VNU Journal of Science: Legal Studies

Journal homepage: https://js.vnu.edu.vn/LS

Original Article

A Risk-Based Approach to Artificial Intelligence Regulation: Lessons from the European Union for Vietnam

Nguyen Van Duong*, Dao Gia Phuc

University of Economics and Law, Vietnam National University, Ho Chi Minh City, No.669 Do Muoi, Quarter 13, Linh Xuan Ward, Ho Chi Minh City, Vietnam

Received 8th July 2025
Revised 17th September 2025; Accepted 7th October 2025

Abstract: This paper examines the risk-based regulatory approach adopted in the European Union's Artificial Intelligence Act. It deciphers the theoretical underpinnings of this approach, including the legal conceptualisation of "risk," its normative foundation, and both the strengths and limitations inherent in its legislative design. Furthermore, the author analyses and assesses Vietnam's legal framework governing artificial intelligence, with particular emphasis on the Draft Law on Artificial Intelligence of Vietnam and the Law on Digital Technology Industry 2025. Drawing on comparative analysis and international best practices, the paper offers recommendations for the development of AI-related regulatory policies in Vietnam.

Keywords: Artificial Intelligence Act, Law on Digital Technology Industry, risk-based approach, European Union, Vietnam.

1. Introduction

In both global and Vietnamese contexts, artificial intelligence (AI) is experiencing rapid expansion and is gradually reshaping society in a positive direction [1]. However, AI also entails certain risks throughout its deployment and application. AI systems are increasingly present in everyday life, ranging from virtual assistants and autonomous vehicles to medical diagnostics, thereby raising concerns over potential negative impacts such as algorithmic bias, privacy infringements and threats to public safety [2]. These developments call for an effective legal framework capable of managing risks while preserving space for technological innovation.

The European Union (EU) has become the first jurisdiction to enact a regulatory instrument specifically designed to govern AI through the European Union's Artificial Intelligence Act

* Corresponding author.
E-mail address: duongnv@uel.edu.vn

https://doi.org/10.25073/2588-1167/vnuls.4799

(AIA). The EU has adopted a risk-based approach to AI governance, under which AI systems are categorised according to the level of risk they pose, with corresponding legal obligations applied to each category [3]. This approach has not only shaped the architecture of AI governance within the EU but has also had a global impact, compelling non-EU companies to comply when supplying AI systems to the EU market.

Given the growing relevance of AI technologies, it is essential for Vietnam to develop a coherent legal framework to address AI-related challenges. At present, Vietnam does not have a dedicated law regulating AI. Instead, relevant provisions are found in existing legal instruments such as the Law on Cybersecurity 2018, the Decree 13 on Personal Data Protection and other related regulations. The absence of a specific legal framework for AI results in a lack of fundamental risk governance structures when AI is applied in practice in Vietnam. Therefore, in the process of developing a legal framework for AI in Vietnam, especially as the Law on Digital Technology Industry 2025 (LDTI 2025) has recently been promulgated, it is necessary to learn from international experience, in particular from the EU, which is widely recognised for its comprehensive and advanced regulatory approach to AI through the adoption of specialised legislation [4]. Furthermore, the AIA's method, along with its human-centred philosophy, can serve as an important model for Vietnam to consider, given the similarities in policy orientation.

Furthermore, the EU's human-centric approach to AI, as embodied in the AIA, may serve as a meaningful model for Vietnam, given certain policy convergence. This paper focuses on analysing the risk-based approach adopted in the AIA and the normative foundations underpinning it. It also examines the practical and conceptual challenges the EU has faced in drafting this regulation. Based on this analysis, the paper offers recommendations for Vietnam to develop a legal framework suited to its national context, capable of balancing innovation promotion with the need to control emerging risks. A comparative legal methodology is employed to assess the EU model in relation to the Vietnamese legal landscape and to extract insights relevant to domestic regulatory development.

