This paper was presented at the 23rd International Conference on Sustainable Development (ICSD 2026), held at the Imperial Hotel and Ba Ria – Vung Tau University in Vung Tau, Vietnam, June 3-4, 2026.

https://doi.org/10.64211/oidaijsd190906

Responsible Generative AI Adoption in Vietnamese Higher Education: A Conceptual Framework Aligned with SDG 4

Nga Thi Thuy Duong [1], Tri Thanh Tu [2], Lien Kim Dang *[3], Toan Chau Phuong [4]
[1,2,3,4] Ho Chi Minh City University of Natural Resources and Environment, Vietnam.
*Corresponding authour: dklien@hcmunre.edu.vn

OIDA International Journal of Sustainable Development, Ontario International Development Agency, Canada.
ISSN 1923-6654 (print) ISSN 1923-6662 (online) www.oidaijsd.com
Also available at https://www.ssrn.com/index.cfm/en/oida-intl-journal-sustainable-dev/

Abstract: The rapid diffusion of generative artificial intelligence (GenAI) has created new opportunities and governance challenges for higher education, particularly in relation to academic integrity, assessment validity, data protection, equity, and AI literacy. In Vietnam, national strategies emphasize artificial intelligence and digital transformation, yet institutional-level guidance for the responsible use of GenAI in teaching, learning, and assessment remains fragmented. This paper develops a conceptual policy framework for responsible GenAI adoption in Vietnamese higher education. Using a conceptual policy analysis and framework development approach, the study synthesizes academic literature, international AI governance guidance, Vietnamese policy documents, and SDG 4 targets. The proposed GenAI Responsible Adoption Framework for Vietnamese Higher Education (GRAV-HE) comprises four interdependent dimensions: ethical policy and AI disclosure, data governance and secure infrastructure, pedagogical transformation, and comprehensive AI literacy. Rather than claiming to replace existing global guidance, GRAV-HE operationalizes Responsible AI and SDG 4 principles for institutional governance in a developing-country higher education context. The framework offers a practical basis for policy dialogue and future empirical validation in Vietnamese and comparable higher education systems.

Keywords: AI governance, Generative AI, Higher education, SDG 4, Vietnam.

Introduction

Generative artificial intelligence (GenAI) has rapidly become a disruptive force in contemporary higher education. Unlike earlier educational technologies that mainly supported communication, content delivery, or information access, GenAI systems can produce text, code, summaries, explanations, feedback, images, and other academic outputs through natural language interaction. Students and faculty now use tools such as ChatGPT, Microsoft Copilot, Google Gemini, Claude, and other GenAI applications for writing support, idea generation, programming assistance, teaching preparation, feedback, and research-related tasks. As a result, GenAI has shifted from being a peripheral digital tool to becoming a direct participant in the production and evaluation of academic work.

The educational potential of GenAI is substantial, but its benefits are not automatic. Large language models may support explanation, reflection, formative feedback, and personalized learning when they are used critically and pedagogically [1]. However, GenAI outputs may contain factual inaccuracies, fabricated references, biased assumptions, and misleading explanations, while students may over-rely on AI-generated content without sufficient verification [1, 2]. UNESCO's guidance therefore emphasizes that GenAI in education requires appropriate regulation, institutional policies, and human capacity development rather than unrestricted adoption [2]. For higher education institutions, the central question is therefore not simply whether GenAI should be allowed, but how it should be governed.

GenAI creates governance challenges across academic integrity, assessment validity, data protection, equity, and institutional accountability. It blurs the boundary between legitimate learning assistance and inappropriate academic

outsourcing, while also weakening the assumption that a polished final submission reliably reflects students' independent reasoning, writing process, or subject understanding. Recent studies suggest that universities need to move beyond plagiarism detection toward clearer AI-use disclosure, assessment redesign, and AI literacy development [3, 4]. Beyond academic integrity, students and faculty may also enter personal information, unpublished research, examination materials, student work, or institutional documents into commercial AI platforms without fully understanding how such data may be stored, processed, or reused. At the same time, unequal access to reliable internet, advanced devices, English-language resources, paid AI subscriptions, and AI literacy may create new forms of educational inequality.

Vietnam provides an important context for examining these issues because the country is expanding digital transformation while still facing uneven readiness across regions, institutions, and learner groups. At the beginning of 2025, Vietnam had approximately 79.8 million internet users, equivalent to 78.8% of the population, but about 21.5 million people remained offline; 59.5% of the population lived in rural areas [5]. Vietnam's higher education system is also large and diverse. In 2024, the country had 243 universities, including 176 public and 67 private institutions, with 2,355,711 students and 88,031 full-time lecturers [6]. These figures show that GenAI governance is not a marginal issue affecting only technology-oriented universities, but a system-level concern involving millions of learners and tens of thousands of lecturers.

At the policy level, Vietnam has positioned artificial intelligence and education digital transformation as national priorities. Decision No. 127/QD-TTg approved the national strategy on research, development, and application of artificial intelligence through 2030, while Decision No. 131/QD-TTg promotes information technology application and digital transformation in education and training for the period 2022–2025, with orientation toward 2030 [7, 8]. The Ministry of Education and Training has also developed the Higher Education Management Information System (HEMIS), which has digitized data from 442 training establishments, more than 152,000 lecturers, and more than 2.1 million students [9]. These developments show that Vietnam has an active digital transformation agenda. However, national digital policy does not automatically provide institutional guidance for GenAI use in course syllabi, assessment, AI-use disclosure, data protection, faculty training, or student AI literacy.

This study frames responsible GenAI adoption through Sustainable Development Goal 4 (SDG 4), which emphasizes inclusive and equitable quality education and lifelong learning opportunities [11]. In Vietnamese higher education, this lens is important because GenAI adoption should be evaluated not only by its potential to improve learning efficiency, but also by whether it supports equity, protects student data, strengthens teacher capacity, and preserves meaningful assessment. If GenAI adoption is left to individual students, instructors, or institutions without shared governance, its benefits may be distributed unevenly and may intensify existing educational inequalities.

Existing global frameworks, including UNESCO guidance and the OECD AI Principles, provide important foundations for responsible AI by emphasizing transparency, accountability, fairness, privacy, safety, human oversight, institutional policy, and human capacity development [2, 12]. Nevertheless, a gap remains between broad global AI governance principles and institution-level implementation in Vietnamese higher education. International guidance does not fully specify how universities in Vietnam should operationalize AI-use disclosure, academic integrity policy, assessment redesign, data governance, faculty training, and student AI literacy under conditions of uneven infrastructure and institutional capacity. Similarly, Vietnam's national AI and digital transformation policies provide macro-level direction, but they do not yet offer detailed mechanisms for governing GenAI in teaching, learning, assessment, and research.

This paper addresses that gap by developing the GenAI Responsible Adoption Framework for Vietnamese Higher Education (GRAV-HE). The study asks: What governance dimensions are required to support responsible, equitable, and SDG 4-aligned GenAI adoption in Vietnamese higher education institutions? The contribution of the paper is threefold. First, it reframes GenAI governance as a sustainable education issue rather than a narrow problem of academic misconduct or technology management. Second, it operationalizes Responsible AI and SDG 4 principles into four institutional governance dimensions: ethical policy and AI disclosure standards, data governance and secure digital infrastructure, pedagogical transformation, and comprehensive AI literacy development. Third, it contextualizes these dimensions for Vietnamese higher education, where digital transformation ambitions, data protection requirements, educational equity concerns, and uneven institutional readiness must be considered together. GRAV-HE does not claim to introduce entirely new AI ethics principles; rather, its originality lies in integrating and operationalizing existing principles into a Vietnam-oriented institutional governance framework that can support policy dialogue, assessment reform, capacity building, and future empirical validation.

Literature Review

GenAI in Higher Education: From Learning Support to Governance Risk

Generative artificial intelligence (GenAI) has rapidly become a major issue in higher education because it changes not only how students access information, but also how they produce, revise, and submit academic work. Unlike earlier educational technologies that mainly supported communication, content delivery, learning management, or information retrieval, GenAI systems can generate original-seeming academic outputs through natural language interaction. These outputs may include essays, summaries, programming code, explanations, feedback, lesson materials, presentation outlines, and research-related text. This generative capacity makes GenAI qualitatively different from many previous digital learning tools: it does not merely support the learning environment, but can directly participate in the production of academic artefacts.

The literature has widely recognized the educational potential of GenAI. For students, large language model-based tools can provide immediate explanations of difficult concepts, support brainstorming, assist with writing and revision, generate examples, summarize complex readings, support programming tasks, and offer formative feedback outside formal classroom hours. For instructors, GenAI may assist in preparing lesson plans, designing rubrics, generating discussion questions, creating case scenarios, adapting explanations to different levels of student understanding, and producing preliminary feedback drafts. Large language models offer significant opportunities for education, while also raising serious pedagogical, ethical, and social challenges [1].

However, the educational value of GenAI depends on how it is embedded into learning design, assessment, and institutional policy. If used critically, GenAI can function as a learning scaffold that supports explanation, reflection, and self-regulated learning. If used uncritically, it may become a shortcut for task completion rather than a tool for knowledge construction. UNESCO's guidance on GenAI in education and research similarly stresses that GenAI should be addressed through regulation, institutional policy, and human capacity building rather than treated as a neutral technology [2]. This distinction is important because GenAI can support learning and, at the same time, undermine the conditions through which learning is assessed.

The central challenge is that GenAI blurs the boundary between assistance and substitution. A student may use GenAI to clarify a concept, but the same tool can generate a complete essay. It may be used to improve grammar, but it can also restructure arguments, fabricate citations, or produce a polished literature review. Similarly, an instructor may use GenAI to generate examples or feedback drafts, but may also face uncertainty about whether submitted work reflects student understanding or machine-generated fluency. This dual-use nature makes GenAI difficult to govern through simple categories of "allowed" or "prohibited" use.

