International Journal of Education, Vocational and Social Science
https://e-journal.citakonsulfindo.or.id/index.php/IJEVSS
E-ISSN : 2963 4954
IJEVSS

Generative AI And Academic Integrity: A Critical Analysis From An Education Policy Perspective

Afifah Faradila
Universitas Negeri Yogyakarta, Indonesia

Corresponding Author: afifahfaradila.2025@student.uny.ac.id

Abstract

The penetration of generative AI in higher education has triggered an epistemological disruption. The changes has given rise to an academic integrity crisis characterized by new practices such as AI-assisted plagiarism, AI-based ghost-writing, and fabrication of data and references due to AI hallucination. The research aims to analyze the impact of generative AI on academic integrity, identify education policy dilemmas, and formulate adaptive governance models in higher education. The research uses a qualitative approach with a literature study design through critical policy analysis. Data was obtained from the policy documents, educational institution regulations, as well as scientific literature. The analysis was carried out using content analysis techniques with an interpretive-critical approach to normative changes in educational governance in the AI era. The research results show that a restrictive approach in the form of a total ban on the use of AI is no longer adequate because it is unrealistic, difficult to implement, and has the potential to widen the AI literacy gap. Academic integrity needs to be reconstructed from a prohibition-based compliance paradigm to a transparency-based accountability paradigm. In this paradigm, integrity is no longer measured using AI, but by how AI is used in a transparent, responsible and scientifically accountable manner.

Keywords: Academic Integrity, Critical Policy Analysis, Education Policy

INTRODUCTION

The development of generative artificial intelligence (AI) has brought fundamental changes to the global education ecosystem. In contrast to the previous wave of educational digitalization which focused on digitizing learning resources, learning management systems, and integrating information technology into the learning process, Generative AI presents new capabilities that

Copyright at authors some right reserved this work is licensed under a Creative Commons Attribution-Share Alike 4.0 International License

allow machines to produce text, images, program code, and natural language-based analysis independently. The presence of various platforms such as ChatGPT, Gemini, Claude, and Microsoft Copilot have changed the way students, educators, and researchers access, process, and produce knowledge in academic environments. (Kasneci, et al., 2023) (Tlili, et al., 2023)

In the educational context, Generative AI offers a variety of strategic opportunities. This technology can function as an adaptive tutor, academic writing assistant, support for personalized learning, and an instrument for increasing research productivity. A number of studies show that the use of Generative AI has the potential to increase the efficiency of the learning process, expand access to knowledge sources, and support the development of 21st century skills (Holmes & Miao, 2023) (Zawacki-Richter, Marín, Bond, & Gouverneur, 2019) (Cotton, Cotton, & ShipwaY, 2024). However, these various benefits are also accompanied by the emergence of increasingly complex ethical issues, especially regarding academic honesty, authenticity of scientific work, intellectual accountability, and the validity of knowledge produced by AI-based systems.

The rapid adoption of Generative AI in education has intensified concerns regarding academic integrity, as its ability to generate essays, reports, and assessment responses has blurred the boundary between legitimate learning support and academic misconduct, giving rise to AI-assisted plagiarism, fabricated references, and digital ghost writing practices. While educational institutions continue to struggle with developing appropriate regulatory frameworks, students and educators have already integrated AI into everyday academic activities, creating a gap between technological innovation and policy responses. Although existing research largely focuses on learning effectiveness and technology acceptance, studies examining the intersection of Generative AI, academic integrity, and educational policy remain limited. Consequently, the key challenge is not whether AI should be used in education, but how institutions can establish governance mechanisms that balance innovation with the core values of honesty, responsibility, fairness, and trust.

The development of Generative AI has significantly changed academic practices at various levels of higher education. This technology is no longer used solely as an information retrieval tool, but has become part of the learning, teaching and research process. The use of Generative AI by students, lecturers and researchers shows that the ongoing digital transformation is not only technological, but also has an impact on academic norms, culture and governance. (Holmes & Miao, 2023) (Dwivedi, et al., 2023)

Generative AI has been widely adopted across higher education by students, lecturers, and researchers. Students use AI to understand complex concepts, develop writing outlines, translate

Copyright at authors some right reserved this work is licensed under a Creative Commons Attribution-Share Alike 4.0 International License

literature, and receive feedback on academic work, thereby enhancing access to learning resources and supporting personalized learning. However, excessive reliance on AI may encourage cognitive offloading, reducing students' cognitive engagement and blurring the distinction between legitimate academic assistance and academic misconduct. At the instructional level, lecturers utilize AI to develop teaching materials, design assessments, generate question variations, and support administrative tasks. While these applications improve efficiency, they also challenge traditional assessment models, as AI-generated outputs make it increasingly difficult to evaluate students' authentic competencies. Consequently, educational practices are shifting toward process-oriented, reflective, collaborative, and contextual forms of assessment.

In research activities, Generative AI supports literature exploration, information synthesis, reference management, data analysis, and preliminary drafting of academic texts, contributing to greater research productivity. Nevertheless, concerns regarding hallucinated references, algorithmic bias, limited transparency, and threats to the originality of scientific work have intensified debates on academic ethics. These developments indicate that restrictive approaches to AI are becoming less feasible, highlighting the need for governance frameworks that promote ethical, transparent, and responsible AI use. From an educational policy perspective, this requires a transition from regulatory control toward adaptive AI governance that establishes clear guidelines for AI use, disclosure practices, academic integrity standards, and AI literacy development.

