

Artificial Intelligence Regulatory Update 2026

The enactment of the Law No. 134/2025/QH15 on Artificial Intelligence (“**AI Law**”), which took effect on 1 March 2026, is anchored in Resolution No. 57-NQ/TW dated 22 December 2024 of the Politburo of the Communist Party of Vietnam. Resolution 57 identifies science, technology, and digital transformation as the “golden key” for Vietnam to escape the middle-income trap. Under this directive, AI is prioritised as a core engine of growth, with the explicit goal of positioning Vietnam as one of the Top 3 AI Research and Development (R&D) hubs in Southeast Asia by 2030.

Since the AI Law’s adoption, the State authorities have moved quickly to put this framework into operation by issuing the first wave of implementing instruments, namely:

- a. Circular No. 05/2026/TT-BKHCN dated 10 March 2026 of the Ministry of Science and Technology (“**MST**”) promulgating the National AI Ethics Framework (“**Circular 05**”);
- b. Decree No. 142/2026/NĐ-CP dated 30 April 2026 of the Government detailing a number of articles of, and measures to implement, the AI Law (“**Decree 142**”);
- c. Decision No. 804/QĐ-TTg dated 6 May 2026 of the Prime Minister promulgating the list of datasets serving AI development in essential sectors (“**Decision 804**”); and
- d. Decision No. 33/2026/QĐ-TTg dated 30 June 2026 of the Prime Minister promulgating the list of high-risk AI systems (“**Decision 33**”).

This regulatory update will focus on the key mechanisms introduced under these instruments, in particular Decree 142 and Decision 33, namely: (i) risk classification and conformity assessment of AI systems, including the newly issued list of high-risk AI systems; (ii) transparency and accountability obligations, covering technical marking, labelling and serious incident reporting; (iii) the regulatory sandbox regime; (iv) AI infrastructure, data and support mechanisms; and (v) the National AI Ethics Framework.

Risk classification and conformity assessment of AI systems

Decree 142 further elaborates the AI Law's risk-based classification of AI systems; the classification does not apply to AI models unless the model is used as a component of a specific AI system. The risk classification is not a one-time exercise. AI systems must be reviewed and, where necessary, re-classified whenever they are modified, integrated with other systems, or their functions or intended use are changed in a manner that may introduce new risks or increase the existing level of risk.

1. Classification of AI systems

(1) High-risk AI system

(a) Classification criteria

Decree 142 sets out detailed criteria for ministries and ministerial-level agencies to apply when proposing AI systems for inclusion in the list of high-risk AI systems issued by the Prime Minister (the “High-Risk List”). Because an AI system is classified as high-risk only if it is included in the High-Risk List, these criteria function as list-building criteria rather than as a stand-alone self-assessment test for distributors. The assessment takes into account the level of impact (i.e., potential harm, degree of autonomy, and level of human involvement in decision-making and implementation), the extent to which the system supports final decision-making, the field of use (in particular the essential sectors identified in Article 6.2 of the AI Law, or fields directly related to public interests), and the scope of users, the scale of the affected population or the degree of connection with critical technical infrastructure.

Notwithstanding the above, an AI system will not be proposed for inclusion in the High-Risk List, and therefore will not be classified as high-risk, solely because it satisfies one or more high-risk indicators, if it:

- only collects, processes, standardises, classifies, translates or improves the quality of data, and does not itself make decisions affecting the lawful rights and interests of organisations or individuals;
- is subject to genuine human oversight, with a person having authority to override the system's output before it takes effect;
- is used solely for internal administrative or governance purposes; or
- serves only an advisory function and its outputs are not used as the sole basis for decision-making.

These exclusions are not absolute: where measurable technical or practical evidence shows that a system presents a serious risk or vulnerability capable of directly harming national defence, national security, or social order and safety, the Prime Minister may nevertheless include it in the High-Risk List.

(b) List of high-risk AI systems

Decision 33 promulgates the High-Risk List, which comprises two component lists: (i) high-risk AI systems; and (ii) high-risk AI systems subject to conformity certification before being put into use.