For this reason, GenAI has shifted from being a learning-support tool to becoming a governance issue for higher education institutions. One major concern is academic integrity. GenAI-generated text may not directly copy existing sources, which limits the effectiveness of conventional plagiarism detection systems. AI-detection tools also remain contested because they may produce false positives and false negatives, raising concerns about fairness and due process. Recent research on GenAI and academic integrity suggests that institutions need to move beyond detection-based responses and develop clearer expectations around disclosure, authorship, assessment design, and responsible use [3, 4].

Assessment validity is another central concern. Many common assessment formats, such as take-home essays, generic literature summaries, coding assignments, and unsupervised written tasks, were designed under the assumption that submitted work largely reflects students' own reasoning and effort. In an AI-rich environment, this assumption becomes less stable. A fluent final product may no longer show how students developed ideas, evaluated evidence, solved problems, or made judgments. The rise of GenAI therefore requires universities to reconsider assessment design, not merely to strengthen misconduct detection. More resilient approaches may include process-based assessment, oral defense, portfolios, staged submissions, reflective writing, and assignments that require students to explain how they used and evaluated AI outputs.

The problem also extends to information quality and academic judgment. GenAI systems can generate persuasive but inaccurate responses, unsupported claims, biased assumptions, or fabricated references. In higher education, this creates a problem not only of factual accuracy but also of epistemic responsibility. Students may accept AI-generated outputs as authoritative without verifying them against scholarly sources. As a result, AI literacy must go beyond technical prompting skills. It should include critical evaluation, source verification, ethical judgment, and awareness of the limits of machine-generated knowledge.

Privacy and equity further complicate GenAI adoption. Students and faculty may enter personal information, unpublished work, institutional documents, or research data into commercial GenAI platforms without fully understanding how such data may be processed, stored, or reused. At the same time, access to GenAI is uneven. Students with paid subscriptions, stronger English proficiency, better devices, reliable internet, or higher digital competence may gain greater benefits than students with fewer resources. In this sense, GenAI may widen educational inequalities if institutions do not provide guidance, training, and equitable access mechanisms.

Taken together, these issues show that GenAI in higher education cannot be treated as a matter of individual choice alone. While students and instructors make daily decisions about tool use, the consequences of GenAI adoption affect institutional responsibilities: academic integrity, assessment quality, data protection, equity, faculty development, and student learning outcomes. The central question is therefore no longer only how GenAI can support learning, but how higher education institutions can govern its use in ways that preserve educational quality, fairness, transparency, and human agency. This shift from learning support to governance risk provides the basis for examining GenAI through Responsible AI principles, SDG 4, and institution-level policy design.

Academic Integrity: Detection-Based Control versus Assessment Redesign

Academic integrity has become one of the most visible and contested issues in the adoption of GenAI in higher education. Traditionally, academic misconduct has often been associated with plagiarism, unauthorized collaboration, contract cheating, fabrication of data, or the reuse of existing materials without proper acknowledgment. GenAI complicates this understanding because it can generate fluent and original-seeming text that may not directly copy any identifiable source. As a result, the boundary between acceptable academic support and inappropriate outsourcing becomes more difficult to define. A student may use GenAI to brainstorm ideas, improve grammar, or clarify a concept, but the same tool can also produce complete essays, solve assignments, generate programming code, or fabricate references. This dual-use character makes academic integrity in the GenAI era a problem of authorship, transparency, responsibility, and assessment design, not merely a problem of plagiarism detection.

Early institutional responses to GenAI have often emphasized control, prohibition, and detection. Such responses are understandable because universities need to protect the credibility of degrees, ensure fairness among students, and maintain confidence in assessment outcomes. In some courses, instructors may prohibit GenAI use entirely, especially when tasks are designed to measure independent writing, reasoning, or problem-solving. In other cases, institutions may rely on AI-detection software to identify suspected AI-generated submissions. However, a detection-based approach has significant limitations. GenAI-generated work may not produce textual similarity with existing sources, which reduces the usefulness of conventional plagiarism detection systems. At the same time, AI-detection tools remain imperfect and may produce both false positives and false negatives [16]. This creates risks of unfairly accusing students or failing to identify inappropriate use.

Recent studies on GenAI and academic integrity indicate that detection alone cannot provide a stable foundation for academic governance. ChatGPT and similar tools create both opportunities and risks for higher education because they challenge conventional assumptions about student authorship and academic honesty [3]. A recent systematic review further demonstrates that GenAI affects academic integrity across multiple dimensions, including student behavior, authorship, plagiarism, and independent learning [4]. Their review covers studies on GenAI in higher education from January 2021 to December 2024, indicating that the issue has quickly become a major concern in the post-ChatGPT period. These findings suggest that institutions should avoid treating GenAI misconduct as only a technical problem to be solved by software.

The deeper issue is assessment validity. Many traditional assignments were designed in an educational environment where submitted work was assumed to be a relatively direct representation of students' own thinking and effort. GenAI weakens this assumption. A polished essay, accurate-looking summary, or well-structured code submission may not reveal whether the student understood the concepts, evaluated evidence, developed arguments, or made decisions independently. Therefore, the key question is not simply whether AI was used, but whether the assessment still measures the intended learning outcomes. If a task can be completed mainly by prompting a GenAI system, then the task itself may need to be redesigned.

This does not mean that universities should abandon written assignments, essays, coding tasks, or take-home projects. Rather, these assessments need to be strengthened so that they capture process, reasoning, judgment, and reflection. Possible approaches include staged submissions, annotated drafts, oral defense, learning journals, portfolios, in-class components, localized case analysis, project-based assessment, and reflective AI-use statements [17]. These approaches shift assessment away from the final product alone and toward the learning process behind the product.

They also make it easier for instructors to evaluate how students use GenAI, whether they understand the material, and whether they can justify their decisions.

AI-use disclosure is central to this shift. Instead of assuming that any AI use is automatically misconduct, universities can distinguish between permitted, restricted, and prohibited uses. For example, using GenAI for grammar checking, brainstorming, or feedback on drafts may be permitted in some contexts, while using it to generate an entire essay or fabricate references should remain prohibited. Disclosure allows students to document which tools they used, for what purpose, and to what extent. It also reinforces the principle that students remain accountable for the accuracy, originality, and integrity of submitted work. In this sense, disclosure is not a substitute for academic integrity; it is a mechanism for making academic responsibility visible.

For Vietnamese higher education institutions, this issue is particularly important because institutional policies on GenAI use are still emerging and may vary across universities, faculties, and individual courses. Without clear guidance, students may receive inconsistent messages about what counts as acceptable AI assistance. Faculty members may also differ in their ability to detect, interpret, or respond to AI-assisted work. A balanced approach is therefore needed: universities should maintain clear rules against misconduct, but they should also support assessment redesign, transparent disclosure practices, and AI literacy development. Academic integrity in the GenAI era should not be framed only as a problem of cheating; it should be understood as a governance challenge involving fairness, transparency, learning quality, and institutional responsibility.

Responsible AI Frameworks: Global Principles and Local Operationalization

Responsible AI provides an important normative foundation for governing GenAI in higher education. Rather than treating AI as a purely technical innovation, Responsible AI emphasizes that AI systems should be designed, adopted, and governed in ways that protect human agency, fairness, transparency, privacy, safety, accountability, and social well-being. These principles are particularly relevant to higher education because GenAI affects not only access to information, but also academic authorship, assessment validity, student data, pedagogical relationships, and educational opportunity. In this sense, Responsible AI shifts the discussion from whether GenAI should be used toward how it can be used ethically, safely, and educationally.

Several international AI governance frameworks provide important reference points for this discussion. The OECD AI Principles promote trustworthy AI that is innovative, human-centered, and respectful of human rights and democratic values [12]. These principles emphasize inclusive growth, human-centered values, transparency, robustness, safety, security, and accountability. UNESCO's Recommendation on the Ethics of Artificial Intelligence provides a global standard for ethical AI and stresses human rights, transparency, fairness, accountability, safety, and social responsibility [19]. The NIST AI Risk Management Framework contributes a risk-based approach for identifying, measuring, managing, and governing AI-related risks [20]. Although these frameworks were not designed exclusively for higher education, they establish the ethical and governance vocabulary needed to evaluate GenAI adoption in universities.

In the education sector, UNESCO's Guidance for generative AI in education and research is especially relevant. It highlights the need for regulation, institutional policy, teacher and learner capacity building, and a human-centered approach to GenAI adoption. This guidance is useful because it recognizes that GenAI cannot be left entirely to market forces, individual users, or ad hoc institutional decisions. It also emphasizes that education systems need to protect learners, support teachers, promote inclusion, and ensure that AI use remains aligned with educational purposes. For higher education institutions, this means that GenAI governance should involve more than academic misconduct control; it should also address pedagogical design, data protection, equity, and human capacity development.

The academic literature on AI ethics shows a similar convergence around core principles. Jobin, Ienca, and Vayena's review of global AI ethics guidelines identifies transparency, justice and fairness, non-maleficence, responsibility, and privacy as recurring themes across many AI ethics documents [21]. Floridi and Cowls also propose a unified framework for AI ethics built around beneficence, non-maleficence, autonomy, justice, and explicability [22]. These principles are valuable because they provide a shared ethical language for evaluating AI systems across different sectors. In higher education, however, such principles need to be translated into concrete institutional practices. For example, transparency becomes meaningful when students disclose whether and how they used GenAI in assignments. Accountability becomes meaningful when universities clarify who is responsible for AI-assisted outputs, grading decisions, feedback, and data handling. Privacy becomes meaningful when institutions define what types of student, research, or institutional data must not be entered into commercial AI platforms.

