Available on: https://gjefm.ep-journals.org/ E-ISSN: 3117-6771, P-ISSN: 3117-6763

Global Journal of Education, Finance and Management

Volume: 02 Issue: 11 November 2026

Responsible Artificial Intelligence Governance in Vietnamese Higher Education: A Risk-Based Institutional Readiness Framework

Nguyen Van Than
Faculty of Economics, Hanoi Industrial and Trade University

Published Date: November 25, 2026 **DOI: 10.65150/EP-gjefm/V2E11/2026-08**

ABSTRACT: The rapid adoption of artificial intelligence, particularly generative artificial intelligence, is creating new opportunities in higher education while intensifying risks related to data, quality, academic integrity, and accountability. This article addresses two questions: which components constitute institutional readiness, and what governance cycle should universities establish to control relevant risks? The study employs a directed literature review, content analysis, policy analysis, and theoretical synthesis. It proposes a five-component institutional readiness framework and a seven-step risk-based governance cycle centred on learners, educational quality, human rights, and human development. The framework distinguishes technology acquisition from organizational capability and links readiness evidence to risk decisions. The article recommends national guidance, interdisciplinary institutional governance, controlled pilots, impact assessment, and empirical validation of the proposed indicators.

KEYWORDS: Artificial intelligence, AI governance, higher education, institutional readiness, risk governance.

1. INTRODUCTION

Artificial intelligence (AI), particularly generative AI, is increasingly used to support learning materials, feedback, research, and university administration. At the same time, generative AI heightens the risks of inaccurate or fabricated information, inappropriate data processing, privacy violations, bias, and challenges to academic integrity (Miao & Holmes, 2023). A broader review of AI in higher education also calls for greater attention to ethics, collaboration, and rigour in AI research and implementation (Bond et al., 2024). The focus should therefore shift from managing individual tools to developing institution-wide AI governance capabilities.

In Vietnam, the Personal Data Protection Law, the Artificial Intelligence Law, and national digital transformation policies have established an enabling foundation (National Assembly of Vietnam, 2025a, 2025b; Prime Minister of Vietnam, 2022). The remaining gap lies in internal institutional frameworks that translate these requirements into procedures for selecting, piloting, evaluating, scaling up, or discontinuing AI applications.

This article has three objectives: to systematize the theoretical and legal foundations; to identify the components of institutional readiness; and to propose a risk-based assessment framework and governance cycle. It addresses two research questions: (1) What components constitute readiness for responsible AI implementation? (2) What governance cycle should higher education institutions establish to control risks related to data, ethics, quality, academic integrity, and accountability? Its contribution is to connect organizational capabilities to risk-based decisions rather than equating the acquisition of AI tools with readiness.

2. RESEARCH CONTENT

2.1. Research methodology

This is a conceptual study employing a directed literature review, content analysis, policy and legal analysis, comparison, synthesis, and theoretical generalization. No surveys, interviews, or experiments were conducted; the framework and indicators are conceptual proposals rather than validated measurement scales. Searches were conducted on 3–4 August 2026 using the terms ‘Responsible AI’, ‘AI governance’, ‘institutional AI readiness’, ‘AI risk management’, ‘university AI policy’, ‘AI literacy’, ‘academic integrity and generative AI’, ‘data governance in higher education’, and the Vietnamese equivalents of ‘AI governance’ and ‘AI implementation readiness’. Sources included Google Scholar, Crossref, ERIC, the official websites of UNESCO, NIST and the European Union, academic publishers, and Vietnam's official legal-document portal. Bibliographic details and document identifiers were checked against official sources. Inclusion required direct relevance to AI governance, ethics, risk, AI competence, higher education, or Vietnamese law, with preference given to peer-reviewed studies and official documents. Blogs, advertisements, mismatched DOIs, and inadequately substantiated sources were excluded. Content was coded into five categories—principles, risks, organizational capabilities, stakeholder roles, and control processes—and compared with Vietnamese law to develop the five-component framework and seven-step cycle. No PRISMA procedure or screening counts were reported, and no empirical validation has been undertaken.