2. Theoretical Considerations on Risk and the Risk-Based Approach in Artificial Intelligence

Risk is inherently associated with adverse consequences or harm. Accordingly, risk in the context of AI refers to the potential for negative or harmful outcomes arising from the practical use of AI. What distinguishes AI-related risks is their multidimensional nature [5]. On the technical side, AI systems may malfunction due to faults in machine learning processes or the use of incomplete or biased training data. These technical flaws can lead to erroneous outputs, flawed reasoning, or incorrect decisions, thereby causing harm when applied [6]. From an ethical and societal perspective, AI has the potential to amplify existing social prejudices and discriminatory practices, thereby violating core principles such as fairness and human rights. This may include racial or gender-based discrimination or the unauthorised processing of personal data [7]. At the institutional level, AI systems may infringe upon fundamental rights, such as the right to privacy, the right to a fair trial, and the right to explanation, rights that are legally recognised and protected.

Moreover, in cases where AI systems cause damage, legal liability remains ambiguous, particularly when it is difficult to determine whether the fault lies with the developer, provider, or end-user. As noted by the United States Federal Trade Commission, AI is best understood as a socio-technical system, meaning that the risks it presents emerge from the interaction between technological components and social or human contexts. In this light, AI risks arise not only from the technical design of algorithms but also from their deployment environment and interaction with human behaviour [8]. Consequently, AI inherently

embodies both direct (technical) risks and indirect (ethical, institutional, and societal) risks.

According to the EU's regulatory perspective, risk is defined as the probability of harm multiplied by its severity.[1] This approach reflects an understanding of AI risk that is not limited to the likelihood of system failure but also encompasses the concrete consequences such failures may have on legally protected rights and interests. This definition underpins a dual-dimensional risk assessment model that combines quantitative (probability of occurrence) and qualitative (severity of harm) elements. It provides the foundation for allocating corresponding legal obligations based on varying levels of risk. This framework serves as the basis for the risk-based classification of AI systems, which is one of the core structural pillars of the AIA.

When approaching the notion of risk from a regulatory perspective, it is essential to draw a clear distinction between technical risk and legal risk in order to identify the object of regulation correctly and to select appropriate regulatory tools. Technical risk primarily falls within the domain of technical standards and refers to the possibility that an AI system may malfunction due to technological issues. Such risks may include incorrect outputs resulting from programming errors, algorithmic inaccuracies, or incomplete training data. These risks are often quantifiable to a certain extent, using indicators such as error probabilities, and potential damage to property, health, or human life.

In contrast, legal risk is qualitative in nature and arises when the operation of an AI system results in consequences that contravene existing legal norms or infringe upon the legitimate rights and interests of individuals, organisations, or communities. Legal risk falls squarely within the scope of legal regulation and requires mechanisms for control, responsibility allocation, and protection of affected parties' rights and interests.

Distinguishing accurately between these two categories of risk is fundamental to constructing a risk-based legal framework for AI regulation. Under this model, legal regulation should focus on governing and sanctioning legal risks, while technical risks, although not subject to direct legal regulation, should nonetheless be managed through technical instruments such as standardisation, testing, and certification. However, to ensure comprehensive and effective AI governance, a coordinated framework is required that harmonises legal and technical regulatory approaches, enabling the identification and mitigation of all types of risk across the entire lifecycle of AI systems.

Traditional models of legal liability, which rely heavily on human conduct, encounter difficulties when applied to AI-generated legal events where decisions are made with minimal or no human intervention. Existing legal frameworks may struggle to keep pace with the rapid development of AI and robotics, and the absence of standardised rules can lead to fragmented and inconsistent regulation [9]. In practice, assigning liability when AI causes harm is extremely complex, in part due to the opacity and "black box" nature of AI systems.[2] [10] The vast volume of personal data processed by AI poses significant risks, including indiscriminate data collection and algorithmic profiling that infringe privacy rights. AI algorithms may also replicate and amplify biases in training datasets, leading to discriminatory outcomes. Furthermore, the lack of transparency in AI decision-making processes complicates efforts to verify compliance with legal standards and detect violations.