Beyond its practical applications, Generative AI is transforming the epistemological foundations of education by reshaping how knowledge is accessed, processed, and produced. Through conversational information retrieval, learners can obtain synthesized information and conceptual explanations instantly, increasing efficiency but also raising concerns regarding misinformation and AI hallucinations. As a result, information literacy and critical evaluation skills become increasingly essential. Furthermore, AI-assisted knowledge processing shifts the human role from primary information processor to evaluator and decision-maker, while excessive dependence may weaken critical thinking, problem-solving abilities, and academic resilience. The most profound transformation occurs in knowledge production, where academic writing and research are increasingly generated through human–AI collaboration, positioning AI not merely as a technical tool but as a cognitive partner within contemporary scholarly activities.

This development raises fundamental questions regarding the concepts of authorship, originality and academic responsibility. If previously academic quality was measured based on an individual's ability to produce work independently, then in the era of Generative AI the focus of

Copyright at authors some right reserved this work is licensed under a Creative Commons Attribution-Share Alike 4.0 International License

assessment shifts to the ability to manage, evaluate and verify output produced through human and machine interaction. Thus, academic competence is no longer solely determined by the ability to produce texts, but also by the ability to direct, criticize and be accountable for the knowledge production process. (Perkins, 2023) (Cotton, Cotton, & ShipwaY, 2024)

This epistemological shift shows that the main challenge of education is no longer simply providing access to information, but rather ensuring that students are able to develop the capacity to think critically, reflectively and ethically in utilizing AI. From the perspective of Educational Policy Science, these conditions require a reorientation of policy from an approach that focuses on knowledge products to an approach that assesses the knowledge construction process.

Therefore, the learning evaluation system needs to shift from assessments that only measure final results to assessments that are able to explore students' thinking, argumentation, reflection and decision-making processes. In addition, regulations are needed that regulate transparency in the use of AI, disclosure of AI contributions in academic assignments, and strengthening AI literacy as an integral part of the education curriculum. Without adaptive policy reform, educational institutions risk producing graduates who have broad access to digital information, but do not have sufficient capacity to assess the validity, meaning and implications of the knowledge they use.

The rapid development of Generative AI has introduced significant ethical challenges for higher education, particularly regarding academic integrity, which is founded on the principles of honesty, responsibility, fairness, trust, and respect. One of the central debates concerns academic originality and plagiarism, as AI-generated content does not always fit traditional definitions of plagiarism yet may obscure the extent of technological contribution when used without disclosure. This has prompted calls to reinterpret academic originality in the era of human–AI collaboration. Additional concerns involve the reliability of AI detection tools, which remain prone to false positives and false negatives, raising issues of academic fairness and limiting their suitability as sole decision-making instruments. Debates have also emerged regarding authorship, accountability, intellectual property, and the use of copyrighted training data, highlighting that the challenges of Generative AI extend beyond technology into the domains of ethics and knowledge governance.

These developments underscore that the primary challenge lies not in the existence of AI itself, but in the absence of effective governance frameworks capable of ensuring its responsible use. The rapid adoption of AI has outpaced policy development, resulting in inconsistent institutional approaches and normative uncertainty regarding acceptable use, disclosure

Copyright at authors some right reserved this work is licensed under a Creative Commons Attribution-Share Alike 4.0 International License

requirements, and assessment standards. Furthermore, concerns regarding AI hallucinations, undisclosed AI-assisted writing, and cognitive offloading emphasize the need for critical verification and appropriate pedagogical design. From an educational policy perspective, these conditions reflect a policy lag that necessitates the development of adaptive AI governance frameworks grounded in transparency, accountability, fairness, and academic responsibility. Such frameworks should regulate AI disclosure practices, define acceptable levels of AI involvement in learning and research, establish adaptive evaluation standards, and protect academic rights, thereby balancing technological innovation with the preservation of academic integrity.

Recent literature on artificial intelligence (AI) in education highlights three major trends: (1) the pedagogical effectiveness of AI in adaptive learning, automated assessment, and digital learning support; (2) stakeholders' perceptions, acceptance, and readiness toward generative AI, including AI literacy, technology acceptance, and user anxiety; and (3) emerging studies on AI ethics, focusing on algorithmic bias, transparency, and academic integrity.

Despite these developments, several gaps remain. First, most studies emphasize individual-level risks, particularly students' use of AI, while connections to educational policy and institutional governance remain underexplored. Second, although AI governance has gained attention internationally, its application within educational institutions is not yet systematically integrated, with limited research on translating governance principles into operational policies and academic integrity mechanisms. Third, the literature lacks adaptive governance frameworks that balance risk mitigation with educational innovation, as existing approaches often lean toward either restrictive regulation or unrestricted adoption.

Addressing these gaps, this study integrates perspectives on AI, academic ethics, and institutional governance to propose a conceptual framework for adaptive AI governance in education. The framework positions educational institutions as key actors in balancing academic integrity, pedagogical innovation, and digital transformation.