License: This is an open access article under the CC BY 4.0 license: https://creativecommons.org/licenses/by/4.0/

2.2. Theoretical foundations of responsible AI governance

AI governance refers to the system of rules, responsibilities, procedures, and evidence used to guide, control, and account for the deployment of AI throughout its lifecycle. Responsible AI takes human rights, educational objectives, safety, fairness, transparency, and accountability as guiding standards. NIST organizes AI risk management around four functions—GOVERN, MAP, MEASURE, and MANAGE—with governance maintained throughout the risk management process. UNESCO emphasizes a human-centred approach and the evaluation of ethical and pedagogical suitability, while European Union guidance focuses on the ethical and responsible use of AI and data in teaching and learning (European Commission, 2026; Miao & Holmes, 2023; Tabassi, 2023).

Institutional readiness is the actual capacity to use AI for clearly defined purposes in a safe and verifiable manner; it is not synonymous with owning technology. A university may have numerous software licences yet remain unprepared if it lacks quality data, internal rules, verification skills, risk classification, or accountable decision-makers. Conversely, an institution with limited resources may undertake a small-scale pilot if its objectives, data arrangements, oversight responsibilities, and termination criteria are clearly specified.

Risk-based governance requires controls proportional to the purpose, data, level of automation, scale, affected groups, reversibility, and consequences of error. A tool that suggests practice questions should not be governed in the same way as a system supporting grading or student ranking. Transparency enables users to know when AI is involved, what data are processed, and what limitations apply. Fairness requires examining differential adverse outcomes across groups; explainability enables authorized personnel to challenge recommendations; and safety covers security, resilience, and incident response. Data governance forms the foundation, AI literacy enables verification, pedagogical innovation must remain aligned with quality assurance, and human decision-makers retain ultimate responsibility for AI-assisted decisions (Chan, 2023; Holmes et al., 2022).

Table 1. Principles of responsible AI governance in higher education

Principle	Core content	Risk to be controlled	Governance implications
Human-centred approach	Support learning and human development	Technical optimization overrides educational aims	Assess impacts on learners
Transparency	Disclose AI use, data and limitations	Learners are unaware of AI influence	Notify, label and describe limitations
Fairness	Reduce disparities across groups	Bias and discrimination	Test by subgroup; offer appeals
Data protection	Purpose limitation, minimization and security	Leaks or unauthorized use	Classify data and control access
Safety	Stable and resilient systems	Errors, attacks, harmful outputs	Testing, logging, incident response
Accountability	Identify responsible actors	Blaming the algorithm	Responsibility matrix; preserve evidence
Human oversight	Power to intervene and override	Automated high-stakes decisions	Human checkpoints and right to review
Academic integrity	Disclose AI contributions	Misconduct, fabricated sources/authorship	Disclosure rules; redesign assessment
Monitoring and correction	Traceable, correctable and stoppable	Undetected or unresolved failures	Logs, periodic review, shutdown mechanism

Source: Compiled by the author.

2.3. Policy and legal foundations in Vietnam

Decision No. 131/QD-TTg of 25 January 2022 approved the project to strengthen information technology applications and digital transformation in education and training for 2022–2025, with an orientation to 2030. It provides strategic direction on platforms, data, governance, and digital capabilities (Prime Minister of Vietnam, 2022), but it does not replace institution-specific procedures for individual AI use cases.

Personal Data Protection Law No. 91/2025/QH15 was promulgated on 26 June 2025 and took effect on 1 January 2026. It provides a legal framework governing the rights of data subjects and the obligations of entities processing, protecting, and transferring personal data. In higher education, these requirements imply the need to specify processing purposes, data categories, access permissions, third-party involvement, retention periods, and mechanisms to protect data-subject rights when educational records, personal files, or research materials are processed through AI tools (National Assembly of Vietnam, 2025a).

Artificial Intelligence Law No. 134/2025/QH15 was promulgated on 10 December 2025 and entered into force on 1 March 2026. It establishes a human-centred, risk-based AI governance framework. In the higher education context, the author applies this approach to recommend stricter controls for AI applications that could substantially affect admissions, grades, disciplinary outcomes, scholarships, or other significant student interests (National Assembly of Vietnam, 2025b).

Vietnam thus has a general legal foundation for AI and personal data. The outstanding gap is institutional: universities must translate legal principles into acceptable-use policies, tool inventories, impact assessment dossiers, validation procedures, incident response processes, and decisions to discontinue systems.