Accordingly, AI governance becomes imperative and should be premised on anticipating risks that may materialise in real-world applications. Nonetheless, the risk governance approach must be carefully balanced with the need to promote innovation in AI and robotics. Any risk-based regulatory framework

[1] Article 3(2) of AIA.

[2] The term "black box effect" in the context of AI refers to a situation in which the decision-making process of an AI system becomes opaque, difficult to understand, or hard to explain.

should integrate ethical considerations into AI development, with an emphasis on enhancing transparency and accountability among all actors across the AI lifecycle [11, 12].

Ultimately, the risk-based approach aims to harness the developmental potential of AI while safeguarding fundamental societal values and rights. In other words, it is a strategy to manage the inherent risks of AI in a manner that maximises its benefits and mitigates its adverse impacts. Furthermore, this approach is widely regarded as striking a balance between risk control and technological advancement. It avoids unnecessarily constraining innovation in low-risk areas while concentrating regulatory resources on AI applications that pose the most significant risks. Moreover, this approach ensures that regulatory controls are proportionate to the level of risk involved, avoiding both overregulation and underregulation. Importantly, it is also technically and technologically feasible to implement [13].

At present, the AIA represents the most explicit embodiment of the risk-based approach. The Act classifies AI systems according to their level of risk, namely, unacceptable risk, high risk, limited risk and minimal risk. These risk levels determine the corresponding legal and regulatory requirements, thereby allowing regulatory efforts to be concentrated on AI applications that pose the greatest risks [14]. The EU's approach to AI risk is derived in part from the regulatory model applied to medical devices in Europe, particularly the principle of AFAP (As Far As Possible).[3] However, the EU interprets the AFAP standard in a narrower sense by removing the criterion that risk should be reduced "as far as possible" without adversely affecting the risk–benefit ratio. In effect, AI developers are required to eliminate risks absolutely, regardless of cost or the potential impact on the utility or performance of the technology. Such an interpretation risks imposing a rigid and overly restrictive application of the AFAP principle. It fails to consider opportunity costs or the implications for technological innovation [15]. Scholars have observed that the EU's risk approach may be excessively stringent, overlooking the need to balance benefits and risks. This, in turn, may increase the compliance burden on businesses and hinder the real-world deployment and development of AI technologies.

3. The Risk-Based Approach in the Legal Framework of the European Union

In the AIA adopted by the European Parliament and the Council in 2024. From the time it was proposed, the AIA by the European Commission has been positioned as a "risk-based" framework aimed at building "trustworthy AI". Accordingly, the AIA establishes a classification system of four levels of AI risk, thereby prescribing a corresponding legal regime for each level. Specifically, these include:

i) Unacceptable risk: This is the highest level of risk, applied to AI systems considered too dangerous or unethical to the extent that they are absolutely prohibited. Article 5 of the AIA imposes a complete prohibition on the use of AI systems that pose an unacceptable risk. The AIA sets out a list of behaviours falling under this category, such as AI that uses subliminal techniques to manipulate human psychology; AI that exploits the vulnerabilities of vulnerable groups (children, persons with disabilities) to cause harm; AI used by public authorities for social scoring of citizens resulting in discriminatory consequences; AI that predicts the likelihood of an individual committing a

[3] The principle of "As Far As Possible" promotes diligence, responsibility, and adaptability in medical practice. It requires healthcare professionals to continuously assess the situation, consider all relevant factors, and make their utmost effort to achieve the best possible outcomes for patients, even when faced with unavoidable challenges and limitations. This principle serves as a reminder that in medicine, the constant pursuit of improvement and the optimisation of care is both an ethical and professional imperative.

crime (a form of "predictive policing"); and the extraction of biometric data remotely in real-time in public spaces for law enforcement purposes (except for certain exceptions such as searching for missing children or preventing terrorism). It can be seen that these are behaviours which gravely and directly violate fundamental values of human beings and society (such as human dignity and human rights). Therefore the EU has adopted an absolute prohibition approach.