Based on the background, problem formulation, and research gaps that have been previously identified, this qualitative research aims to analyze the implications of using generative artificial intelligence on academic integrity, especially in the context of changes in scientific writing practices, shifts in the concept of originality, as well as risks arising from the use of artificial intelligence without adequate disclosure (undisclosed AI-assisted academic writing), including the potential for inaccurate information produced by generative artificial intelligence systems. In addition, this research aims to identify ethical dilemmas in the use of generative artificial intelligence in the academic environment, which includes tensions between technological

Copyright at authors some right reserved this work is licensed under a Creative Commons Attribution-Share Alike 4.0 International License

efficiency and the principles of academic honesty, as well as differences in perceptions between students, lecturers and institutions in responding to the use of this technology in the learning and academic evaluation process. Furthermore, this research also aims to analyze the implications of educational policy for the governance of the use of generative artificial intelligence at the institutional level, by examining various developing approaches, such as restrictive and adaptive approaches, as well as formulating principles that can support the development of a governance framework that is more contextual, transparent and oriented towards strengthening academic integrity as well as pedagogical innovation.

This research uses a qualitative approach with a conceptual-systematic study design that combines literature study and critical policy analysis. This approach is used to examine in depth various scientific literature and policy documents related to generative artificial intelligence in education. Methodologically, this research follows a structured qualitative synthesis approach which emphasizes a systematic process in collecting, sorting and analysing literature to produce conceptual conclusions that can be scientifically justified . In addition, this research uses a critical analysis approach that not only describes existing policies, but also examines how a problem is represented in policy along with its social and ethical implications. (Bacchi, 2025) (Fischer, Torgerson, Durnová, & Orsini, 2015)

The data in this research is entirely sourced from secondary data consisting of policy documents and reputable scientific literature. The main data includes official policy documents and guidelines published by international institutions, such as UNESCO (2023) in its Guidelines for Generative Artificial Intelligence in Education and Research, as well as the OECD (2019; 2024) which discusses the principles and policy direction of artificial intelligence in education. In addition, institutional policy documents such as academic ethics guidelines, rector's decrees, and rules for the use of artificial intelligence in higher education are also used as part of the main data. Secondary data consists of scientific articles obtained from the Scopus, Web of Science, and SINTA 1–2 databases that are relevant to the themes of artificial intelligence in education, academic integrity, and higher education policy. The literature used is limited to articles that have gone through a peer review process and were published within the 2015–2025 time frame to remain relevant to the latest developments.

Data collection was carried out through documentation methods with systematic literature search techniques on various scientific databases, such as Scopus, Web of Science, SINTA, as well as official websites of international institutions. The search process was carried out using keywords such as "generative artificial intelligence", "academic integrity", "education

Copyright at authors some right reserved this work is licensed under a Creative Commons Attribution-Share Alike 4.0 International License

policy", "higher education governance", and "learning assessment". This search technique follows systematic search principles that emphasize traceability, thoroughness, and consistency in literature selection (Xiao & Watson, 2019). Literature selection criteria include: (1) scientific articles that have gone through a peer review process, (2) published in the 2015–2025 time period, and (3) have a direct connection to the topic of artificial intelligence in education and educational policy. Literature that was not relevant or did not meet scientific standards was excluded from the analysis.

Data analysis was carried out qualitatively through two complementary techniques, namely content analysis and critical policy analysis. Content analysis is carried out to systematically identify, group and interpret data from policy documents and scientific literature. This technique refers to Krippendorff and Elo & Kyngäs, which includes initial coding stages, grouping themes, and drawing conceptual conclusions. This analysis is used to map forms of artificial intelligence-based academic integrity violations as well as patterns of policy responses that develop in various educational institutions. (Krippendorff, 2018) (Elo & Kyngäs, 2008)

Next, critical policy analysis is used to examine the ideological, structural, and normative dimensions of educational policies related to artificial intelligence. This approach refers to Bacchi (2009) through the What's the Problem Represented to be approach, as well as Fischer (2015) in the tradition of interpretive policy analysis. This analysis focuses on the way problems are represented in policy, power relations in educational settings, and their impact on lecturers and students as the main actors in education. The results of the two analysis techniques are then synthesized to build a conceptual framework for artificial intelligence governance in education that balances aspects of academic integrity, learning innovation, and fair access to technology.

OVERVIEW OF GENERATIVE AI

Generative AI is a branch of artificial intelligence that focuses on the ability of systems to generate new content such as text, images, code, audio and video based on patterns learned from large-scale data. According to UNESCO (2023), Generative AI is different from traditional AI systems that are analytical or predictive in nature, because of its ability to produce new output that is contextual and resembles human expression.

Technically, Generative AI is based on large-scale neural networks models such as transformer architecture, which are trained using a self-supervised learning approach and strengthened through alignment techniques such as Reinforcement Learning from Human Feedback (RLHF). In practice, these systems do not have semantic understanding like humans, but

Copyright at authors some right reserved this work is licensed under a Creative Commons Attribution-Share Alike 4.0 International License

instead generate responses based on statistical representations of very large training data. (Kasneci, et al., 2023)

Based on Kasneci et al. (2023), Generative AI has a number of key characteristics that differentiate it from previous educational technologies:

- Advanced natural language processing capabilities, which enable the system to understand context-based instructions (prompts) and produce coherent responses in a variety of communication styles.
- Context-based adaptive capabilities, namely the ability to adjust responses based on conversation history and user instructions within a single interaction session.
- Multimodal support, which allows integration of multiple forms of data such as text, images, and code in one generative system.
- Epistemic limitations, namely reliance on statistical patterns of training data which can lead to inaccurate information or hallucinated outputs, thus requiring human verification in an academic context.