The two component lists differ by conformity assessment route. For high-risk AI systems not included in the second of those lists (the “**certification sub-list**”), distributors may either self-assess conformity, by preparing a technical dossier and assuming legal responsibility for the assessment result, or engage a conformity assessment body registered or recognised under the legislation on standards and technical regulations. For systems included in the certification sub-list, assessment by such a body is mandatory before use. Foreign distributors supplying high-risk AI systems in Vietnam must also maintain a lawful contact point in Vietnam and, for systems on the certification sub-list, must have either a commercial presence or an authorised representative in Vietnam.

The High-Risk List currently covers 46 AI systems across six sectors: education, healthcare, banking, transport, judicial proceedings, and ethnic and religious affairs, with transport accounting for the largest share (31 systems). Of these, only six AI systems are subject to conformity certification, comprising systems in ethnic and religious affairs (4), healthcare (1), and judicial proceedings (1).

The table below sets out illustrative examples for each of the six regulated sectors.

Sector	Illustrative high-risk AI systems
Education	- Self-study using uncontrolled data sources. - Test, assess and rank learners. - Monitor and analyse learner behaviour using biometric data.
Healthcare	Assist with or perform surgery through an integrated surgical robot, including a robot that directly participates in, guides, or applies treatment during an intervention, pursuant to a medical order and without healthcare professional confirmation at each step.
Banking	- Execute high-value electronic transactions without human control. - Approve or deny prescribed credit without independent approval by a credit officer.
Transport	- Operate vehicles without continuous human control. - Control or automatically operate traffic signals or manage road, railway or air traffic. - Operate or control transport works and critical technical infrastructure, including road tunnels, major bridges, expressways and transport hubs, as well as water supply, drainage and urban lighting systems. - Assess, score, rank, or classify applications for licences, certificates, or operating authorisations without independent review or approval by a competent authority. - Detect violations or perform real-time biometric identification for enforcement purposes.

Sector	Illustrative high-risk AI systems
	Control aviation operations relating to, e.g., unmanned aircraft (drones/UAVs), baggage, passengers, and weather conditions.
Judicial proceedings	Conduct large-scale biometric identification for resolution of public interest civil cases.
Ethnic & religious affairs	- Process applications to determine beneficiaries and areas of ethnic policies. - Verify and decide the validity of ethnicity- and religion-related information. - Approve or reject applications or extend relevant administrative decisions. - Disburse or recover benefits or suspend eligibility for benefits. - Infer or classify ethnicity, religion, or beliefs for the purpose of applying administrative measures. - Allocate or coordinate budgets or public policies.

Decision 33 sets regulatory compliance deadlines for high-risk AI systems that are already operating or newly deployed, as follows:

High-risk AI system	Sector	Deadline
Operating before 15 August 2026 ¹	Healthcare, education, finance	Before 1 September 2027
	Others	Before 1 March 2027
Deployed within six months from 15 August 2026	All	—

¹ 15 August 2026 is the effective date of Decision 33.

These deadlines reflect the transitional timetable in Article 35 of the AI Law for systems already operating when the AI Law took effect on 1 March 2026. During the transitional period, the relevant systems may continue to operate unless the competent State authority determines that a system poses a risk of serious harm and requires its suspension or termination under Article 35.2 of the AI Law.

(2) Medium-risk AI system

In addition to the criteria in the AI Law (e.g., the ability to mislead, influence or manipulate users who are unaware that they are interacting with AI systems or AI-generated outputs), Decree 142 introduces an additional exclusionary criterion for classifying medium-risk AI systems, namely that the system does not fall within the High-Risk List.

The classification of medium-risk AI systems is also subject to a carve-out: an AI system that meets the criteria will nevertheless not be classified as medium-risk where it:

- merely performs technical editing of form or presentation, without generation of new content or alteration of subject identity;
- is a clearly identifiable office productivity tool that does not simulate identities or events;

- does not directly interact with or provide the outputs to the public (including through intermediaries);
- is used in artistic, cinematic, video-game or other creative activities where the context of publication makes the fictional nature of the content apparent; or
- solely processes data or optimises the operation of technical systems, without direct interaction with users, generation of outputs to the public, or direct interaction with the physical environment in the manner of the primary system.