Fairness is also central to the higher education context. A general commitment to fairness is not sufficient unless institutions address unequal access to AI tools, devices, reliable internet, English-language resources, and AI literacy. Students with stronger digital skills or access to paid GenAI tools may gain advantages over those with fewer resources. Therefore, Responsible AI in education must be understood not only as a matter of avoiding algorithmic bias, but also as a matter of educational equity. Similarly, human oversight is not simply a technical principle; it requires universities to ensure that AI does not replace academic judgment, pedagogical responsibility, or meaningful teacher-student interaction. In this way, Responsible AI principles become institutionally relevant only when they are connected to academic policies, assessment design, data governance, and capacity building.

Despite their importance, global AI governance frameworks have limitations when applied directly to Vietnamese higher education. Their principles are often broad and normative, but they do not fully specify how universities should design course-level AI-use rules, disclosure templates, assessment procedures, data protection protocols, faculty training, or student AI literacy programs. They also tend to assume a level of institutional capacity that may not be evenly available across developing-country higher education systems. In Vietnam, universities differ in digital infrastructure, legal and technical expertise, faculty readiness, student access, and experience with AI governance. As a result, international frameworks such as UNESCO and OECD provide necessary guidance, but they require contextual translation before they can become actionable at the institutional level.

This gap between global principles and local implementation is central to the present study. The purpose of GRAV-HE is not to replace UNESCO, OECD, NIST, or existing AI ethics frameworks. Rather, its contribution lies in operationalizing their principles for Vietnamese higher education by organizing them into four institutional governance dimensions: ethical policy and AI disclosure, data governance and secure infrastructure, pedagogical transformation, and comprehensive AI literacy. In doing so, GRAV-HE serves as a bridge between abstract Responsible AI principles and practical university-level governance. This approach is particularly important for Vietnam and comparable developing-country contexts, where responsible GenAI adoption requires not only ethical principles, but also feasible institutional mechanisms that can be adapted to uneven levels of readiness and resources.

Table 1: Translating Responsible AI principles into higher education governance mechanisms

Responsible AI principle	Meaning in GenAI governance	Possible HEI mechanism
Transparency	Users should know and disclose when AI is used	AI-use disclosure form; syllabus statement
Accountability	Human actors remain responsible for AI-assisted work	Student responsibility declaration; faculty grading policy
Privacy	Personal and institutional data must be protected	Data classification; restricted data input rules
Fairness	AI adoption should not widen educational inequality	Equitable access support; AI literacy training
Safety and reliability	AI outputs should be critically checked	Verification requirement; source-checking practice
Human oversight	AI should not replace academic judgment	Instructor review; oral defense; reflective assessment

Developing-Country and Vietnamese Higher Education Contexts

The governance of GenAI in higher education cannot be separated from institutional and national context. Much of the early discussion on GenAI policies, guidelines, and responsible adoption has been shaped by universities and education systems with relatively strong digital infrastructure, established academic governance mechanisms, and greater access to technological resources [23, 24]. Policy models developed in such contexts may not be directly transferable to developing-country higher education systems, where institutional capacity, digital infrastructure, faculty readiness, student access, and regulatory implementation may vary considerably.

In developing-country higher education contexts, GenAI presents a dual challenge. On the one hand, it may help expand access to academic support, provide language assistance, support formative feedback, and reduce some barriers to learning resources. These affordances are particularly relevant where large class sizes, limited tutoring support, and uneven access to academic resources are persistent challenges. On the other hand, GenAI may deepen existing inequalities if its effective use depends on paid subscriptions, strong English proficiency, reliable internet

access, advanced devices, or prior digital competence. The governance question is therefore not simply whether GenAI is useful, but whether institutions can prevent technological advantage from becoming educational advantage.

Vietnam illustrates this dual condition clearly. National digital indicators show both significant connectivity and remaining exclusion. In early 2025, Vietnam had 79.8 million internet users, but approximately 21.5 million people still did not use the internet. Mobile connectivity was extensive, with 127 million cellular mobile connections, equivalent to 126% of the population, but mobile connection figures do not necessarily mean that all learners have stable, affordable, and educationally effective access to GenAI tools [25]. DataReportal also notes that some mobile connections may only include voice or SMS services and should not be treated as a direct proxy for internet use. This distinction matters for GenAI governance because using GenAI effectively often requires stable connectivity, suitable devices, digital skills, and the ability to evaluate AI-generated outputs.

Recent official communications also show rapid infrastructure development. Vietnam reported 110.5 million mobile broadband subscriptions, with 91.3% of subscribers using smartphones, and the proportion of households using fiber-optic internet reaching 85.8% [25]. These figures indicate a strong digital infrastructure base for AI-enabled education. However, the same policy environment still recognizes the need to eliminate coverage gaps and strengthen digital infrastructure. For universities, this means that GenAI governance should not assume equal readiness across all learners and institutions, even when national connectivity indicators appear strong.

The higher education sector further demonstrates why a context-sensitive approach is needed. Vietnam's higher education system included 243 universities and 2,355,711 students in 2024, with 1,819,416 students in public universities and 536,295 students in private universities. The system also included 88,031 full-time lecturers [6]. These figures show that GenAI governance must operate across a large and heterogeneous sector rather than within a small number of elite institutions. Differences between public and private universities, urban and regional institutions, research-intensive and teaching-oriented universities, and well-resourced and resource-constrained programs may affect how GenAI policies can be implemented.

Vietnam's education digital transformation also provides an important foundation for GenAI governance. The Ministry of Education and Training has implemented the Higher Education Management Information System (HEMIS), digitizing data from 442 training establishments, more than 152,000 lecturers, and more than 2.1 million students. This suggests that Vietnam already has substantial digital governance initiatives in higher education. However, the existence of digital data systems does not automatically resolve GenAI-specific governance issues. HEMIS may support institutional data management, but universities still need rules for AI-use disclosure, acceptable and prohibited GenAI use, student data protection, AI-assisted assessment, faculty training, and responsible platform selection.

The gap between national digital transformation and classroom-level governance is especially important. A national AI strategy can encourage innovation, but it does not tell an instructor whether students may use GenAI for grammar correction, translation, coding support, literature summaries, thesis writing, or examination preparation. It also does not specify how students should disclose AI use, how faculty should evaluate AI-assisted work, or what kinds of student or research data should not be entered into commercial AI platforms. In the absence of institutional guidance, GenAI practices may be governed informally by individual instructors, creating inconsistency across courses and programs.

For these reasons, Vietnam should not be treated merely as a background setting for applying global AI principles. It represents a higher education context where rapid digital transformation, large system scale, uneven readiness, emerging data protection requirements, and SDG 4-oriented equity concerns intersect. This context justifies the need for GRAV-HE as a framework that translates global Responsible AI principles into feasible institutional governance mechanisms for Vietnamese HEIs. The framework does not claim to be empirically validated across all Vietnamese universities; rather, it provides a policy-informed structure that can guide future stakeholder consultation, institutional case studies, and pilot implementation.

SDG 4 as a Governance Lens rather than a Label

Sustainable Development Goal 4 (SDG 4) provides a normative foundation for examining GenAI governance in higher education. SDG 4 aims to ensure inclusive and equitable quality education and promote lifelong learning opportunities for all [11]. It therefore places educational quality, equity, inclusion, skills development, and institutional capacity at the center of sustainable development. In the context of GenAI, this means that responsible adoption should not be evaluated only by whether universities allow or prohibit AI tools. It should also be evaluated by whether GenAI use strengthens meaningful learning, protects students, reduces inequality, supports teachers, and contributes to safe and effective learning environments.

This framing is important because GenAI can both support and undermine the aims of SDG 4. On the positive side, GenAI may expand access to learning support by providing explanations, feedback, examples, language assistance, and programming support beyond formal classroom hours. It may also assist instructors in designing teaching materials, adapting explanations, creating learning activities, and supporting formative feedback. These possibilities are relevant to SDG 4 because they may contribute to more flexible, responsive, and inclusive learning environments, particularly where institutional resources or instructor time are limited. UNESCO's guidance on GenAI in education similarly emphasizes that GenAI should be governed through policy, regulation, and human capacity building to maintain a human-centered vision of education [2].

However, GenAI may also weaken SDG 4 if it is adopted without appropriate governance. Unequal access to paid AI tools, reliable internet, high-quality devices, English-language resources, and AI literacy may widen existing educational inequalities. Students with stronger digital competence or access to premium tools may gain advantages over those with fewer resources. In addition, overreliance on AI-generated outputs may reduce opportunities for independent thinking, writing development, problem-solving, and critical judgment. GenAI may also compromise educational quality when fabricated references, biased assumptions, or inaccurate explanations are accepted without verification. From an assessment perspective, AI-generated submissions may make it more difficult for universities to determine whether students have achieved intended learning outcomes. These risks show that GenAI governance is essential not only for academic integrity, but also for protecting the quality and equity of higher education.

Several SDG 4 targets are directly relevant to GenAI governance. Target 4.4 emphasizes the development of relevant skills for employment, decent work, and entrepreneurship. In the GenAI era, this target supports the need for AI literacy, digital competence, critical evaluation, and responsible use skills among students and faculty. Target 4.7 emphasizes education for sustainable development, human rights, responsible citizenship, and ethical awareness. This target supports the inclusion of AI ethics, disclosure, academic responsibility, and critical reflection in university policies and curricula. Target 4.a focuses on safe, inclusive, and effective learning environments, which is relevant to data protection, privacy, secure digital infrastructure, and equitable access to AI-supported learning. Target 4.c emphasizes the supply and capacity of qualified teachers, which connects directly to faculty development for AI-aware pedagogy, assessment redesign, and responsible tool use. These targets show that SDG 4 is not external to GenAI governance; it provides criteria for judging whether GenAI adoption is educationally responsible.