OECD (2024) identified that the use of Generative AI in education is growing in several main domains, namely learning, research, academic writing and assessment.

In a learning context, Generative AI has the potential to function as a learning aid that supports independent learning and concept exploration, although its pedagogical effectiveness is highly dependent on the learning design and digital literacy of the user. In the realm of research, AI can be used as a tool in initial literature synthesis, idea exploration, and data analysis support. However, its use still requires critical validation of the source and results produced.

In academic writing, Generative AI can help increase efficiency in technical aspects such as language editing, text structure, and improving writing style. However, issues related to transparency in the use of AI (AI disclosure) are an important concern in maintaining academic integrity. In the context of assessment, AI is starting to be used as a tool in the development of initial evaluation and feedback instruments. However, the OECD (2024) emphasizes that the use of AI in assessments must remain within the framework of pedagogical oversight to maintain the validity and reliability of evaluations.

OVERVIEW OF ACADEMIC INTEGRITY

According to the idea of constructivism, people learn through observation and research. By engaging in life experiences and observing the results, people construct their understanding and knowledge of the world. New information must be compared to our prior knowledge and

Copyright at authors some right reserved this work is licensed under a Creative Commons Attribution-Share Alike 4.0 International License

opinions, which may demand us to alter our belief systems or reject and consider the new information as insignificant. In any case, we are active participants in the creation of our knowledge and experience. Ask questions, investigate and evaluate what we already know to achieve this. The constructive approach of classroom learning can refer to a variety of different teaching practices (Bereiter, 1994; Tam, 2000). According to Driscoll (2005), the constructivist model suggests that understanding can only occur within the human brain. It does not have to correspond to any fact. Students would be constantly attempting to construct their possible mental aspect of the entire globe based on their perceptions of it. Students will keep updating their thought processes to represent new knowledge as they recognize each novel experience, and will thus build their perception of reality (Spector, 2001). The second idea is that learning is an active process instead of a passive one. Students must encounter their comprehension in view about what they experience in the new setting. If what students experience contradicts their existing knowledge, their insight can shift to match the new information. Students stay operational during this phase, applying existing knowledge, noting essential aspects in novel situations, judging the coherence of previous and arising understanding, and modifying required knowledge based on that judgment (Gopnik & Wellman, 2012; Moons & De Backer, 2013; Mvududu, 2005; Naylor & Keogh, 1999; Phillips, 1995).

EDUCATIONAL POLICY SCIENCE PERSPECTIVE

In public policy studies, education policy functions not only as a set of formal regulations but also as a governance instrument that shapes interactions among actors, institutions, and technologies within the education system. In the era of Generative AI, education policy plays a strategic role in balancing technological innovation with academic values. Conceptually, it serves three key functions: as a regulatory instrument that establishes standards, boundaries, and compliance mechanisms to mitigate risks such as AI-assisted academic misconduct; as a governance mechanism that coordinates roles, responsibilities, and accountability among stakeholders; and as a socio-technical adaptation mechanism that bridges the gap between rapid technological change and the institutional and ethical readiness of educational communities. Consistent with UNESCO (2023) and OECD (2019; 2024), effective AI governance should be adaptive, principle-based, and responsive to emerging challenges.

This study draws on three complementary policy perspectives. First, Policy Cycle Theory (Lasswell) explains policy development through the stages of agenda setting, formulation, adoption, implementation, and evaluation, highlighting the policy lag that often occurs when

Copyright at authors some right reserved this work is licensed under a Creative Commons Attribution-Share Alike 4.0 International License

regulatory responses fail to keep pace with technological innovation. Second, the Multiple Streams Framework (Kingdon) suggests that policy change emerges when problem, policy, and political streams converge. In the context of Generative AI, concerns regarding academic integrity and assessment validity have increased, and various policy alternatives have been proposed, yet comprehensive and coordinated policy responses remain fragmented. Third, Governance Theory, particularly Ostrom's concept of polycentric governance and Ansell and Gash's collaborative governance model, emphasizes that AI governance cannot be managed by a single institution. Instead, it requires collaboration among governments, educational institutions, technology developers, and civil society. Consequently, education policy in the AI era is evolving from a hierarchical regulatory model toward an adaptive, participatory, and multi-stakeholder governance framework that promotes transparency, accountability, and ethical AI use.

DISCUSSION

The emergence of Generative AI has transformed the epistemological landscape of higher education, shifting knowledge production from a traditional knowledge creation paradigm toward a more hybrid knowledge generation model. Rather than replacing human cognition, AI redistributes cognitive functions through interactions between humans and large language models, where users guide the process and AI generates content based on learned language patterns. This transformation accelerates knowledge production and alters the role of humans from primary knowledge producers to directors and evaluators of technology-assisted outputs within the digital learning ecosystem.