(3) Low-risk AI system

Systems that do not fall within the high-risk or medium-risk categories described above.

2. Automated risk-classification support tool

MST must establish and maintain an automated risk-classification support tool on the National AI One-Stop Portal (“**AI Portal**”) to help distributors and deployers self-assess and classify their AI systems. Use of the tool is optional, and no authority may require its output to be submitted as a mandatory component of any administrative procedure. That said, a classification result generated by the tool is treated as a valid electronic document forming part of the distributor's risk-classification dossier and may be relied upon by a competent authority as one of the bases for assessing compliance with risk-classification requirements. Where the risk level remains unclear after use of the tool, the distributor may request classification guidance from MST based on the system's technical dossier.

3. Risk-classification dossier

A distributor must prepare a risk-classification dossier for each high-risk or medium-risk AI system before the system is put into use. The dossier includes (i) system identification information; (ii) a description of the system and its intended context of use; (iii) a general description of the principal categories of input data used in the operation of the system; and (iv) a summary of the risk management, safety, and transparency measures implemented in relation to the system. Where the distributor develops, deploys, or operates the system using a third-party platform, model, or service, the distributor is only required to provide technical and data-related information that is within its lawful access and control. This dossier is distinct from the technical dossier that a distributor prepares to support a self-assessment of conformity.

The preparation of the dossier does not require the disclosure of source code, model parameters, detailed algorithms, raw training data, or any State, business or technology secrets. Technical documentation prepared in accordance with internationally recognised standards may be used to satisfy this requirement, provided that such documentation contains the information described above.

Where the system processes personal data, a personal data processing impact assessment prepared in accordance with applicable personal data protection laws may be used to replace, or be incorporated into, the risk-classification dossier, with a view to reducing administrative compliance burdens. The distributor must retain the risk-classification dossier for the entire operational life cycle of the system.

4. Publication and notification of results

Decree 142 further specifies the information about the results of high-risk AI system conformity assessment to be published on the AI Portal before the system is put into use, and again whenever an updated re-assessment result is obtained. The information comprises:

- Name, version, and unique identifier of the AI system.
- Details of the distributor.
- Details of conformity assessment, including conformity assessment body, assessment method, conclusion on conformity, date of assessment completion, date of result recognition.
- Date of most recent information update.

A conformity assessment must be conducted again if a significant change is made to the AI system during its use and that change could affect the original assessment outcome. A significant change includes: changes to the system's main functions, intended purpose, or application area that create new risks or change the assessed risk level; changes to the system architecture, AI model, or key technical configuration that may affect accuracy, reliability, safety, or controllability; changes to data sources, key input data, or data processing methods that materially affect system performance or risk levels; integration of the system with another system or deployment in a new operating environment that may generate new risks or alter the evaluated operating conditions; or any other change that materially affects compliance with the requirements of Article 14 of the AI Law on the management of high-risk AI systems.

Routine data updates, bug fixes, performance improvements, and version upgrades are not considered significant changes, provided they do not affect compliance with those requirements.

Conformity assessment results issued by foreign conformity assessment bodies may be recognised by the Vietnamese authorities in accordance with Vietnamese laws on national and technical standards and relevant treaties.

Decree 142 also establishes a non-duplication principle of particular importance for products that incorporate AI components. Where a high-risk AI system is subject to sector-specific legislation, or to laws governing standards, technical regulations, or product and goods quality, any testing, inspection, conformity assessment, or certification results obtained under those frameworks may be relied upon to demonstrate compliance with the corresponding requirements of the AI Law, and competent authorities may not require those matters to be reassessed. Only requirements not covered by the earlier assessment must be assessed separately.

Distributors of medium-risk and high-risk AI systems must electronically notify their risk-classification results via the AI Portal before the system is put into use, either by direct submission through an online form on the AI Portal, or through an application programming interface or other appropriate electronic method. Upon receipt, the AI Portal automatically records the submission, issues a system identifier, and sends an electronic confirmation to the distributor.