ii) High risk: This group includes AI systems that are not prohibited but pose a significant risk to human health, safety, or fundamental rights. Article 6, Chapter III and Annex III of the AIA provide a detailed list of two categories of high-risk AI systems: 1) AI that is a safety component of products subject to strict regulation (such as AI in medical devices or industrial machinery); 2) AI used in sensitive areas (such as education, recruitment, financial services, criminal justice, biometric identification, migration management, etc.). High-risk AI systems must comply with a set of stringent obligations under Chapter III of the AIA, including: developers of high-risk AI must implement a risk management system throughout the AI lifecycle, and assess and mitigate risks both before and after the system is placed on the market. They must ensure data governance and the quality of training data (to avoid bias or distortion). Technical documentation must be comprehensive, instructions for use must be provided, and users must be clearly informed when interacting with AI to ensure transparency at this level. In addition, human oversight must be ensured for this category of risk, so that humans can monitor, intervene in, or deactivate the AI if necessary when high risk arises. The system must meet requirements regarding accuracy, cybersecurity, and reliability, and must undergo conformity assessment or self-assessment before being placed on the market. Furthermore, providers must register high-risk AI systems in the EU database and establish a post-market mechanism to collect feedback and address issues. Not only developers but also other actors in the AI supply chain (such as importers, distributors, and deployers) bear corresponding responsibilities.

iii) Limited risk: This group consists of AI systems that pose limited risks but still require certain transparency safeguards. According to Article 52 of the AIA, AI systems that interact with humans or generate content that may cause confusion must comply with an obligation to inform users that they are interacting with AI. The EU stops short of imposing legal obligations at this level and instead encourages developers to adopt voluntary codes of conduct for such AI systems to ensure they remain user-friendly and trustworthy.

iv) Minimal or no risk: According to Article 69 of the AIA, this group includes the vast majority of AI systems, particularly common applications such as spam filters, AI-based video games, or smart office tools, since these tools are considered to pose little or no significant risk. For this group, the AIA imposes no additional legal obligations. Governance is primarily left to developers and other actors guided by ethical principles, thereby avoiding unnecessary burdens on innovation.

The AIA is built upon the "New Legislative Framework" (NLF) of the European Union, which serves as the overarching legal framework for product-related legislation and establishes standards for health and consumer safety [16]. In particular, the prohibition of AI systems deemed to pose an unacceptable risk exemplifies an ex-ante approach (prior to occurrence - preventive regulation) [17]. This means that the EU is prepared to intervene and prevent such AI systems from being placed on the market or put into use from the outset, based on the determination that these systems are likely to cause serious harm and conflict with core societal values, rather than waiting for adverse consequences to materialise [18]. This reflects the EU's prioritisation of protecting citizens and social values from identified threats posed by AI.

In addition, the principle of proportionality, one of the general principles of EU law, is enshrined in Article 5(4) of the Treaty on European Union (TEU), which requires that "the

content and form of Union action shall not exceed what is necessary to achieve the objectives of the Treaties". The principle of proportionality demands that burdens (financial or administrative) placed on economic operators and citizens must be minimised in proportion to the goals to be attained [19].

Furthermore, the EU places strong emphasis on human-centric AI, which constitutes one of the core motivations behind the AIA [20]. This objective demonstrates Europe's aspiration to become a global leader in the development and deployment of advanced, ethical and safe AI, promoting a human-centred approach in the global context. The EU seeks to distinguish itself from other major powers by emphasising its ethical, value-based, and human-focused model [21]. The risk classification under the AI Act draws directly from the Charter of Fundamental Rights of the European Union as its primary normative reference [22].