License: This is an open access article under the CC BY 4.0 license: https://creativecommons.org/licenses/by/4.0/

2.4. Risks associated with AI use

Data risks arise when personal information, educational records, or research content are entered into tools without clarity about purposes, processing rights, or storage locations. Controls should cover data classification, permissible inputs, role-based access, impact assessments, and contractual provisions. Incomplete, outdated, or historically unequal training data can distort predictions and perpetuate disadvantage. Universities should review provenance, representativeness, error rates, and outcomes across groups; predictions must not be the sole basis for decisions affecting students. Generative AI can produce inaccurate or fabricated information, including citations to nonexistent studies. UNESCO also warns of inappropriate data processing, privacy breaches, and bias. Source verification and human responsibility for output accuracy are therefore essential; internal policies may require disclosure of AI contributions to student work (Chan, 2023; Miao & Holmes, 2023). Insufficient explainability weakens the right to challenge decisions. For grades, scholarships, admissions, or student grouping, learners should be informed and able to request human review. Unequal access reflects differences in devices, subscription costs, language, accessibility for students with disabilities, and digital competence; institutions should provide equivalent alternatives and appropriate support.

Academic-integrity risks include concealing AI assistance, fabricating data or references, and delegating assessed work entirely to AI. Reliance on AI-content detectors alone is inappropriate because evidence of their effectiveness remains limited. Policies should distinguish permitted uses, disclosure requirements, and prohibited practices, while redesigning assessment to measure understanding, critical thinking, and analytical ability; oral examinations or controlled assessment settings may complement this approach (Chan, 2023; Miao & Holmes, 2023). Vendor dependence may lock data and processes into a proprietary ecosystem. Contracts should provide for data export, notification of changes, confidentiality, audit rights, and exit plans. AI must not undermine educators' professional autonomy: evaluation criteria and academic decisions remain the responsibility of authorized people.

Accountability risks occur when ownership of business processes, data, technology, and approval authority is unclear. Each use case requires a responsibility matrix, a reporting contact, and a shutdown mechanism. Even a technically effective tool may be inappropriate for its educational purpose. Validation must begin with the problem to be solved; choosing not to deploy AI is reasonable when expected benefits do not outweigh risks.

2.5. Components of institutional readiness

The proposed framework comprises five mutually related components, to be assessed through verifiable evidence rather than declarations.

Table 2. Suggested criteria for assessing institutional readiness for AI deployment

Component	Suggested indicators	Evidence to collect	Risk if unmet	Responsible unit
Leadership and strategy	Goals, priorities, approval, resources	Strategy, policies, decisions, minutes	Fragmented investments; unclear ownership	University council; executive board
Technology, data and safety	Inventory, access rights, logs, tests	Data maps, security records, contracts	Leakage, poor data, vendor lock-in	IT; data owners; legal affairs
Human capabilities	Competency standards, training, support, verification	Curricula, assessments, support records	Misuse, dependence, undetected errors	HR; faculties; support units
Pedagogy and quality assurance	Learning outcomes, criteria, alternatives	Syllabi, activities, feedback, reports	Low value; reduced learner autonomy	Faculties; assessment; QA
Ethics, law and accountability	Notice, impact assessment, appeals	Policies, forms, logs, incident reports	Rights violations; opacity	Legal affairs; internal control; leadership

Source: Proposed by the author based on the literature synthesis.

Leadership and strategy encompass vision, objectives, policies, resources, responsibilities, and participation. Suggested indicators include a prioritized use-case portfolio, approval mechanisms, lifecycle budgets, an interdisciplinary committee, and student consultation. The university council and executive management bear primary responsibility.

Technology, data, and safety encompass infrastructure, integration, data quality, access rights, cybersecurity, retention, vendor controls, and traceability. Essential evidence includes data inventories, system logs, access controls, incident-response plans, and vendor exit arrangements. Information technology teams should coordinate with data owners and legal staff.

Human capabilities cover academic staff, students, and administrators. Faculty members design learning activities and validate outputs; students verify sources and disclose AI use; administrators understand risks, contracts, and evidence. Indicators include role-specific competency standards, professional development, and on-site support. Ng et al. (2021) conceptualize AI literacy through four dimensions: knowing and understanding AI; using and applying AI; evaluating and creating with AI; and recognizing ethical issues.