5. Risk management obligations for high-risk AI systems

In addition to classification and conformity assessment requirements, Decree 142 requires distributors of high-risk AI systems to establish and maintain a risk management system that reflects the system's intended use, deployment scope, and risk level. At a minimum, the risk management system must:

- identify and assess risks to human rights, safety, security, and public interests arising from the design, development, and supply of the system;
- ensure the quality, suitability, and representativeness of training, testing, and evaluation data;
- provide appropriate human oversight and intervention mechanisms;
- implement technical or organisational measures to prevent, mitigate, or control identified risks; and
- review and update those measures whenever significant changes are made to the system, its data, or its intended use.

The distributor must also provide the deployer with sufficient information to operate the system safely and as intended.

The deployer must use the system only for its intended purpose and within the conditions specified by the distributor. It must monitor the system for risks, incidents, or unexpected behaviour, maintain appropriate human oversight, implement risk-mitigation measures within its control, and cooperate with the distributor in reviewing and updating those measures. If the deployer identifies a risk of serious harm to life, health, human rights, property, cybersecurity, social order and safety, or other public interests, it must immediately take appropriate risk-mitigation measures and notify both the distributor and the competent authority.

Transparency and accountability principles

Distributors and deployers of AI systems must discharge transparency obligations in respect of AI systems and their content. They must also provide explanations upon the written request of competent authorities. In fulfilling these obligations, they are not required to disclose source code, algorithms, training data, model parameters or trade secrets, unless otherwise required by law, and must still comply with regulations on personal data protection.

1. Technical marking of AI-generated content

Distributors must apply a technical solution to mark audio, image, or video outputs in a machine-readable format (text outputs are excluded, unless otherwise required by law). This may be done by embedding an identifier into the file structure, content data, or metadata, using a digital signature or e-signature, or another equivalent technical solution, confirming that the content was AI-created or AI-edited. The marking must be maintained throughout the creation, export, and delivery of the content. Where the AI system is supplied as open source or free of charge, the distributor is treated as having discharged this obligation if it either builds the marking function into the system or publicly releases the tools, standard configurations, application programming interfaces or technical documentation enabling the deployer to configure and operate that function; in that case, a deployer that uses the system to deliver content to the public must itself apply the technical marking.

2. Notification and labelling requirements

Deployers must clearly notify the public when delivering AI-generated or AI-edited content that could cause confusion as to the authenticity of an event, person, or content origin. A label is required where the content (e.g., audio, image, or video) simulates a real person's appearance or voice, or recreates a real event. For cinematic or artistic works, the deployer that publicly releases the content bears the labelling obligation; the label may appear in the opening, the ending, the credits, the description, or an attachment.

Where content is generated or modified by an AI system during its production, the person or organisation responsible for generating or modifying the content must provide the information needed for the person or organisation that publicly releases the content to comply with the labelling requirement. The notification or label must be clear, easy to understand, and easy to notice. It must be provided before, or at the time, the recipient accesses the content. It must not be presented in a way that hides or reduces awareness that the content was generated or modified by AI, and it must not significantly interfere with the display, functionality, or use of the content.

The method of notification or labelling must be appropriate to the type of content and the means by which the content is made available, and may include:

- a notice or label displayed directly on the content or in its title or description;
- a caption accompanying the content;
- a notice through the platform interface;

- an audio notification; or
- other equivalent methods.

That said, the labelling requirement does not apply in the following circumstances:

- Technical edits to audio, image or video solely for enhancing quality, without altering their essential nature or primary context;
- Correction of spelling or grammar, or summarisation, paraphrasing or translation that does not distort the fundamental meaning of the original text;
- Content used solely for internal purposes; and
- Content in the course of research, development or controlled testing that is not made available to the public.

3. Serious AI incident reporting and handling

Decree 142 clarifies the substantive meaning of a serious incident as defined under the AI Law, being an event arising during the operation of an AI system that causes:

- loss of human life or serious harm to human health;
- significant damage to property or serious impact on the operations of an organisation;
- serious infringement of legitimate rights and interests of individuals or organisations;
- serious disruption to provision of public services or essential services; or
- impact on national security, or on social order and safety.