Finally, the EU values responsible innovation across the legislative process and its implementation by relevant stakeholders. It has introduced the AI Innovation Package and the regulatory sandbox mechanism to support businesses in testing low-risk AI in a controlled environment. This reflects a balance between regulatory oversight and the encouragement of innovation, indicating a long-term vision: managing risk to build user trust and thereby promote the sustainable development of AI.

Although the AIA is grounded in relatively robust theoretical foundations for a risk-based approach, it nonetheless presents certain strengths and weaknesses.

3.1. Advantages of the Risk-Based Approach in the EU Artificial Intelligence Act

The AIA's approach focuses regulatory efforts on the most critical areas, specifically targeting AI systems likely to pose high risks to human health, safety, and fundamental rights. This ensures that strict regulations are applied to those AI systems that present the greatest risks. For other systems, mechanisms such as "soft law" instruments are employed instead, thereby avoiding unnecessary regulatory burdens across the board.

The EU also seeks to strike a balance between control and the promotion of innovation. Its objective is to reach an optimal point between minimising AI-related risks and encouraging technological advancement. This developmental axis helps legislators to avoid overregulation and respect the principle of proportionality during the formulation of legally binding rules before a law is finalised. Fundamentally, the AIA provides a relatively clear legal structure for classifying AI systems by risk levels. In other words, it constructs a "risk-level pyramid" that offers a visual and structured method for identifying and categorising risks.

By adopting a human-centric risk-based approach, the EU is able to pursue the goal of building public and societal trust in the notion of "trustworthy AI made in Europe". This, in turn, facilitates the acceptance and sustainable development of AI and enhances the competitive advantage of European AI companies in comparison to other major powers such as the United States or China.

3.2. Disadvantages of the Risk-Based Approach in the AIA

The classification and differentiation of risk levels under the AIA remain the subject of considerable debate. Risk categorisation relies heavily on the socially acceptable level of risk, which is not a purely quantitative exercise defined by clear criteria, but rather a socio-cultural assessment dependent on public acceptance [20]. In other words, the identification of risk is a complex matter and can lead to confusion between different risk levels. Any risk determination must therefore take into account social values and fundamental rights. Consequently, the assessment of risk levels rests with competent authorities, meaning that AI providers may face uncertainty regarding the costs they must incur to ensure legal compliance. A notable illustration of this uncertainty can be found in a survey of 106 AI systems conducted

within the EU, which reported that 18% of AI systems fell into the high-risk category, 42% into the low-risk category, while 40% could not be clearly classified as high-risk or not [23].

At present, technical standards such as codes of practice or technical guidelines, intended to ensure that AI systems operate safely, transparently, and reliably remain incomplete [25]. This affects the ability to ensure compliance with the AIA. Under EU law, if a company adheres to officially recognised EU technical standards (known as "harmonised standards"), it benefits from the presumption of conformity. In other words, compliance with such standards serves as proof of legality, and no further justification is required. However, to date, no harmonised standards have been adopted for AI, especially with respect to risk management in high-risk systems [24]. In the absence of harmonised standards, those developed by ISO/IEC JTC 1/SC 42 (the subcommittee on AI) may serve as a temporary reference, but they do not carry legal weight and cannot trigger the presumption of conformity [16]. This means that even if companies apply such standards, they are still required to demonstrate separately that their systems meet the legal requirements.

The current classification of AI risk levels under the AIA continues to present limitations and inconsistencies for stakeholders. One of the primary reasons is the lack of clear guidance on how to determine which risks are considered "acceptable" under the existing legal framework. In this context of regulatory ambiguity, the responsibility for assessing and classifying risk levels is often transferred to AI providers, namely, those entities developing or deploying AI systems. Other actors in the AI usage chain, including purchasers, deployers, and end-users, may be unaware of the scope of their legal obligations in managing and addressing residual risks. Specifically, they may not be adequately equipped with a legal basis to determine whether they are required to implement additional mitigation measures or to issue warnings to affected individuals about remaining risks after deployment. Such a lack of clarity regarding responsibility can result in two serious consequences: (1) these actors are forced to make legal assessments and decisions under conditions of limited information, increasing the risk of error or non-compliance; and (2) in the event of harm, they may be exposed to unforeseen legal liability. This constitutes a clear example of "legal uncertainty", which poses a significant barrier to the widespread and safe deployment of AI systems in practice.