Using SDG 4 as a governance lens changes the design of a GenAI framework in several ways. Without SDG 4, institutional governance might focus mainly on compliance, misconduct prevention, or technical control. With SDG 4, the framework must also ask whether GenAI adoption improves or undermines inclusive and equitable quality education. This shifts attention from narrow control toward broader educational development. For example, an AI-use policy should not only define prohibited behavior; it should also help students learn responsible AI practices. A data governance policy should not only reduce legal risk; it should also protect learners and maintain trust in the educational environment. Assessment redesign should not only prevent cheating; it should preserve meaningful evidence of learning. Faculty training should not only introduce AI tools; it should strengthen teachers' capacity to guide students ethically and pedagogically.

This SDG 4 lens is especially important for Vietnamese higher education because digital transformation is a national priority, but institutional readiness remains uneven. If GenAI is adopted primarily by students with better resources, stronger English proficiency, or access to paid tools, it may intensify existing educational inequalities. If universities respond only through prohibition and detection, they may miss the opportunity to develop AI literacy and future-oriented skills. Conversely, if GenAI is allowed without institutional safeguards, it may undermine academic integrity, data protection, and assessment validity. Therefore, SDG 4 requires a balanced governance approach that supports innovation while protecting equity, quality, and human agency.

Although SDG 4 is the central normative anchor of this study, GenAI governance also has implications for other Sustainable Development Goals. SDG 10 is relevant because unequal access to GenAI tools and AI literacy may deepen educational inequality. SDG 9 is relevant because responsible GenAI adoption depends on digital infrastructure, innovation capacity, and secure technological systems. SDG 16 is also relevant because transparent policies, accountability mechanisms, and responsible institutional governance are necessary for trust and fairness in AI-mediated education. More broadly, research on AI and the SDGs shows that AI can support many sustainable development targets, but it may also inhibit progress if governance, transparency, accountability, and equity are neglected [26].

Taken together, SDG 4 provides more than a thematic connection to sustainable development. It functions as a normative filter for deciding what responsible GenAI adoption should include. A GenAI governance framework

aligned with SDG 4 must therefore address not only academic integrity and policy compliance, but also equitable access, AI literacy, data protection, safe learning environments, teacher capacity, and meaningful assessment. This understanding provides the basis for the GRAV-HE framework, which translates Responsible AI and SDG 4 principles into institutional governance dimensions for Vietnamese higher education.

Methodology: Conceptual Policy Analysis and Framework Development

Research Design

This study adopts a conceptual policy analysis and framework development approach. It is not designed as a systematic review, scoping review, empirical survey, interview study, or institutional case study. This design is appropriate because the purpose of the paper is not to test hypotheses, measure stakeholder perceptions, or evaluate the effectiveness of a specific GenAI intervention. Rather, the study aims to integrate existing academic literature, international AI governance guidance, Vietnamese policy documents, and SDG 4-oriented education principles in order to develop a context-sensitive governance framework for responsible GenAI adoption in Vietnamese higher education.

A conceptual framework development approach is suitable for emerging and interdisciplinary research areas where concepts, risks, and institutional practices are still evolving [27]. GenAI governance in higher education is such an area because it involves overlapping concerns related to academic integrity, assessment validity, data protection, digital equity, AI literacy, pedagogical transformation, and sustainable education. These issues cannot be adequately addressed through a single disciplinary lens. Therefore, this study synthesizes insights from education research, Responsible AI literature, policy documents, and sustainable development frameworks to identify governance dimensions relevant to Vietnamese higher education institutions [28].

The methodological contribution of the study lies in organizing dispersed academic and policy knowledge into a coherent institutional framework. The resulting GRAV-HE framework should therefore be understood as a conceptual and policy-informed model. It is intended to support institutional dialogue, policy design, assessment reform, faculty development, and future empirical validation, rather than to function as a fully tested implementation model.

Source Selection

The source selection was organized around five thematic domains corresponding to the purpose of developing a responsible GenAI governance framework for Vietnamese higher education: GenAI in higher education, academic integrity and assessment, Responsible AI governance, Vietnamese digital transformation and personal data protection policy, and SDG 4-oriented sustainable education. This domain-based selection logic was used because the study aims to construct an interdisciplinary conceptual framework rather than to provide an exhaustive systematic review of all publications on GenAI [28, 29].

Sources on GenAI in higher education were used to identify the educational opportunities, risks, and institutional implications of large language model-based tools. These studies helped clarify how GenAI functions not only as a learning support technology, but also as a governance challenge for universities. Literature on academic integrity and assessment was then used to examine issues such as AI-assisted writing, authorship, plagiarism detection, AI-use disclosure, assessment validity, and assessment redesign. This body of work informed the argument that GenAI governance cannot rely solely on prohibition or detection.

Responsible AI and AI governance sources provided the ethical foundation for the framework. International guidance and academic literature on transparency, accountability, fairness, privacy, safety, and human oversight were used to identify the normative principles underlying responsible GenAI adoption. Vietnamese policy and legal documents were included to contextualize the framework in relation to Vietnam's national AI strategy, education digital transformation agenda, and personal data protection requirements. These documents were especially important for connecting global AI governance principles with the institutional and regulatory conditions of Vietnamese higher education.

SDG 4 and sustainable education sources provided the broader developmental orientation of the framework. They were used to link GenAI governance with inclusive and equitable quality education, digital and AI-related skills, safe learning environments, and teacher capacity. This helped ensure that the framework was not limited to academic

misconduct prevention or technical risk management, but also addressed educational equity, capacity building, and sustainable higher education development.

Sources were prioritized when they met one or more of the following criteria: direct relevance to GenAI in higher education; contribution to academic integrity, assessment, AI ethics, or AI governance debates; policy relevance to Vietnam or international education governance; and relevance to SDG 4, digital equity, or sustainable education. Particular attention was given to literature published from 2019 onward, with emphasis on post-2022 studies due to the rapid public adoption of GenAI tools following the release of advanced large language model applications. Foundational sources published before 2019 were included when they provided important conceptual or methodological grounding for AI ethics, conceptual framework development, or literature review methodology. Table 2 summarizes the five source domains and their respective roles in the framework development process.

Table 2: Source Domains for Conceptual Policy Analysis

Source Domain	Type of Sources	Purpose in Framework Development
GenAI in Higher Education	Peer-reviewed articles, systematic reviews, major education reports	Identify educational opportunities, risks, and institutional implications of GenAI.
Academic Integrity & Assessment	Journal articles, academic integrity studies, assessment reform reports	Inform AI disclosure, authorship, assessment validity, and assessment redesign.
Responsible AI Governance	UNESCO, OECD, NIST, AI ethics literature	Provide principles such as transparency, accountability, fairness, privacy, safety, and human oversight.
Vietnamese Policy & Legal Context	Decision No. 127/QD-TTg, Decision No. 131/QD-TTg, Decree No. 13/2023/ND-CP	Contextualize the framework within Vietnam's AI strategy, education digital transformation, and data protection environment.
SDG 4 & Sustainable Education	UN, UNESCO, SDG-related literature	Anchor the framework in inclusive, equitable, and quality education.

Analytical Procedure

The analytical procedure followed a six-step framework development process. First, relevant source domains were identified based on the study aim, including GenAI in higher education, academic integrity and assessment, Responsible AI governance, Vietnamese policy, and SDG 4-oriented sustainable education. Second, recurring governance concerns were extracted from the selected sources, including academic integrity, AI-use disclosure, assessment validity, privacy, platform security, unequal access, AI literacy, faculty readiness, misinformation, and pedagogical redesign. Third, these concerns were grouped into thematic clusters. Issues related to authorship and hidden AI use were grouped under ethical policy and disclosure; issues related to data, platforms, and access were grouped under data governance and secure infrastructure; concerns about assessment validity and learning evidence were grouped under pedagogical transformation; and concerns about student and faculty capacity were grouped under AI literacy.

Fourth, the thematic clusters were mapped against Responsible AI principles and SDG 4 targets. This step connected transparency, accountability, fairness, privacy, safety, and human oversight with SDG-oriented concerns such as quality education, relevant skills, safe learning environments, and teacher capacity. Fifth, the mapped themes were interpreted in relation to Vietnamese higher education, including national AI policy, education digital transformation, personal data protection, uneven institutional readiness, and differences in student and faculty capacity. Finally, the clusters were consolidated into the four dimensions of the GRAV-HE framework: ethical policy and AI disclosure standards, data governance and secure digital infrastructure, pedagogical transformation, and comprehensive AI literacy development.

Table 3: Analytical procedure for developing the GRAV-HE framework

Step	Analytical focus	Main analytical action	Output
Step 1. Source-domain identification	Define the conceptual boundaries of the study	Identify source domains related to GenAI in higher education, academic integrity, Responsible AI, Vietnamese policy, and SDG 4	Five source domains
Step 2. Governance concern extraction	Identify recurring governance issues	Extract concerns such as academic integrity, disclosure, privacy, equity, assessment validity, AI literacy, and faculty readiness	Initial list of governance concerns
Step 3. Thematic clustering	Organize related concerns into broader categories	Group similar concerns into institutional governance themes	Preliminary governance clusters
Step 4. Normative mapping	Connect themes with ethical and SDG-oriented principles	Map clusters to Responsible AI principles and SDG 4 targets	Ethical and sustainable development alignment
Step 5. Vietnamese contextual interpretation	Adapt the mapped themes to Vietnam's HE context	Interpret themes in relation to national policies, data protection, digital infrastructure, and institutional readiness	Context-sensitive governance priorities
Step 6. Framework construction	Consolidate the final framework	Organize governance priorities into four GRAV-HE dimensions	Four-dimensional conceptual framework

Methodological Scope and Limitations

The methodological scope of this study is conceptual and policy-oriented. The paper does not collect primary data through surveys, interviews, focus groups, experiments, or institutional case studies. It also does not claim to provide a systematic review or a comprehensive mapping of all literature on GenAI in higher education. Instead, it synthesizes selected academic, policy, and SDG-related sources to develop a theoretically grounded and policy-informed framework for responsible GenAI adoption in Vietnamese higher education.