Although the definition is broken down into specific categories, Decree 142 does not quantify the threshold at which an impact or damage becomes “serious”. The determination of seriousness will therefore turn, first, on whether the incident falls within one of the categories above and, thereafter, on a case-by-case analysis of the specific circumstances.

When a serious incident of an AI system occurs, the deployer and user must promptly record the incident, limit its consequences, and notify the distributor for remediation, while the distributor must apply technical measures to prevent and remediate the harm. In addition, a preliminary report must be submitted to the competent authority via the AI Portal within 72 hours of confirmation for time-critical incidents (those in paragraphs (a), (d) and (e) above, together with paragraph (c) cases that cannot be brought under control) or within five working days for other serious incidents. Within 15 calendar days from the submission of the preliminary report, a formal remediation report must be submitted, and system logs must be retained. Decree 142 also provides prescribed forms of serious incident reports for both individuals and organisations.

An incident is considered confirmed when the reporting party has enough initial information to conclude that the incident has occurred and is highly likely to have been caused by a fault in the AI system. It is not necessary to wait for a full technical investigation to determine the exact cause. Decree 142 makes clear that submitting a preliminary report within the required timeframe does not constitute an admission of technical fault or legal liability. If the deployer is unable to contact the distributor, the deployer must submit the report. Where the same incident must also be reported under cybersecurity, personal data protection, or other sector-specific laws, the report must be made in accordance with those laws. In such cases, competent authorities must coordinate and share information, and may not require information that is already available through connected systems or databases.



Regulatory sandbox

Building on the sandbox framework introduced under the AI Law, Decree 142 establishes a detailed regime allowing organisations to test AI systems under regulatory supervision before commercial deployment. The sandbox mechanism does not, however, extend to the research, development or testing of AI systems in simulated or closed environments, or to internal testing involving no real-world participants and producing no impact outside the testing environment.

1. Classification of testing levels

Decree 142 classifies AI systems participating in the sandbox into three testing levels based on factors including the AI system's risk level, the type of data processed, the scale of deployment and its potential impact.

Level 1

covers limited testing where AI outputs are not used in real operations, are not made available to users and are not relied upon in decision-making ("**Level 1**").

Level 2

applies where real users participate and AI outputs are used in operational activities, made available to users or used as a reference in decision-making, but do not directly create, modify or terminate legal rights or obligations ("**Level 2**").

Level 3

applies to higher-risk AI systems, including those processing sensitive personal data, core data, children's data, State secrets or other protected data, or those connected to national critical information infrastructure ("**Level 3**").

2. Entry requirements and approval

Applicants must demonstrate that the proposed AI system involves an innovative technology or deployment model and submit a testing proposal describing the objectives, scope, duration and risk management measures. They must also implement appropriate safeguards for affected persons during the testing process. Where an AI system could directly cause significant personal injury or large-scale property damage, the applicant must maintain liability insurance or an equivalent financial

safeguard. Decree 142 also provides simplified application requirements for start-ups, SMEs, science and technology organisations and research groups applying for Level 1 testing.

Applicants that satisfy the entry requirements will be issued a sandbox participation certificate, which authorises them to conduct testing within the approved scope. The competent authority varies depending on the testing level and deployment scope. Provincial People's Committees oversee Level 1 and Level 2 projects deployed within a single province. Ministries and ministerial-level agencies supervise Level 1 and Level 2 projects falling within their sector of State management that are deployed across two or more provinces, or that are deployed by entities under their direct management, while the Ministry of Public Security (“**MPS**”) is responsible for Level 3 projects. Applications are filed electronically through the National Public Service Portal and are then synchronised to the AI Portal.

3. Ongoing obligations

During the testing period, participants remain subject to ongoing regulatory oversight. They must submit periodic reports every six months for Level 1 and Level 2 projects, and every three months for Level 3 projects. Serious incidents or any testing conducted beyond the approved scope must be reported within 72 hours. Participants must also submit a final report no later than 15 days before the end of the approved testing period. Decree 142 also introduces the relevant prescribed forms for both periodic reports and final reports. All reports must be submitted via the AI Portal. The testing period itself may be extended: the request must be filed at least 15 working days before expiry, together with a final report on the testing carried out, and the authority must appraise and decide within 10 working days.