At the same time, digital transformation is reshaping students' academic behaviors, particularly in information seeking and scientific writing. AI-powered conversational search enables rapid summarization, categorization, and synthesis of information, increasing learning efficiency but also creating risks of cognitive offloading and reduced critical engagement. Furthermore, the ability of AI to generate coherent academic texts challenges the validity of product-based assessments, as the quality of written outputs may no longer accurately reflect students' conceptual understanding. Consequently, higher education institutions need to shift assessment practices toward evaluating learning processes, transparency in AI use, critical literacy, and reflective thinking to ensure the authentic measurement of academic competence.

According to the Big Indonesian Dictionary (KBBI), plagiarism is "plagiarism that violates copyright" or taking other people's work, opinions, or ideas without citing a valid source (Language Agency, Ministry of Education and Culture). Its core lies in the recognition of intellectual

Copyright at authors some right reserved this work is licensed under a Creative Commons Attribution-Share Alike 4.0 International License

ownership. However, the emergence of generative AI is shifting this practice, as Large Language Models (LLMs) expand academic integrity violations from direct text copying to automated reproduction of ideas through generative systems (Kasneci et al., 2023; Pudasaini et al., 2024).

This development transforms plagiarism from textual copying into AI-assisted or generative plagiarism. While traditional plagiarism can be detected through lexical similarity, AI can generate linguistically original text while preserving underlying idea structures through large-scale semantic paraphrasing (Bender et al., 2021; Floridi & Chiriatti, 2020). As a result, plagiarism shifts from copying to idea appropriation that is difficult to detect using similarity-based tools, further complicated by blurred boundaries between human and AI authorship (Dwivedi et al., 2023; Utami et al., 2025). This reflects regulatory lag in academic institutions and reduces the relevance of product-based assessments (Cotton et al., 2024).

In another dimension, ghostwriting is defined as "writing by another party on behalf of someone who did not write it," representing intellectual dishonesty due to a mismatch between listed and actual authors. In higher education, this has evolved into AI-based ghostwriting, where the entire writing process is delegated to AI systems without significant cognitive involvement from students (Cotton et al., 2024; Perkins, 2023).

LLMs such as ChatGPT, Claude, and Gemini have made academic writing outsourcing nearly effortless through simple prompting (Kasneci et al., 2023; Dwivedi et al., 2023). Although outputs appear unique, they are probabilistically generated and often preserve latent argument patterns, creating an illusion of originality and academic ability (Bender et al., 2021). Meanwhile, AI detection tools based on perplexity and burstiness remain unreliable, producing false positives and negatives, especially for paraphrased texts (Liang et al., 2023). This demonstrates that AI-based ghostwriting not only reduces students' cognitive engagement but also creates a detection crisis in conventional academic integrity systems.

Digital ghost-writing creates an epistemological crisis in higher education because it removes students' intellectual agency in the learning process. From a pedagogical perspective, writing is not only an end product, but also a cognitive process involving analysis, synthesis, and reflection. When this process is transferred to machines, the function of education as knowledge formation and cognitive development becomes distorted (UNESCO, 2023). The main implications that emerge are:

a. Dissociation between Products and Competencies

Students can produce high-quality work without having a substantial understanding of the content of the writing.

Copyright at authors some right reserved this work is licensed under a Creative Commons Attribution-Share Alike 4.0 International License

b. Academic Evaluation Validity Crisis

Evaluation systems based on essays, reports and theses are becoming less valid in measuring students' actual abilities. (Kasneci, et al., 2023)

c. Erosion of the Concept of "Academic Writer"

The concept of the writer as an intellectual subject has shifted to being simply an AI system operator, which has an impact on the redefinition of academic literacy in the digital era.

In the Big Indonesian Dictionary (KBBI), hallucinations are defined as "perceptions or responses without any sensory stimulation." In generative AI, this concept refers to outputs that appear convincing but are not grounded in reality or empirical verification. In the literature, this is known as AI hallucination, where large language models produce false, inaccurate, or unverified information (Ji et al., 2023). Unlike plagiarism and ghostwriting, which concern authorship and intellectual ownership, hallucinations and fabrications challenge the epistemological validity of scientific knowledge itself. In academic contexts, they often appear as fabricated but structurally plausible references. LLMs can generate complete citations (author, year, journal, DOI) that do not exist in databases such as Scopus or Google Scholar because they operate through pattern prediction rather than source retrieval (Ji et al., 2023), increasing risks of false references, higher verification burdens, and weakened academic accountability.

Generative AI can also produce fictitious empirical data that appears statistically and narratively valid, including survey results, regression outputs, or even entirely imaginary interview transcripts (Bender et al., 2021; Dwivedi et al., 2023). Although internally coherent, such data lacks empirical grounding, creating a "synthetic plausibility" problem where information seems correct but cannot be verified in reality (Ji et al., 2023). These hallucinations and fabrications contribute to an epistemic crisis in knowledge production systems. First, scientific trust declines as the authenticity of data and references becomes uncertain, challenging universities as verification institutions (UNESCO, 2023). Second, peer review systems face a verification gap because they are not designed to systematically validate all references and datasets (Kasneci et al., 2023). Third, there is a risk of hallucination citation cascades, where AI-generated errors spread through chains of citation and form pseudo-scientific literature (Rawte et al., 2023).