Several limitations follow from this scope. First, the GRAV-HE framework has not yet been empirically validated with Vietnamese higher education stakeholders, including students, faculty members, university leaders, educational technologists, legal officers, and policymakers. Therefore, it should not be interpreted as a tested implementation model. Second, Vietnamese HEIs differ in size, resources, digital infrastructure, governance capacity, disciplinary focus, and student populations. The framework should therefore be understood as an adaptable model rather than a one-size-fits-all prescription. Third, because GenAI tools, regulations, and institutional practices are evolving rapidly, the framework should remain open to revision as new evidence and policy developments emerge.

These limitations clarify the appropriate interpretation of the study's contribution. GRAV-HE is intended to organize existing academic, policy, and SDG-oriented knowledge into a context-sensitive governance model that can support institutional dialogue, policy development, assessment reform, data governance, faculty development, and future empirical research. Subsequent studies should refine and test the framework through expert interviews, Delphi studies, student and faculty surveys, institutional case studies, and pilot implementation in selected Vietnamese higher education institutions.

The GRAV-HE Framework

Positioning GRAV-HE in Relation to Existing Global Guidance

The GRAV-HE framework is not proposed as a replacement for existing international AI governance guidance. Rather, it builds on and operationalizes major global frameworks on Responsible AI and GenAI in education. UNESCO's guidance on GenAI in education and research provides important direction on regulation, institutional policy, human capacity building, and responsible educational use [2]. The OECD AI Principles establish broader standards for trustworthy and human-centered AI, including transparency, accountability, robustness, safety, fairness, and respect for human rights [12]. The NIST AI Risk Management Framework further contributes a risk-based approach for identifying, measuring, managing, and governing AI-related risks [20]. Together, these frameworks provide the normative foundation for responsible GenAI governance in higher education.

However, these global frameworks are necessarily broad. They provide principles and policy directions, but they do not fully specify how universities in developing-country contexts should translate them into everyday institutional practices. For example, global guidance does not usually define how a Vietnamese university should write AI-use statements in course syllabi, design AI disclosure forms, classify restricted data, decide which GenAI tools are acceptable for academic use, redesign assessment tasks, train faculty members, or support students with unequal access to AI tools. This creates a gap between international principles and local implementation.

GRAV-HE addresses this gap by functioning as an intermediate institutional framework. Its purpose is not to introduce entirely new AI ethics principles, but to translate existing Responsible AI and SDG 4 principles into governance dimensions that are meaningful for Vietnamese higher education institutions. In this framework, transparency is operationalized through AI-use disclosure and syllabus-level guidance. Accountability is operationalized through student responsibility, faculty oversight, and institutional policy. Privacy and safety are operationalized through data governance and secure platform use. Fairness is operationalized through attention to unequal access, digital infrastructure, and AI literacy. Human oversight is operationalized through assessment redesign, pedagogical judgment, and faculty capacity building.

The originality of GRAV-HE therefore lies in its contextual and operational contribution. It integrates Responsible AI principles, SDG 4 targets, Vietnamese policy conditions, and higher education governance needs into a four-dimensional framework: ethical policy and AI disclosure standards, data governance and secure digital infrastructure, pedagogical transformation, and comprehensive AI literacy development. This positioning clarifies that GRAV-HE should be understood as a Vietnam-oriented conceptual policy framework rather than as a universal AI governance model or an empirically validated implementation tool.

Next page

Table 4: Positioning GRAV-HE in relation to existing global AI governance guidance

Aspect	Global guidance	Gap for Vietnamese HEIs	GRAV-HE response
Ethical principles	Broad principles on transparency, accountability, fairness, privacy, safety, and human oversight	Limited guidance for course-level GenAI rules	AI-use categories, disclosure templates, syllabus statements, and responsibility rules
Academic integrity	Responsible educational use and institutional policy are emphasized	Unclear distinction between permitted, restricted, and prohibited AI use	Ethical policy and AI disclosure standards for assignments, theses, projects, and research outputs
Data protection	Risk management, robustness, safety, and accountability are emphasized	Limited rules for student data, research drafts, classroom data, and institutional documents	Restricted data categories, approved tools, anonymization, and secure platform guidance
Pedagogical use	Human-centered GenAI use in education is encouraged	Limited guidance for assessment redesign in AI-rich classrooms	Process-based assessment, oral defense, portfolios, staged submissions, and reflective AI-use statements
Capacity building	Human capacity development is recommended	Student and faculty AI literacy needs differ across institutions	Functional, critical, ethical, pedagogical, and data-related AI literacy
SDG orientation	Inclusion, equity, and quality education are supported at a broad level	SDG 4 may remain a policy label without institutional mechanisms	Governance actions mapped to SDG 4.4, 4.7, 4.a, 4.c, with links to SDG 9, 10, and 16
Developing-country context	Global frameworks are designed for broad applicability	Vietnamese HEIs vary in infrastructure, faculty readiness, policy capacity, and student access	Adaptable governance model for institutions with different readiness levels

As shown in Table 4, GRAV-HE is positioned as an operational and contextual framework rather than as a new set of AI ethics principles. Its contribution lies in connecting global Responsible AI guidance with the institutional realities of Vietnamese higher education. By translating broad principles into governance mechanisms related to disclosure, data protection, assessment redesign, and AI literacy, the framework provides a practical structure for universities seeking to adopt GenAI responsibly under conditions of uneven readiness and resources.

Overview of the Framework

Building on the conceptual policy analysis presented in Section 3, the GenAI Responsible Adoption Framework for Vietnamese Higher Education (GRAV-HE) is designed to support responsible, equitable, and SDG 4-aligned GenAI adoption in Vietnamese higher education institutions. The framework responds to the need for institutional governance mechanisms that can translate global Responsible AI principles into practical policies and practices within the Vietnamese higher education context.

GRAV-HE consists of four interdependent dimensions: ethical policy and AI disclosure standards, data governance and secure digital infrastructure, pedagogical transformation, and comprehensive AI literacy development. These dimensions are not separate or sequential components. Rather, they function as mutually reinforcing areas of institutional governance. Ethical policy clarifies acceptable and unacceptable AI use. Data governance protects privacy

and supports safe platform use. Pedagogical transformation ensures that GenAI does not undermine assessment validity or learning quality. AI literacy equips students and faculty with the capacity to use GenAI critically, ethically, and responsibly.

The framework is grounded in two normative foundations. The first is Responsible AI, which emphasizes transparency, accountability, fairness, privacy, safety, and human oversight. The second is SDG 4, which emphasizes inclusive and equitable quality education, relevant skills development, safe learning environments, and teacher capacity. In the Vietnamese context, these foundations must be interpreted in relation to national digital transformation policies, personal data protection requirements, uneven institutional readiness, and differences in student and faculty capacity.

Figure 1: The GRAV-HE framework for responsible GenAI adoption in Vietnamese higher education

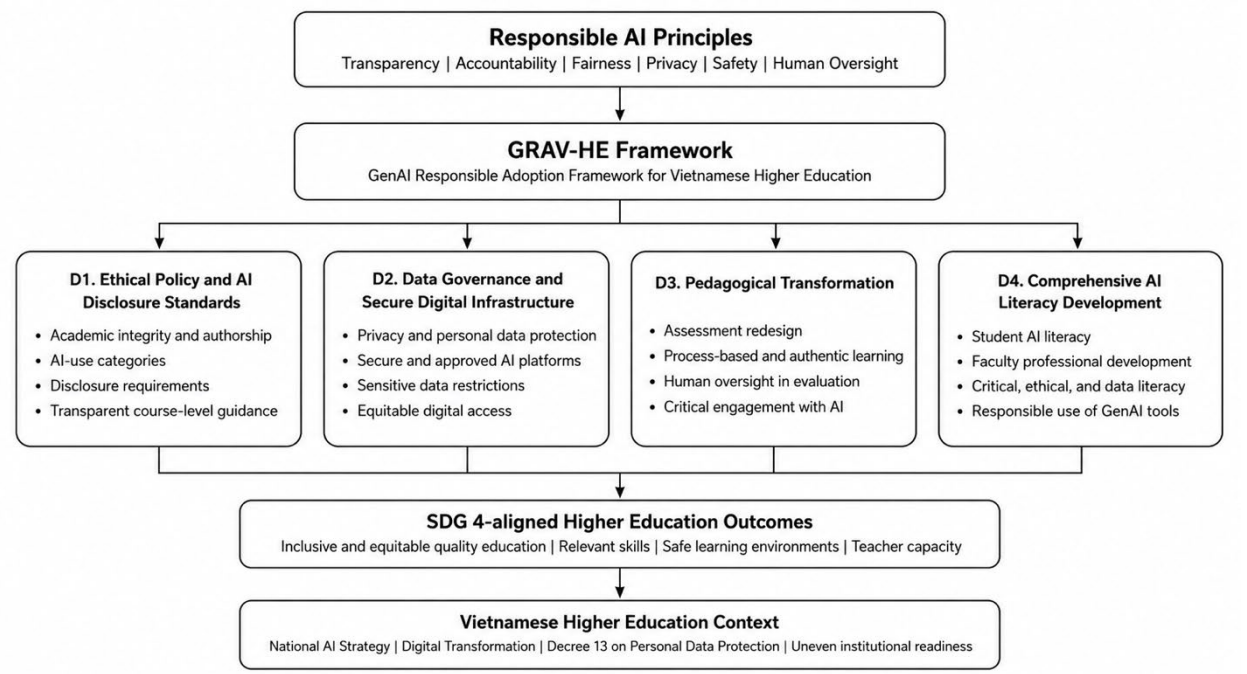

Figure 1 positions GRAV-HE between global Responsible AI principles and SDG 4-aligned higher education outcomes. The four dimensions translate broad principles such as transparency, accountability, fairness, privacy, safety, and human oversight into institutional governance mechanisms for Vietnamese higher education. The framework also reflects Vietnam's policy and institutional context, including national AI strategy, digital transformation, personal data protection, and uneven institutional readiness.