4. Post-testing period

Within 15 working days of receiving the final report, the competent authority must evaluate the testing results and either issue a sandbox completion certificate or refuse to do so by reasoned written decision. The completion certificate may support subsequent AI conformity assessment and provide the basis for exemptions, reductions or adjustments to certain compliance obligations under the AI Law.

Participants may continue operating the AI system during a transition period of up to 12 months from the date of the completion certificate (extendable once for up to six months), provided the system remains within the approved scope and risk profile. The sandbox completion certificate does not replace any licences, approvals or other regulatory requirements under applicable laws. Any change that exceeds the approved scope, conditions, or limits is treated as an expansion of operations and requires a new approval. If the expansion changes the relevant province, sector, or testing level, approval must be obtained from the authority with jurisdiction over the expanded activity. Once the transition period ends, the participant must comply fully with the AI Law, Decree 142, and any other applicable laws before continuing to provide the product or service or operate the AI system.

5. Liability waiver and participant safeguards

Organisations and individuals participating in an approved regulatory sandbox may benefit from certain exemptions from civil, administrative and criminal liability arising from the testing activities,

provided they comply with the approved testing conditions and applicable legal requirements. These protections do not apply where participants knew, or should have known, of material risks but failed to promptly report them to the competent authority or implement appropriate mitigation measures. However, participants remain liable for damage caused to third parties and must continue to fulfil their obligations to protect test participants during the testing period.

Separately, subject to the scope, duration, and conditions set out in the participation certificate, the competent authority may grant participants certain regulatory flexibilities, including: allowing alternative notification or transparency measures instead of certain risk-classification, marking, or labelling requirements, provided that all testing participants are clearly informed; postponing conformity assessment until after the testing phase, where conformity assessment would otherwise be required; and giving priority access to AI support programmes, resources, and government support policies.

6. Financial support

Participants in the regulatory sandbox may be granted an AI development support voucher (the “**AI Voucher**”), an electronic instrument recording a funding limit drawn from the National AI Development Fund and other lawful sources, covering costs directly attributable to the testing activities. The AI Voucher may be used to cover part of the costs of (i) computing, storage and platform infrastructure; (ii) use of data and AI models; and (iii) relevant services (e.g., technical consulting, system testing, security and risk assessment), capped at 50% of the actual, eligible costs.

The AI Voucher scheme is not limited to sandbox participants. Under Decree 142, AI Vouchers may also be granted to: innovative start-ups and small and medium-sized enterprises (SMEs); science and technology organisations and research groups with a viable innovation project; and other organisations or individuals carrying out an AI research, development, or application project under an approved programme or support scheme. Applications must be submitted through the National Public Service Portal. AI Vouchers are non-transferable, cannot be exchanged for cash, and are paid directly to the service provider rather than to the recipient.

A graphic showing a large, dark, rectangular chip with a large white 'A' on top, connected to a dense array of gold-colored pins or connectors. The background is a dark red gradient.

AI infrastructure, data and support mechanisms

Decree 142 establishes the AI Portal, operated by MST, as the central platform for AI-related regulatory administration. The AI Portal will facilitate the submission and management of AI-related notifications, reports and conformity assessment results, assign identification numbers to AI systems, and publish information on AI infrastructure, technical guidance and government support programmes.

Importantly, Decree 142 includes a transitional arrangement for the period before the AI Portal becomes operational. Until MST officially announces that the AI Portal is in operation, all notifications and reports required under the Decree may be submitted through electronic systems, online public services, or other channels designated by MST. Submissions made through these channels have the same legal effect as submissions made through the AI Portal. During this period, any information that must be disclosed under the Decree should be published on the official website or portal of the authority or organisation responsible for making the disclosure.