Pedagogy and quality assurance require AI uses to align with learning outcomes, learning activities, and assessment. Evidence includes course outlines, assignments, success criteria, non-AI alternatives, feedback, and quality reports. Faculties, departments, assessment units, and quality assurance teams share responsibility.

Ethics, law, and accountability encompass privacy, transparency, fairness, explainability, disclosure of AI use, complaints, human oversight, and error correction. Evidence consists of policies, impact assessments, logs, and incident reports. Legal counsel, internal control staff, and institutional leaders should coordinate; ultimate responsibility cannot be delegated to an algorithm.

License: This is an open access article under the CC BY 4.0 license: https://creativecommons.org/licenses/by/4.0/



Figure 1. Institutional readiness framework for responsible AI deployment
Source: Model proposed by the author based on the literature synthesis.

Figure 1 demonstrates that no component is sufficient on its own. Leadership sets priorities; technology and data provide the foundation; human capabilities support verification; pedagogy establishes educational value; and ethics and law establish boundaries and accountability. A critical weakness in any one component can increase risks across the entire system.

2.6. Risk-Based AI governance framework

The proposed governance cycle consists of seven steps and should be applied to each use case, since a single AI tool may carry different levels of risk in different contexts.

Step 1 – Identify the need. The proposing unit describes the problem, objectives, beneficiaries, non-AI alternatives, and success criteria.

Step 2 – Assess readiness. The five components are reviewed using documented evidence to identify gaps, corrective measures, and pilot limitations.

Step 3 – Review and classify risks. An interdisciplinary committee examines the purpose, data, degree of automation, scale, affected groups, explainability, and consequences of error. Applications affecting admissions, grading, discipline, or scholarships warrant stricter controls.

Step 4 – Conduct a controlled pilot. Scope, duration, datasets, access rights, participants, and stop criteria are specified in advance; incidents must be documented.

Step 5 – Assess impacts. Review output quality, pedagogical value, fairness, privacy, safety, accessibility, workload, and academic integrity.

Step 6 – Decide whether to scale up, revise, or discontinue. Expansion is justified only when benefits are evident, controls are feasible, and responsibilities are assigned. Deployment should cease when risks are unacceptable or objectives remain unmet.

Step 7 – Monitor and improve continuously. Track incidents, complaints, group-level disparities, model version changes, effectiveness, and costs; findings feed back into the identification of needs.



Figure 2. Risk-Based AI governance cycle in higher education institutions
Source: Model proposed by the author based on the literature synthesis.

The university council oversees strategic risks; executive leadership approves deployment; IT departments review technical aspects; assessment and quality assurance units evaluate impacts; faculties, departments, and lecturers retain academic responsibility; students provide feedback and exercise their rights; legal and internal-control units review compliance; and vendors support audits and incident resolution. Decisions that materially affect learners remain the responsibility of authorized institutional personnel.

2.7. Discussion

The first research question is addressed through five components: leadership and strategy; technology, data, and safety; human capabilities; pedagogy and quality assurance; and ethics, law, and accountability. Like Chan (2023), the framework connects pedagogy, governance, and operations, while adding indicators, evidence, risks, and responsible organizational units. It also integrates NIST risk management, UNESCO's

License: This is an open access article under the CC BY 4.0 license: https://creativecommons.org/licenses/by/4.0/

human-centred approach, and Vietnamese legal requirements (Miao & Holmes, 2023; National Assembly of Vietnam, 2025a, 2025b; Tabassi, 2023).

The second question is addressed by the seven-step cycle, which places readiness assessment before piloting and returns monitoring results to needs identification. Scaling up is not automatic; modification or termination are legitimate outcomes. The key distinction is between deploying tools and building organizational capability. A licence, a single training course, or a general policy does not establish readiness without evidence of data quality, pedagogical value, rights protection, and accountability.

The framework's theoretical contribution is to connect institutional readiness to risk-based decisions; its practical value is to help universities begin with bounded applications. Lecturers continue to define objectives, verify outputs, and evaluate learning, while students are protected and involved through feedback, information rights, and opportunities to request review.