Dimension 1: Ethical Policy and AI Disclosure Standards

The first dimension of GRAV-HE addresses the need for clear institutional policies on acceptable, restricted, and prohibited uses of GenAI in academic work. This dimension responds to one of the most immediate governance problems created by GenAI: the boundary between legitimate learning support and inappropriate academic outsourcing. Without explicit guidance, the same AI-assisted practice may be interpreted differently across courses, faculties, and institutions. For example, using GenAI for grammar correction may be permitted in one course, restricted in another, and treated as misconduct elsewhere. Such inconsistency can weaken fairness, transparency, and student trust.

Ethical policy should therefore move beyond general warnings about AI misuse. Universities need to define categories of AI use at the institutional and course levels. Permitted uses may include brainstorming, grammar checking, concept clarification, language support, or feedback on drafts when these uses are consistent with learning outcomes. Restricted uses may include AI-assisted writing, translation, coding support, data analysis, or literature summarization that must be disclosed and critically reviewed. Prohibited uses should include submitting AI-generated work as one's own, fabricating references, generating false data, or using GenAI to complete assessments designed to measure independent performance.

AI-use disclosure is central to this dimension. Disclosure should not be framed automatically as evidence of misconduct. Instead, it should be treated as a mechanism for transparency and academic responsibility. A disclosure template may ask students to identify the AI tool used, explain the purpose of use, describe the extent of assistance, and confirm that they remain responsible for the accuracy, originality, and integrity of the submitted work. This approach allows universities to distinguish between responsible AI-assisted learning and inappropriate outsourcing.

For Vietnamese higher education institutions, this dimension is especially important because institutional GenAI policies are still emerging and may vary across universities and programs. A clear policy structure can reduce uncertainty for both students and faculty. It can also support more consistent academic integrity procedures by clarifying expectations before misconduct occurs. In this sense, ethical policy and disclosure standards operationalize the Responsible AI principles of transparency and accountability, while also supporting SDG 4.7 through education for ethical awareness, responsible citizenship, and sustainable development.

Dimension 2: Data Governance and Secure Digital Infrastructure

The second dimension focuses on data governance, privacy, security, platform use, and equitable access. GenAI platforms often operate through commercial infrastructures that process user inputs on external servers. When students or faculty enter personal data, unpublished research, classroom materials, institutional documents, examination content, or student submissions into these systems, they may expose sensitive information to risks that are not fully controlled by the university. For this reason, data governance is not only a technical issue, but a core component of responsible GenAI adoption.

Universities should establish clear rules about what types of data may or may not be entered into GenAI tools. Restricted data should include identifiable student information, examination materials, unpublished research data, confidential institutional documents, student records, and any personal data protected under relevant regulations. Where possible, institutions should identify approved tools, recommend safer alternatives, and provide guidance on data minimization, anonymization, and responsible platform use. Even when universities lack the resources to develop local AI systems, basic data-use rules can reduce avoidable risks.

In Vietnam, this dimension is particularly relevant because Decree No. 13/2023/ND-CP on personal data protection creates a legal and institutional basis for protecting personal data [10]. However, legal requirements must be translated into everyday academic practice. Faculty members need to know whether they may use GenAI to summarize student submissions, generate feedback, process classroom data, or support grading preparation. Students need to understand why they should not upload personal information, unpublished theses, internal documents, or identifiable data into public AI platforms.

Data governance is also linked to educational equity. Responsible GenAI adoption requires not only protecting users from risk, but also ensuring that students have fair opportunities to benefit from AI-supported learning. If only students with paid subscriptions, strong English proficiency, reliable internet, and advanced devices can use GenAI effectively, AI adoption may widen existing inequalities. Universities should therefore consider shared access models, institutional accounts, library-supported AI guidance, low-cost training, and support for students with limited digital resources. This dimension operationalizes the Responsible AI principles of privacy, safety, and *fairness, while supporting SDG 4.a on safe, inclusive, and effective learning environments.*

Dimension 3: Pedagogical Transformation

The third dimension emphasizes that GenAI governance cannot rely only on policy, prohibition, or detection. The deeper challenge concerns whether existing teaching and assessment practices remain valid in an AI-rich environment. Many traditional assessment formats, such as take-home essays, generic literature reviews, coding assignments, and unsupervised written tasks, were designed under the assumption that submitted work largely reflects students' own reasoning and effort. GenAI weakens this assumption because it can generate fluent and structured outputs that may appear to demonstrate understanding without revealing the student's actual learning process.

Pedagogical transformation therefore begins with assessment redesign. The purpose is not to eliminate written work or prohibit all AI use, but to ensure that assessment captures reasoning, judgment, process, and reflection. More resilient assessment approaches may include staged submissions, annotated drafts, oral defense, learning journals, portfolios, in-class writing, localized case studies, project-based tasks, and reflective AI-use statements [17]. These formats make it harder for students to outsource the entire task to GenAI and easier for instructors to evaluate how students think, decide, revise, and justify their work.

GenAI can also be integrated constructively into learning activities. Instructors may ask students to critique AI-generated answers, identify hallucinated references, compare AI outputs with scholarly sources, revise weak AI-generated arguments, or reflect on the ethical boundaries of AI assistance. Such activities shift GenAI from being an invisible shortcut to becoming an object of critical engagement. This approach supports higher-order thinking because students must evaluate, question, and improve AI-generated content rather than accept it passively.

For Vietnamese higher education institutions, pedagogical transformation must be realistic and adaptable. Faculty members may face large classes, limited time, uneven digital resources, and varying levels of AI familiarity. Therefore, assessment redesign should not be framed as a complete replacement of existing practice. Instead, institutions can begin with feasible changes such as adding disclosure statements, requiring draft history, incorporating short oral explanations, using localized prompts, or asking students to reflect on how they verified AI-generated content. This dimension operationalizes the Responsible AI principles of accountability and human oversight, while supporting SDG 4.4 by developing relevant skills such as critical thinking, digital competence, communication, and responsible AI use.

Dimension 4: Comprehensive AI Literacy Development

The fourth dimension addresses the need for systematic AI literacy development among students and faculty. Access to GenAI tools does not automatically produce responsible or effective use. Students may accept AI-generated outputs without verification, cite fabricated references, misunderstand the boundary between assistance and misconduct, or share sensitive data with public platforms. Faculty members may also differ in their ability to design AI-aware assessments, guide students, evaluate AI-assisted work, and protect student data. Therefore, AI literacy should be treated as a core governance requirement rather than as an optional technical skill.

Comprehensive AI literacy includes functional, critical, ethical, and data-related competencies. Functional literacy refers to the ability to use GenAI appropriately for tasks such as brainstorming, concept clarification, feedback, coding support, and language assistance. Critical literacy involves evaluating AI-generated outputs, identifying errors, checking references, recognizing bias, and comparing AI responses with credible sources. Ethical literacy includes understanding disclosure, authorship, academic integrity, fairness, and responsible use. Data literacy involves knowing what information should not be entered into AI systems and how to protect personal, academic, and institutional data.

For students, AI literacy can be integrated into orientation programs, academic skills courses, research methods courses, writing courses, and discipline-specific modules. Students should learn not only how to prompt AI systems, but also how to verify outputs, document AI assistance, protect data, and remain accountable for their academic work. For faculty, AI literacy should be linked to professional development. Training should cover AI-use policy design, syllabus statements, assessment redesign, data protection, responsible feedback practices, and discipline-specific examples of acceptable and unacceptable AI use.

Next Page

Table 5: Suggested practices for comprehensive AI literacy development

AI literacy component	Core meaning	Example practices
Functional literacy	Ability to use GenAI appropriately for academic and teaching-related tasks	Brainstorming, concept clarification, feedback, coding support, language assistance, and teaching material preparation
Critical literacy	Ability to evaluate, question, and verify AI-generated outputs	Check claims and references, identify errors or hallucinations, recognize bias, and compare AI responses with scholarly sources
Ethical literacy	Understanding of disclosure, authorship, academic integrity, and responsible use	Declare AI use, respect assignment boundaries, take responsibility for submitted work, and model transparent AI use
Data literacy	Awareness of what personal, academic, research, or institutional data should not be shared with AI platforms	Avoid uploading confidential data, student records, unpublished work, or identifiable information; apply data minimization
Pedagogical literacy	Understanding of AI as a learning support tool rather than a substitute for thinking or teaching judgment	Use AI to support reflection and learning processes; design AI-aware tasks, portfolios, oral defense, and reflective AI-use statements

As shown in Table 5, AI literacy in GRAV-HE is not limited to technical prompting skills. It includes the ability to use GenAI appropriately, evaluate its outputs critically, understand ethical and academic boundaries, protect sensitive data, and integrate AI into learning without replacing human judgment. This broader understanding is essential for Vietnamese HEIs because students and faculty may differ in digital competence, English proficiency, access to paid tools, and familiarity with AI limitations.