4. Current Legal Framework for AI Risk Classification and Some Recommendations for Vietnam

4.1. The Current State of Vietnamese Law on Risk Classification in Artificial Intelligence

Prior to 14 June 2025, Vietnam had not enacted a dedicated law governing AI. Issues relating to AI were only regulated in part under broader legal frameworks concerning cybersecurity, information safety, and privacy, as found in normative instruments such as the Law on Cybersecurity 2018, the Law on Network Information Security, Decree No. 13/2023/ND-CP, Law on Personal Data Protection 2025, among others. In practice, these instruments may only be interpreted as indirectly applicable to AI to the extent that their regulatory subjects relate to AI technologies. This has led policy experts to assess that Vietnam's policy and legal framework on AI remains limited in light of practical needs. The development of AI has significantly outpaced both legal provisions and ethical guidelines in this area [17, 25][4].

[4] Over the past years, a series of strategic policy documents on artificial intelligence have been issued in Vietnam, including: (1) Resolution No. 52-NQ/TW on certain guidelines and policies for proactive participation in the Fourth Industrial Revolution, (2) Decision No. 127/QD-TTg on the National Strategy for Research, Development and Application of Artificial Intelligence to 2030, (3) Decision No. 569/QD-TTg on the issuance of the Strategy for Science, Technology and Innovation Development to 2030, (4) Resolution No. 57-NQ/TW on Breakthrough

The Law on Cybersecurity 2018 is primarily concerned with protecting national security and public order in cyberspace. While the law does not explicitly address AI, it establishes regulatory measures based on the level of risk posed to critical information systems. In addition, it prohibits and sanctions numerous harmful activities in cyberspace, including cyberterrorism, cyberattacks, and the use of cyberspace for criminal or subversive purposes. These prohibited acts could encompass malicious uses of AI, such as disseminating disinformation or using deepfakes for fraudulent purposes. Therefore, if an AI application currently causes harm to cybersecurity or public order, it can only be addressed through the general provisions of this law.

Decree No. 13/2023/ND-CP and Law on Personal Data Protection 2025 do not directly regulate AI-specific matters, but they exert a substantial influence on AI applications, since most AI systems involve data collection and processing. Accordingly, any AI system that collects or analyses personal data within Vietnam must comply with the provisions of Decree 13 and LPDP 2025. For instance, an AI system used in recruitment or healthcare that processes sensitive personal data such as health or biometric information must obtain the explicit consent of the data subject and adopt appropriate security measures. This results in a de facto risk classification based on the nature of the data used. Where sensitive data is involved, privacy risks are considered higher, thereby triggering stricter regulatory requirements.

Thus, Vietnam has not previously established a mechanism for classifying AI according to risk. Legal instruments such as the Law on Information Technology, the Law on Access to Information, the Law on Network Information Security, and other currently applicable regulations do not contain any provisions concerning the classification or management of AI based on risk levels. As such, it may be said that the existing legal framework for AI in Vietnam remains incomplete. However, from the perspective of "soft law" principles, the Ministry of Science and Technology issued Decision No. 1290/QD-BKHCN, which provides guidance on principles for the responsible development of AI systems. This document outlines a vision of building a human-centred society, with a balanced approach to the benefits and risks of AI. The Ministry proposed nine core principles to be observed in the research, design, development and provision of AI systems, including: 1) Cooperation and innovation; 2) Transparency; 3) Human oversight; 4) Safety; 5) Security; 6) Privacy; 7) Respect for human rights and dignity; 8) User assistance; and 9) Accountability.