From the perspective of Educational Policy Science, the adoption of generative artificial intelligence is not neutral, as it produces structural consequences in the distribution of learning opportunities. Literature suggests that digital innovation often reproduces inequality in new forms (Dijk, 2020). In this context, inequality is no longer limited to access to technology, but extends to

Copyright at authors some right reserved this work is licensed under a Creative Commons Attribution-Share Alike 4.0 International License

differences in capability and quality of use, referred to as The New Digital Inequality (Ragnedda & Muschert, 2013). A first dimension of inequality emerges from unequal access to free and paid AI services. Premium models typically offer greater context capacity, higher accuracy, and stronger reasoning abilities compared to free versions (Kasneci et al., 2023), creating a situation where economically advantaged students benefit from more powerful academic augmentation tools.

A second dimension is the AI literacy gap, referring to differences in the ability to use AI effectively and ethically. This includes understanding system mechanisms, prompt design, critical evaluation skills, and the ability to identify bias and errors (Long & Magerko, 2020; Ng et al., 2021). In practice, this gap appears in uneven prompting skills, low critical verification capacity, and increasing cognitive dependency on automated systems (Selwyn, 2019), along with intergenerational differences between students and lecturers in pedagogical AI use. From a policy perspective, this emerging inequality indicates that generative AI may reinforce social stratification through unequal access to educational enhancement tools (Ragnedda & Muschert, 2013). It also raises concerns about assessment fairness, since product-based evaluation systems often ignore disparities in access and AI literacy. Therefore, policy interventions are needed that emphasize inclusive access, stronger digital literacy development, transparency in AI use, and process-based assessment models as recommended by UNESCO (2023).

The unpreparedness of educational regulations in responding to generative AI disruption has created complex policy dilemmas in higher education institutions. From the perspective of policy process theory, particularly Kingdon's Multiple Streams Framework (1995) and Lasswell's policy cycle (1956), the emergence of generative AI reveals an asymmetry between problem identification, policy alternatives, and political dynamics. This mismatch has produced fragmented institutional responses that can generally be categorized into three approaches: restrictive, permissive, and adaptive. Each reflects different assumptions about technology, academic integrity, and institutional control.

The restrictive approach treats generative AI as a direct threat to academic integrity and therefore seeks to prohibit its use entirely in learning and assessment. Several institutions initially adopted bans on tools such as ChatGPT as a precaution against academic misconduct (Cotton et al., 2024). Theoretically, this aligns with academic integrity protectionism, which emphasizes strict control over external inputs in knowledge production (Bretag, 2016). While this approach strengthens control over plagiarism and ghostwriting and maintains clear ethical boundaries, it is difficult to enforce in open digital environments and often becomes unenforceable in practice. It may also encourage covert AI use, undermine trust between lecturers and students, and reduce

Copyright at authors some right reserved this work is licensed under a Creative Commons Attribution-Share Alike 4.0 International License

students' preparedness for AI-integrated futures.

In contrast, the permissive approach allows unrestricted use of generative AI, assuming that technology is neutral and can be guided by individual ethics. Rooted in technological optimism (Feenberg, 1991), this approach promotes pedagogical innovation, rapid adoption, and academic experimentation. However, without regulatory frameworks, it risks moral hazard, over-reliance on automated systems, and declining validity of academic competence (Kasneci et al., 2023; Dwivedi et al., 2023). This creates a regulatory vacuum that enables misuse such as AI-driven ghostwriting and data fabrication, making academic integrity difficult to sustain systematically.

Between these extremes, the adaptive approach offers a more balanced alternative through agile governance, emphasizing flexibility, responsiveness, and contextual regulation (Taeihagh, 2021; Kuhlmann & Rip, 2018). Rather than prohibiting AI use, it regulates transparency through mechanisms such as disclosure of AI assistance, prompts, and usage documentation. UNESCO (2023) recommends this model as a key strategy in educational AI governance because it balances innovation with integrity. However, it requires strong institutional capacity, continuous regulatory adaptation, and a shift from product-based to process-based assessment systems.

Overall, these three approaches demonstrate that the central policy dilemma is no longer whether AI should be used, but how human–machine relations in knowledge production should be governed. This signals a shift in educational policy from technological control to ethical governance. Reconstructing academic integrity in the AI era therefore requires moving from prohibition-based regulation to transparency-based governance. This shift reframes the key question from whether students use AI to how AI is used transparently, responsibly, and verifiably in academic work. It is grounded in three pillars: radical transparency through AI disclosure, responsible use that preserves human cognitive agency, and scientific accountability that ensures full responsibility for AI-assisted outputs. These principles respond to challenges such as AI hallucinations and reinforce the need for verification, process-based assessment, and human oversight in academic evaluation (Ji et al., 2023; UNESCO, 2023).