In the Vietnamese higher education context, comprehensive AI literacy can help transform GenAI from an unmanaged risk into a guided learning opportunity. This dimension operationalizes Responsible AI principles of transparency, fairness, and human agency, while supporting SDG 4.4 on relevant skills and SDG 4.c on teacher capacity.

Governance Logic of the GRAV-HE Framework

The four dimensions of GRAV-HE are interconnected. Ethical policy establishes expectations for responsible use, but it cannot be effective without AI literacy. Data governance protects privacy and safety, but it must be supported by infrastructure and user awareness. Pedagogical transformation protects assessment validity, but it depends on faculty capacity and institutional support. AI literacy enables responsible use, but it requires clear policy and meaningful learning design. Therefore, GRAV-HE should be understood as an integrated governance model rather than a list of separate recommendations.

Next page

Table 6: Core governance logic of the GRAV-HE framework

GRAV-HE dimension	Core governance problem	Main institutional response	SDG 4 link
Ethical policy and AI disclosure standards	Unclear authorship, hidden AI use, and inconsistent course-level expectations	Define permitted, restricted, and prohibited AI use; require disclosure; include AI-use rules in syllabi	SDG 4.7
Data governance and secure digital infrastructure	Privacy risks, unsafe platform use, and unequal access to AI-supported learning	Establish restricted data categories; identify approved tools; promote anonymization and equitable access	SDG 4.a
Pedagogical transformation	Reduced assessment validity and weak evidence of students' own reasoning	Redesign assessment through portfolios, oral defense, staged submissions, authentic tasks, and AI-use reflection	SDG 4.4
Comprehensive AI literacy development	Uneven student and faculty capacity to use GenAI critically and responsibly	Provide student AI literacy modules, faculty training, source verification practices, and data-awareness guidance	SDG 4.4; SDG 4.c

As shown in Table 6, GRAV-HE connects each governance dimension to a specific institutional problem. Ethical policy addresses uncertainty around authorship and acceptable use. Data governance responds to privacy, platform, and access risks. Pedagogical transformation protects assessment validity and meaningful learning. AI literacy supports the capacity of students and faculty to use GenAI critically, ethically, and responsibly. This structure helps position GRAV-HE as an operational governance framework rather than a general list of AI ethics principles.

Implementation Roadmap and Discussion

Readiness-Based Implementation Roadmap for Vietnamese HEIs

The GRAV-HE framework should be implemented as a gradual and adaptive process rather than as a fixed compliance checklist. Vietnamese higher education institutions differ in digital infrastructure, governance capacity, faculty readiness, student access, financial resources, and experience with AI-related policy development. For this reason, a readiness-based roadmap is more appropriate than a uniform timeline. Since this study is conceptual and policy-informed rather than empirically validated through institutional pilot implementation, the roadmap is presented as a sequence of readiness stages rather than as a fixed schedule of months.

At the baseline stage, universities should establish minimum governance expectations for GenAI use. This can begin with relatively low-cost actions such as developing AI-use guidelines, defining permitted, restricted, and prohibited uses, creating disclosure templates, updating academic integrity statements, and providing model syllabus language. The purpose of this stage is to reduce ambiguity and ensure that students and faculty have a shared understanding of responsible GenAI use.

Once basic governance expectations are established, institutions should move toward capacity building. Training should support students, faculty members, academic advisors, librarians, IT staff, and quality assurance units in understanding responsible GenAI use. Key topics may include AI-use disclosure, verification of AI-generated outputs, data protection, ethical boundaries, and appropriate academic use. Faculty development is especially important because instructors translate institutional policy into course-level rules, learning activities, and assessment practices.

As institutional capacity develops, departments and programs should review assessment and curriculum practices that are vulnerable to inappropriate AI outsourcing. The aim is not to remove GenAI from learning, but to ensure that assessment continues to provide valid evidence of student understanding, reasoning, and judgment. Practical strategies may include portfolios, staged submissions, oral defense, reflective writing, localized case studies, in-class components, and assignments requiring students to explain and evaluate their use of AI.

Finally, GenAI governance should become a process of monitoring and continuous adaptation. Because GenAI tools, student practices, platform policies, and regulatory expectations evolve rapidly, universities should regularly collect

feedback, review policy effectiveness, monitor equity and data risks, and update guidance on approved or restricted tools. In this way, GRAV-HE can support an ongoing governance process rather than a one-time institutional response.

Table 7: Readiness-based implementation roadmap for GRAV-HE

Stage	Main objective	Priority actions	Lead actors
Minimum governance baseline	Establish clear expectations for GenAI use	AI-use guidelines; disclosure templates; syllabus statements; academic integrity updates	University leadership; academic affairs; QA units
Capacity building	Build responsible AI awareness and literacy	Faculty workshops; student orientation; data protection guidance; library and IT support	Faculty development centers; departments; IT office; libraries
Assessment and curriculum redesign	Align assessment with GenAI realities	Portfolios; oral defense; staged submissions; localized tasks; reflective AI-use statements	Faculties; program coordinators; instructors
Monitoring and adaptation	Refine governance through feedback and review	Feedback collection; policy review; tool updates; monitoring equity and data risks	Governance board; QA units; IT office; student representatives

As shown in Table 7, the roadmap avoids prescribing a universal implementation timeline. Instead, it organizes implementation according to institutional readiness. This approach recognizes that Vietnamese HEIs may begin from different levels of governance capacity, digital infrastructure, and faculty readiness. Some institutions may start with low-cost policy clarification and disclosure templates, while others may be ready to invest in faculty development, approved platforms, and assessment redesign.

Theoretical Contributions

The GRAV-HE framework contributes to the literature in three main ways. First, it reframes GenAI governance as a sustainable higher education issue rather than a narrow problem of academic misconduct, technological adoption, or compliance. Much of the early discussion on GenAI in higher education has focused on cheating, detection, and acceptable use. While these issues remain important, they do not fully capture the broader implications of GenAI for educational quality, equity, data protection, assessment validity, and human capacity development. By positioning GenAI governance through an SDG 4 lens, this paper connects responsible AI adoption with inclusive and equitable quality education.

Second, the framework contributes by operationalizing Responsible AI principles for higher education governance. Principles such as transparency, accountability, fairness, privacy, safety, and human oversight are widely recognized in global AI governance literature, but they often remain broad and normative. GRAV-HE translates these principles into institution-level mechanisms, including AI-use disclosure, syllabus guidance, data governance, approved platform use, assessment redesign, and AI literacy development. The contribution therefore lies not in proposing entirely new AI ethics principles, but in organizing existing principles into actionable governance dimensions for universities.

Third, the framework provides a contextual contribution by extending GenAI governance discussions into Vietnamese higher education and comparable developing-country contexts. Global guidance is valuable, but it does not fully address uneven institutional readiness, differences in digital infrastructure, data protection requirements, faculty capacity, student access, and AI literacy. GRAV-HE addresses this gap by proposing an adaptable framework that connects global Responsible AI and SDG 4 principles with the policy and institutional conditions of Vietnamese HEIs.

Policy Implications

The GRAV-HE framework has several policy implications for Vietnamese higher education. At the national level, the framework suggests the need for more specific guidance on GenAI use in higher education. Existing national strategies on artificial intelligence and education digital transformation provide broad policy direction, but universities still require practical guidance on how GenAI should be governed in teaching, learning, assessment, research, and

administration. A national or sector-level guideline on GenAI in higher education could help institutions develop consistent policies while still allowing adaptation to disciplinary and institutional contexts.

At the institutional level, universities should move beyond informal or instructor-specific responses to GenAI. Institutional policies should define acceptable, restricted, and prohibited uses of GenAI; establish disclosure requirements; clarify student and faculty responsibilities; and specify how AI-related misconduct should be handled. Such policies should avoid two extremes: banning GenAI entirely without recognizing its educational potential, or allowing unrestricted use without safeguards. A balanced policy approach should support responsible innovation while protecting academic integrity, data privacy, fairness, and educational quality.

The framework also implies that GenAI governance should be integrated into quality assurance and data governance systems. Universities need to review whether existing assessment regulations remain valid in an AI-rich environment, particularly for take-home assignments, thesis writing, coding tasks, online examinations, and project-based learning. In parallel, internal rules should clarify what types of personal, academic, research, and institutional data should not be entered into public or commercial GenAI platforms. These policy mechanisms are necessary to translate broad Responsible AI principles into everyday institutional practice.

Practical Implications for Universities and Faculty

For Vietnamese HEIs, GRAV-HE provides a practical structure for organizing institutional responses to GenAI. A feasible starting point is the creation of a university-level GenAI working group involving academic affairs, quality assurance, information technology, legal or compliance units, librarians, faculty representatives, and students. This group can coordinate policy development, identify institutional risks, recommend approved tools, and support implementation across faculties and departments. Importantly, universities do not need to begin with costly technical solutions. Low-cost measures such as AI-use guidelines, disclosure templates, syllabus statements, and awareness workshops can provide an immediate governance baseline.

For academic affairs and quality assurance units, the framework highlights the need to update assessment guidance and academic integrity procedures. Assignment instructions should clearly state whether GenAI use is permitted, restricted, or prohibited. Where AI use is allowed, students should explain how they used the tool and how they verified its outputs. Quality assurance processes should also encourage programs to review assessment formats that are highly vulnerable to inappropriate AI outsourcing. This does not mean eliminating written assignments, but redesigning them to include process evidence, reflection, oral explanation, or contextualized tasks.