The framework has not been empirically validated. Its indicators should be developed into measurement scales, piloted across different types of institutions, and assessed for reliability and construct validity. As technology and regulation continue to change, the criteria will require periodic updates.

3. CONCLUSION AND RECOMMENDATIONS

Readiness for responsible AI implementation consists of five components: leadership and strategy; technology, data, and safety; human capabilities; pedagogy and quality assurance; and ethics, law, and accountability. These dimensions should be assessed through evidence and considered as interdependent.

The governance cycle comprises seven steps: identifying the need; assessing readiness; reviewing and classifying risks; conducting controlled pilots; assessing impacts; deciding to scale up, modify, or discontinue; and continuously monitoring and improving. It ties decisions to evidence, proportionate controls, and ultimate human accountability. Although not empirically validated, the framework can serve as a preliminary set of assessment criteria.

Government authorities should issue sector-specific guidance on AI governance in higher education, clarifying risk classification, data protection, AI-assisted assessment, and the right to human review. They should develop role-specific AI competency standards and support institutions with limited technological resources.

Higher education institutions should establish interdisciplinary governance, adopt AI-use policies, inventory tools and data flows, classify risks, train users, conduct controlled pilots, evaluate impacts before scaling, and create complaint channels and incident response plans. Contracts must address data handling, confidentiality, audit access, data portability, and service termination. AI outputs must never serve as the sole basis for high-stakes decisions affecting learners.

REFERENCES

1) Bond, M., Khosravi, H., De Laat, M., Bergdahl, N., Negrea, V., Oxley, E., Pham, P., Chong, S. W., & Siemens, G. (2024). A meta systematic review of artificial intelligence in higher education: A call for increased ethics, collaboration, and rigour. International Journal of Educational Technology in Higher Education, 21, Article 4. https://doi.org/10.1186/s41239-023-00436-z
2) Chan, C. K. Y. (2023). A comprehensive AI policy education framework for university teaching and learning. International Journal of Educational Technology in Higher Education, 20, Article 38. https://doi.org/10.1186/s41239-023-00408-3
3) European Commission, Directorate-General for Education, Youth, Sport and Culture. (2026). Guidelines on the ethical use of artificial intelligence and data in teaching and learning for educators. Publications Office of the European Union. https://doi.org/10.2766/7967834
4) Holmes, W., Porayska-Pomsta, K., Holstein, K., Sutherland, E., Baker, T., Buckingham Shum, S., Santos, O. C., Rodrigo, M. T., Cukurova, M., Bittencourt, I. I., & Koedinger, K. R. (2022). Ethics of AI in education: Towards a community-wide framework. International Journal of Artificial Intelligence in Education, 32(3), 504–526. https://doi.org/10.1007/s40593-021-00239-1
5) Miao, F., & Holmes, W. (2023). Guidance for generative AI in education and research. UNESCO. https://unesdoc.unesco.org/ark:/48223/pf0000386693
6) Ng, D. T. K., Leung, J. K. L., Chu, S. K. W., & Qiao, M. S. (2021). Conceptualizing AI literacy: An exploratory review. Computers and Education: Artificial Intelligence, 2, Article 100041. https://doi.org/10.1016/j.caeai.2021.100041
7) National Assembly of Vietnam. (2025a, June 26). Personal Data Protection Law (Law No. 91/2025/QH15). https://vanban.chinhphu.vn/?classid=1&docid=214590&pageid=27160&typegroupid=3
8) National Assembly of Vietnam. (2025b, December 10). Artificial Intelligence Law (Law No. 134/2025/QH15). https://vanban.chinhphu.vn/?classid=1&docid=216334&pageid=27160&typegroupid=3
9) Tabassi, E. (2023). Artificial intelligence risk management framework (AI RMF 1.0) (NIST AI 100-1). National Institute of Standards and Technology. https://doi.org/10.6028/NIST.AI.100-1
10) Prime Minister of Vietnam. (2022, January 25). Decision No. 131/QD-TTg approving the project to strengthen information technology applications and digital transformation in education and training during 2022–2025, with an orientation to 2030. https://vanban.chinhphu.vn/default.aspx?docid=205236&pageid=27160

License: This is an open access article under the CC BY 4.0 license: https://creativecommons.org/licenses/by/4.0/

Corresponding Author: Nguyen Van Than