As a conceptual synthesis of the previous analyses, this section articulates an adaptive education policy model for generative artificial intelligence within an agile governance framework. The model is not merely a technical response to technological disruption, but a systemic reconstruction of academic integrity governance in the era of knowledge automation. In line with UNESCO (2023) and OECD (2021), this approach emphasizes not only regulatory control, but also human capacity building, process transparency, and pedagogical transformation. The first pillar is the rule-making framework, which provides normative clarity by classifying AI use into permitted

Copyright at authors some right reserved this work is licensed under a Creative Commons Attribution-Share Alike 4.0 International License

(e.g., proofreading), limited (e.g., supporting analysis or argument structuring), and prohibited uses (e.g., full ghostwriting and data fabrication), thereby reducing ambiguity and subjectivity in academic judgment (Cotton et al., 2024; Bretag, 2016). The second pillar is AI literacy cultivation, which strengthens the ability of users to understand system limitations, verify outputs, and design ethical prompts, while addressing intergenerational literacy gaps between students and lecturers (Long & Magerko, 2020; Ng et al., 2021; Selwyn, 2019). The third pillar is transparency and disclosure, requiring mandatory AI disclosure statements that document platforms used, purposes, and levels of AI contribution, including prompt logs to ensure process-based accountability (Floridi & Chiriatti, 2020).

The fourth pillar is assessment redesign, which responds to the limitations of product-based evaluation in the context of AI automation by emphasizing authentic and process-oriented assessments such as oral examinations, case presentations, research logbooks, and project-based learning (Kasneci et al., 2023; Perkins, 2023; Wiggins, 1998). The fifth pillar is the strengthening of academic ethics through internalized integrity culture, including ICAI values, lecturer exemplarity, and the development of an academic ecosystem centered on intellectual processes rather than outputs (Bretag, 2016; Macfarlane et al., 2014). This aligns with Biesta's (2010) view of education as a subjectification process that forms critical and responsible individuals. Overall, this five-pillar model demonstrates that AI governance in education extends beyond technical regulation and instead represents a comprehensive transformation of higher education systems, integrating regulation, literacy, transparency, assessment reform, and ethical culture into an adaptive and inclusive framework for sustaining academic integrity.

CONCLUSION

Based on a conceptual-critical analysis of the dynamics of generative AI in the higher education ecosystem, this research concludes that the technological transformation of artificial intelligence is not merely instrumental, but also epistemological and normative, as it reshapes how knowledge is produced, validated, and evaluated within academic systems. First, generative AI has fundamentally altered the paradigm of academic integrity by shifting knowledge production from a human agent-based model to a probabilistic, machine-mediated system. This shift has generated new academic pathologies, including AI-assisted plagiarism in semantic form, AI-based ghostwriting that reduces human cognitive agency, and hallucination or fabrication phenomena that produce unverifiable references and scientific data, indicating that academic integrity issues now extend beyond textual misconduct into deeper epistemic concerns. Second, a total

Copyright at authors some right reserved this work is licensed under a Creative Commons Attribution-Share Alike 4.0 International License

prohibition or restrictive approach is no longer sufficient in contemporary education governance, as banning AI use is largely impractical within an open digital ecosystem, risks widening the AI literacy gap, and may hinder students' technological competencies, making such approaches reactive rather than adaptive to structural changes in higher education. Third, academic integrity must be reconstructed from a compliance-based paradigm to a transparency- and accountability-based paradigm, where integrity is defined not by the absence of AI use, but by the capacity to document, explain, and justify how AI is used in knowledge production, grounded in principles of full disclosure, process auditability, and intellectual responsibility that shift integrity from prohibition to ethical transparency and process verification. Fourth, educational policy transformation must move toward an adaptive, inclusive AI governance model grounded in critical digital literacy, repositioning educational institutions not merely as regulators but as enablers of technology-integrated learning ecosystems; this is operationalized through five interconnected pillars, namely clear AI usage regulation, strengthening AI literacy among students and lecturers, implementing AI disclosure mechanisms, redesigning assessment toward process-based evaluation, and reinforcing an internalized and sustainable academic ethics culture, ultimately emphasizing that success in the AI era depends not only on technological control, but on the ability of institutions to build adaptive and reflective integrity ecosystems.

REFERENCES

Ansell, C., & Gash, A. (2008). Collaborative Governance in Theory and Practice. Journal of Public Administration Research and Theory.

Ashish, V. (2017). Attention is all you need. dvances in neural information processing systems 30.

Bacchi, C. (2025). What's the Problem Represented to Be? A New Thinking Paradigm. London: Routledge.

Bender, E., Gebru, T., Major, A. M., & Shmitchell, S. (2021). On the dangers of stochastic parrots: Can language models be too big? ACM conference on fairness, accountability, and transparency.

Bretag, T. (2016). Handbook of academic integrity.

Chan, C. K. (2023). Is AI changing the rules of academic misconduct? An in-depth look at students' perceptions of'AI-giarism. airxiv.

Cotton, D. R., Cotton, P. A., & ShipwaY, J. R. (2024). Chatting and cheating: Ensuring academic integrity in the era of ChatGPT. Innovations in Education and Teaching International .

Dijk, J. V. (2020). The digital divide. John Wiley & Sons.