For faculty members, the most important practical implication is the shift from a detection-oriented approach to an assessment design approach. Instructors should not rely solely on AI-detection tools to determine misconduct. Instead, they can redesign assignments so that students must demonstrate reasoning, judgment, and learning processes. Practical strategies include staged submissions, draft annotations, oral defense, learning journals, local case-based analysis, project portfolios, and reflective statements on AI use. Faculty can also use GenAI constructively by asking students to critique AI-generated responses, identify errors, compare outputs with academic sources, and reflect on ethical use.

For students and support units, GRAV-HE emphasizes that responsible AI use must be taught rather than assumed. Students need guidance on how to use GenAI for legitimate academic support, how to disclose AI assistance, how to verify AI-generated claims, and how to avoid sharing sensitive data. Libraries, IT offices, and learning support centers can play an important role by offering workshops on source verification, prompt design, data privacy, and responsible academic writing with AI support. Together, these practices can help universities move from fragmented responses to coordinated GenAI governance.

Limitations and Future Research

The GRAV-HE framework should be interpreted as a conceptual and policy-informed model rather than a fully validated implementation framework. Since this study is based on a structured synthesis of academic and policy literature, its contribution lies in organizing existing knowledge into a context-sensitive governance model for Vietnamese higher education. Further empirical work is needed to examine how the framework operates in real institutional settings and how different stakeholders perceive its relevance, feasibility, and usefulness.

Future research can extend this study by engaging university leaders, faculty members, students, educational technologists, and policymakers through expert interviews, Delphi studies, surveys, or institutional case studies. Such studies would help refine the framework, assess its applicability across different types of Vietnamese higher education institutions, and identify discipline-specific requirements for GenAI governance. Pilot implementation in selected

universities would also provide valuable evidence on how AI-use policies, disclosure practices, assessment redesign, data governance, and AI literacy programs can be adapted in practice.

Given the rapid development of GenAI tools and related policy debates, the GRAV-HE framework should remain adaptable. Future studies should continue updating and testing the framework as technologies, regulations, and institutional practices evolve. In this way, the present study provides a foundation for ongoing research and practical experimentation rather than a final or fixed model.

Conclusion

The rapid diffusion of GenAI has created a new governance challenge for higher education. While these tools can support learning, teaching, feedback, and academic productivity, they also complicate assumptions about authorship, assessment, privacy, equity, and institutional responsibility. For Vietnamese higher education, the challenge is especially significant because national ambitions for AI and digital transformation must be translated into practical policies and pedagogical practices within institutions that differ in capacity, infrastructure, and readiness.

This paper responded to that challenge by developing the GRAV-HE framework, a context-sensitive model for responsible GenAI adoption in Vietnamese higher education. The framework integrates four dimensions: ethical policy and AI disclosure standards, data governance and secure digital infrastructure, pedagogical transformation, and comprehensive AI literacy development. These dimensions translate Responsible AI principles and SDG 4 commitments into institutional practices that can guide policy design, assessment reform, data protection, faculty development, and student learning.

The contribution of GRAV-HE lies in positioning GenAI governance as a sustainable education issue rather than a narrow technology management problem. Responsible adoption requires universities to move beyond reactive responses and toward coordinated governance that protects integrity, supports innovation, strengthens human capacity, and reduces educational inequality. Although further empirical validation is needed, the framework offers a practical foundation for institutional dialogue, policy development, and future research. For Vietnam and comparable developing-country contexts, sustainable GenAI adoption will depend not only on technological access, but on the capacity of higher education institutions to govern AI ethically, equitably, and pedagogically.

Author contributions

All authors contributed to the conceptualization, literature synthesis, framework development, manuscript drafting, and revision. All authors have read and approved the final manuscript.

References

[1] Kasneci, E., Sessler, K., Küchemann, S., Bannert, M., Dementieva, D., Fischer, F., Gasser, U., Groh, G., Günnemann, S., Hüllermeier, E., Krusche, S., Kutyniok, G., Michaeli, T., Nerdel, C., Pfeffer, J., Poquet, O., Sailer, M., Schmidt, A., Seidel, T., Stadler, M., Weller, J., Kuhn, J., & Kasneci, G. (2023). ChatGPT for good? On opportunities and challenges of large language models for education. Learning and Individual Differences, 103, Article 102274. https://doi.org/10.1016/j.lindif.2023.102274

[2] Miao, F., & Holmes, W. (2023). Guidance for generative AI in education and research. UNESCO. https://doi.org/10.54675/EWZM9535

[3] Cotton, D. R. E., Cotton, P. A., & Shipway, J. R. (2024). Chatting and cheating: Ensuring academic integrity in the era of ChatGPT. Innovations in Education and Teaching International, 61(2), 228–239. https://doi.org/10.1080/14703297.2023.2190148

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

[5] DataReportal. (2025). Digital 2025: Vietnam. DataReportal – Global Digital Insights. https://datareportal.com/reports/digital-2025-vietnam

[6] Vietnam.vn. (2025, May 22). The size of higher education increased by more than 300,000 students in the 2023–2024 academic year. https://www.vietnam.vn/en/quy-mo-giao-duc-dh-tang-hon-300-000-sinh-vien-trong-nam-hoc-2023-2024

[7] Prime Minister of Vietnam. (2021). Decision No. 127/QD-TTg: National strategy on research, development and application of artificial intelligence until 2030. Government of Vietnam.

[8] Prime Minister of Vietnam. (2022). Decision No. 131/QD-TTg: Project on strengthening the application of information technology and digital transformation in education and training in the period 2022–2025, with orientation to 2030. Government of Vietnam.
[9] Vietnam News. (2023, October 7). Basic administrative procedures in education put online. https://vietnamnews.vn/society/1594876/basic-administrative-procedures-in-education-put-online.html
[10] Government of Vietnam. (2023). Decree No. 13/2023/ND-CP on personal data protection. Government of Vietnam.
[11] United Nations. (2015). Transforming our world: The 2030 agenda for sustainable development. United Nations.
[12] OECD. (2019). OECD AI principles. Organisation for Economic Co-operation and Development. https://oecd.ai/en/ai-principles
[13] Holmes, W., Bialik, M., & Fadel, C. (2019). Artificial intelligence in education: Promises and implications for teaching and learning. Center for Curriculum Redesign.
[14] 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, 10, Article 15. https://doi.org/10.1186/s40561-023-00237-x
[15] OECD. (2023). OECD digital education outlook 2023: Towards an effective digital education ecosystem. OECD Publishing. https://doi.org/10.1787/c74f03de-en
[16] Perkins, M., Roe, J., Postma, D., McGaughran, J., & Hickerson, D. (2024). Detection of GPT-4 generated text in higher education: Combining academic judgement and software to identify generative AI tool misuse. Journal of Academic Ethics, 22(1), 89–113. https://doi.org/10.1007/s10805-023-09492-6
[17] Lodge, J. M., Howard, S., Bearman, M. L., Dawson, P., Agostinho, S., Buckingham Shum, S. J., Deneen, C., Ellis, C., Fawns, T., Gniel, H., Harper, R., Henderson, M., Liu, D. Y. T., Markauskaite, L., McLean, J., Perrotta, C., Schuwirth, L., Slade, C., & Associates. (2023). Assessment reform for the age of artificial intelligence. Tertiary Education Quality and Standards Agency.
[18] Dignum, V. (2019). Responsible artificial intelligence: How to develop and use AI in a responsible way. Springer. https://doi.org/10.1007/978-3-030-30371-6
[19] UNESCO. (2021). Recommendation on the ethics of artificial intelligence. UNESCO.
[20] National Institute of Standards and Technology. (2023). Artificial intelligence risk management framework (AI RMF 1.0). U.S. Department of Commerce. https://doi.org/10.6028/NIST.AI.100-1
[21] Jobin, A., Ienca, M., & Vayena, E. (2019). The global landscape of AI ethics guidelines. Nature Machine Intelligence, 1, 389–399. https://doi.org/10.1038/s42256-019-0088-2
[22] Floridi, L., & Cowls, J. (2019). A unified framework of five principles for AI in society. Harvard Data Science Review, 1(1). https://doi.org/10.1162/99608f92.8cd550d1
[23] Jin, Y., Yan, L., Echeverria, V., Gašević, D., & Martinez-Maldonado, R. (2025). Generative AI in higher education: A global perspective of institutional adoption policies and guidelines. Computers and Education: Artificial Intelligence, 8, Article 100348. https://doi.org/10.1016/j.caeai.2024.100348
[24] Xie, Q., Li, M., & Enkhtur, A. (2024). Exploring generative AI policies in higher education: A comparative perspective from China, Japan, Mongolia, and the USA. arXiv. https://doi.org/10.48550/arXiv.2407.08986
[25] Ministry of Science and Technology. (2026). Vietnam has 110.5 million mobile internet subscriptions. https://beta-en.mic.gov.vn/vietnam-has-1105-million-mobile-internet-subscriptions-197260506151607177.htm
[26] Vinuesa, R., Azizpour, H., Leite, I., Balaam, M., Dignum, V., Domisch, S., Fellander, A., Langhans, S. D., Tegmark, M., & Fuso Nerini, F. (2020). The role of artificial intelligence in achieving the Sustainable Development Goals. Nature Communications, 11, Article 233. https://doi.org/10.1038/s41467-019-14108-y
[27] Jabareen, Y. (2009). Building a conceptual framework: Philosophy, definitions, and procedure. International Journal of Qualitative Methods, 8(4), 49–62. https://doi.org/10.1177/160940690900800406
[28] Snyder, H. (2019). Literature review as a research methodology: An overview and guidelines. Journal of Business Research, 104, 333–339. https://doi.org/10.1016/j.jbusres.2019.07.039
[29] Tranfield, D., Denyer, D., & Smart, P. (2003). Towards a methodology for developing evidence-informed management knowledge by means of systematic review. British Journal of Management, 14(3), 207–222. https://doi.org/10.1111/1467-8551.00375