The Decree further provides for the development of a national AI infrastructure under a model combining State and private investment. The framework aims to provide shared computing resources, AI platforms and other technical resources for AI research, development and deployment, while promoting interoperability and resource sharing. Institutional responsibility is split: MST leads the components serving research, development and innovation, while MPS leads the infrastructure serving State management, deployed centrally at the National Data Centre, and also coordinates the national AI infrastructure network as a whole together with MST.

In addition, Decree 142 establishes a framework for AI-serving databases, comprising open data, conditionally open data and commercial data. The Decree sets out the principles for the management, sharing and use of such data and requires compliance with applicable laws. To promote AI development, Decree 142 also introduces non-financial support measures for organisations and individuals that voluntarily contribute data, AI models or research outputs, including priority access to shared AI infrastructure and technical resources.

The national AI database is managed by MPS through the National Data Centre. Decision 804 provides the first detailed framework for this database.

For the period 2026–2030, the Decision identifies 15 groups of datasets to support AI development in essential sectors, and a lead ministry or agency is assigned to each dataset group. The Decision also identifies 39 priority data categories, covering: development of Vietnamese large language models (LLMs); computer vision; AI system testing and evaluation; AI applications in essential sectors; and AI safety and trustworthiness assessment. MPS is responsible for overseeing implementation of the framework and proposing updates to the dataset list when necessary.



National **AI Ethics** Framework

Circular 05 promulgates the National AI Ethics Framework (“**AI Ethics Framework**”), which directly applies to State authorities and to organisations and individuals acting as developers, distributors, deployers or users of AI systems serving State management or public services. Other Vietnamese and foreign organisations and individuals researching, developing, supplying, deploying or using AI systems in Vietnam are encouraged, but not required, to apply the AI Ethics Framework; in practice it serves as the reference basis for enterprises to build internal AI ethics frameworks and to self-assess their compliance with AI ethics principles.

The AI Ethics Framework rests on four groups of principles as follows:

Safety, reliability and security

ensuring the safety, reliability and non-harmfulness of AI systems, with testing mechanisms, human oversight and system security measures;

Human rights, fairness and transparency

respecting human rights, ensuring fairness and transparency, mitigating data bias, protecting privacy and establishing clear accountability;

Social benefit and sustainable development

orienting AI towards social benefit and sustainable development, ensuring that AI systems deliver value to the community, minimise negative impacts and align with cultural norms; and

Innovation and social responsibility

encouraging innovation coupled with social responsibility through training, cooperation and the responsible development of AI.

Circular 05 also includes an appendix providing practical guidance for different stakeholder groups, including AI developers, distributors, deployers, users, State authorities and public service providers.

The AI Ethics Framework is to be reviewed and updated every three years, or upon a major change in technology, law or management practice, and MST is required to issue governance and compliance-assessment guidance and to build a digital platform and automated self-assessment tool to assist ministries, localities and enterprises.

Authors



Phong Nguyen

Partner
phong.nguyen@kinstellarsea.com



Cao Pham

Associate
cao.pham@kinstellarsea.com



Khanh Linh Nguyen

Associate
khanhlinh.nguyen@kinstellarsea.com



Ngan Le

Associate
ngan.le@kinstellarsea.com

About Kinstellar Southeast Asia (Vietnam)

Kinstellar Southeast Asia, in cooperation with Asia Counsel Vietnam Law Company Limited, is a leading law firm based in Ho Chi Minh City with over a decade of experience advising international and Vietnamese investors. Our team of more than 30 legal operatives combines deep local insight with international standards to deliver precise, business-focused legal advice.

In our corporate and regulatory practice, we advise investors, financial institutions and technology companies on digital assets, fintech, and emerging regulatory frameworks in Vietnam, from licensing and market entry to compliance, tax and dispute resolution. We also advise on the digital asset aspects of M&A and fintech partnerships, including cross-border transactions and engagement with relevant regulators.

Our clients span multinational corporates, funds, financial institutions and Vietnamese enterprises across a range of sectors.

Kinstellar Southeast Asia (Vietnam)

www.kinstellarsea.com

Level 9, Deutsches Haus, 33 Le Duan Boulevard, Sai Gon Ward, Ho Chi Minh City