellis, C., van Haeringen, K., Harper, R., Bretag, T., Zucker, I., Newton, P., & Rozenberg, P. (2020). Does authentic assessment assure academic integrity? Evidence from contract cheating data. *Higher Education Research & Development*, 39(3), 454–469.
<https://doi.org/10.1080/07294360.2019.1680956>
- Harper, R., Bretag, T., & Rundle, K. (2020). Detecting contract cheating: Examining the role of assessment type. *Higher Education Research & Development*, 39(7), 1367–1381.
<https://doi.org/10.1080/07294360.2020.1724899>
- International Center for Academic Integrity. (2021). *The fundamental values of academic integrity* (3rd ed.).
<https://academicintegrity.org>
- Kasneci, E., Sessler, K., Küchemann, S., Bannert, M., Dementieva, D., Fischer, F., Gasser, U., Großschedl, J., Günnemann, S., Hüllermeier, E., Krusche, S., Narciss, S., Pinkwart, N., Salehi, S., Schmidt, A., Seidel, T., Stadler, M., & Kasneci, G. (2023). ChatGPT for good? On opportunities and challenges of large language models for education. *Learning and Individual Differences*, 103, 102274.
<https://doi.org/10.1016/j.lindif.2023.102274>
- Lancaster, T., & Cotarlan, C. (2021). Contract cheating by STEM students through a file sharing website: A COVID-19 pandemic perspective. *International Journal for Educational Integrity*, 17(3).
<https://doi.org/10.1007/s40979-021-00083-9>
- Liang, P. P., Wu, C., Morency, L.-P., & Salakhutdinov, R. (2023). GPT detectors are biased against non-native English writers. *Patterns*, 4(9), 100779.
<https://doi.org/10.1016/j.patter.2023.100779>
- Montacute, R. (2023). *Artificial intelligence and education: The implications of ChatGPT for teachers and students*. Sutton Trust.
- OECD. (2024). *AI in education: Ethical and policy implications*. Organisation for Economic Co-operation and Development.
- OpenAI. (2023). New AI classifier for indicating AI-written text. <https://openai.com>
- QAA. (2020). *Contracting to cheat in higher education: How to address contract cheating, the use of third-party services and essay mills*. Quality Assurance Agency for Higher Education.
- QAA. (2023). *Generative artificial intelligence in education: Guidance for higher education providers*. Quality Assurance Agency for Higher Education.
- Selwyn, N., Hillman, T., Eynon, R., Ferreira, G., Knox, J., Macgilchrist, F., & Sancho-Gil, J. M. (2023). What's next for Ed-Tech? Critical hopes and concerns for the 2020s. *Learning, Media and Technology*, 48(1), 1–12.
<https://doi.org/10.1080/17439884.2022.2139395>
- TEQSA. (2023). *Assessment reform for the age of artificial intelligence*. Tertiary Education Quality and Standards Agency.
- Tlili, A., Shehata, B., Adarkwah, M. A., Bozkurt, A., Hickey, D. T., Huang, R., Johnson, N., & Denden, M. (2023). What if the devil is my guardian angel: ChatGPT as a case study of using chatbots in education. *Smart Learning Environments*, 10(15).
<https://doi.org/10.1186/s40561-023-00237-x>
- UNESCO. (2023). *Guidance for generative AI in education and research*. United Nations Educational, Scientific and Cultural Organization.
- Zawacki-Richter, O., Marín, V. I., Bond, M., & Gouverneur, F. (2019). Systematic review of research on artificial intelligence applications in higher education – Where are the educators? *International Journal of Educational Technology in Higher Education*, 16(1).
<https://doi.org/10.1186/s41239-019-0171-0>