Copyright at authors some right reserved this work is licensed under a Creative Commons Attribution-Share Alike 4.0 International License

Dwivedi, Y. K., Kshetri, N., Hughes, L., Slade, E. L., Jeyaraj, A., Karf, A. K., . . . Albashrawi, M. A. (2023). So what if ChatGPT wrote it?" Multidisciplinary perspectives on opportunities, challenges and implications of generative conversational AI for research, practice and policy. International Journal of Information Management.

Elo, S., & Kyngäs, H. (2008). The qualitative content analysis process. Journal of advanced nursing, 107-115.

Fawns, T., Henderson, M., Matthews, K., Oberg, G., Liang, Y., Walton, J., . . . McLean, J. (2024). Gen AI and student perspectives of use and ambiguity. Proceedings of Australasian Society for Computers in Learning in Tertiary Education (pp. 132-134). Society for Computers in Learning in Tertiary Education.

Feenberg, A. (1991). Critical theory of technology. New York: Oxford University Press.

Fischer, F., Torgerson, D., Durnová, A., & Orsini, M. (2015). Introduction to critical policy studies. Edward Elgar Publishing.

Floridi, L., & Chiriatti., M. (2020). GPT-3: Its nature, scope, limits, and consequences. Minds and machines.

Holmes, W., & Miao, F. (2023). Guidance for generative AI in education and research. UNESCO.

Ji, Z., Lee, N., Frieske, R., Yu, T., Su, D., Xu, Y., . . . Fung, P. (2023). Survey of hallucination in natural language generation. ACM computing surveys.

Kasneci, E., Sessler, K., Küchemann, S., Bannert, M., Dementieva, D., Fischer, F., . . . Michaeli, T. (2023). ChatGPT for good? On opportunities and challenges of large language models for education. Learning and individual differences.

Kingdon, J. W., & Stano, E. (1984). Agendas, Alternatives, And Public Policies. Boston: Little, Brown.

Krippendorff, K. (2018). Content analysis: An introduction to its methodology. Sage publications.

Kuhlmann, S., & Rip, A. (2018). Next-generation innovation policy and grand challenges. Science and public policy, 448-454.

Liang, W., Yuksekgonul, M., Mao, Y., Wu, E., & Zou, J. (2023). GPT detectors are biased against non-native English writers. Patterns.

Long, D., & Magerko, B. (2020). What is AI literacy? Competencies and design considerations. Proceedings of the 2020 CHI conference on human factors in computing systems.

Luckin, R., & Holmes, W. (2016). Intelligence unleashed: An argument for AI in education.

Macfarlane, B., Zhang, J., & Pun, A. (2014). Academic integrity: A review of the literature. Studies in higher education.

Ng, D., Leung, J., Chu, K., & Qiao, M. (2021). AI literacy: Definition, teaching, evaluation and ethical

Copyright at authors some right reserved this work is licensed under a Creative Commons Attribution-Share Alike 4.0 International License

issues. Proceedings of the association for information science and technology.

OECD. (2019). OECD principles on artificial intelligence. OECD.

Ostrom, E. (2010). Beyond Markets and States: Polycentric Governance of Complex Economic Systems . American Economic Review, 641–72.

Papaioannou, D., Sutton, A., & Booth, A. (2016). Systematic approaches to a successful literature review.

Perkins, M. (2023). Academic integrity considerations of AI large language models in the post-pandemic era: ChatGPT and beyond. Journal of University Teaching and Learning Practice.

Pudasaini, S., Miralles-Pechuán, L., Lillis, D., & Salvador, M. (2024). Survey on plagiarism detection in large language models: The impact of chatgpt and gemini on academic integrity. Arxiv.

Ragnedda, M., & Muschert, G. W. (2013). The digital divide. Florence: KY: Routledge.

Selwyn, N. (2019). Should robots replace teachers?: AI and the future of education. John Wiley & Sons.

Taeihagh, A. (2021). Governance of artificial intelligence. Policy and society 40.2, 137-157.

Tlili, A., Huang, R., Mustafa, M. Y., Zhao, J., Bozkurt, A., Xu, L., . . . Burgos, D. (2023). Speaking of transparency: Are all artificial intelligence (AI) literature reviews in education transparent? Journal of Applied Learning & Teaching.

Tlili, A., Shehata, B., Adarkwah, M. A., Bozkurt, A., Hickey, D. T., Huang, R., & Agyemang, B. (2023). What if the devil is my guardian angel: ChatGPT as a case study of using chatbots in education. Smart Learning Environments.

Utami, P. N., Alhamid, I. N., Fadhilah, N. P., Tabitha, A. H., & Aini, A. Q. (2025). Dinamika Perubahan Bentuk Plagiarisme di Kalangan Mahasiswa DKI Jakarta: Studi Kasus Enam Perguruan Tinggi. Risenologi.

Wiggins, G. (1998). Educative Assessment. Designing Assessments To Inform and Improve Student Performance. San Francisco: Jossey-Bass Publishers.

Williamson, B., & Eynon, R. (2020). Historical threads, missing links, and future directions in AI in education. Learning, Media and Technology.

Xiao, Y., & Watson, M. (2019). Guidance on conducting a systematic literature review. Journal of planning education and research, 93-112.

Zawacki-Richter, O., Marín, V., 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.

Copyright at authors some right reserved this work is licensed under a Creative Commons Attribution-Share Alike 4.0 International License