This can be regarded as the first ethical code of conduct for AI in Vietnam, although its official title is a "principles guidance document". The Ministry of Science and Technology encourages organisations and individuals engaged in the field of AI to apply these principles voluntarily. The content of these principles aligns with the universal values reflected in international ethical AI guidelines.

Recognising the importance of developing AI in a controlled and responsible manner, the Law on Digital Technology Industry No. 71/2025/QH15 was passed by the National Assembly on 14 June 2025 and will take effect on 1 January 2026. However, because it contains only five provisions, the LDTI 2025 does not yet clearly classify different levels of risk associated with AI systems [26]. Accordingly, it can be argued that AI has moved far beyond the traditional notion of the digital industry and technology, as it is not only viewed through a technological or industrial policy lens but also

Development in Science, Technology, Innovation, and National Digital Transformation, and (5) Resolution No. 03/NQ-CP on the Government's Action Programme for implementing Resolution No. 57-NQ/TW. This report focuses in greater detail on Resolution No. 57-NQ/TW and Decision No. 127/QD-TTg. Although Vietnam has issued numerous policy documents on AI, to a certain extent the practical implementation remains slow and such policies have yet to be fully incorporated into domestic law through detailed legal provisions.

understood as a multilayered political and social phenomenon [27]. For this reason, the Draft Law on AI was introduced to address this gap Under Article 35 on transitional provisions of the Draft AI Law dated 25 November 2025, the provisions on AI in the LDTI 2025 are expected to be repealed. Unlike the Law on Digital Technology Industry, Article 43 of LDTI 2025 categorises AI systems into four types: i) high-risk AI systems, ii) non-high-risk AI systems, iii) AI systems with significant impact and iv) other AI systems. By contrast, the Draft Law on AI dedicates the entirety of Chapter II to the classification and risk-based management of AI systems, adopting a different four-tier structure: 1) unacceptable-risk AI systems; 2) high-risk AI systems; 3) medium-risk AI systems; and 4) low-risk AI systems. This demonstrates a clear policy choice in favour of a risk-based regulatory approach in Vietnam's Draft AI Law. In addition, when compared with the EU AIA, Vietnam's approach exhibits notable conceptual alignment with this legislative model.

According to Clause 1, Article 7 of the Draft Law on AI, unacceptable-risk AI systems are those capable of causing severe and irreparable harm to human rights, national security, social order, or public safety, or those used for activities prohibited by law. Accordingly, Article 11 of the Draft AI Law absolutely prohibits the development, provision, deployment, or use of unacceptable-risk AI systems in Vietnam. These include systems that: engage in conduct prohibited by law; use impersonation or simulation of persons or events to deceive or systematically manipulate cognition or behaviour; exploit the vulnerabilities of vulnerable groups; or create or disseminate falsified content that poses serious threats to national security or social order. This represents a significant improvement in Vietnam's efforts to draw on the EU AIA's experience, as the relevant provisions focus more deeply on structurally harmful behaviours capable of causing irreparable damage. This incorporation is evident in the adoption of an impact based risk assessment mechanism, which is a core principle of the EU AI Act. Under this approach, risk levels are not determined by the AI technology itself but by its intended purpose, its operational context, and its potential to generate adverse consequences for human rights and social order.

However, several issues that are regulated under the EU AIA, such as social scoring systems, systems that assess or predict the likelihood of an individual committing a crime, systems that create or expand facial recognition databases through untargeted scraping of facial images from the Internet or from camera or CCTV sources, emotion recognition systems, biometric categorisation systems, and remote real time biometric identification systems, have not yet been incorporated into the Vietnamese Draft AI Law. This omission results from differences in the rights-based regulatory approach of the European Union, which has increasingly placed strong emphasis on the protection of human rights. Therefore, the inclusion of these matters in the near future will be necessary if Vietnam moves towards a legislative philosophy aligned with that of the European Union.

Furthermore, the absolute prohibition of AI systems that manipulate behaviour, deceive users systematically, or exploit vulnerable groups reflects a modern legislative mindset grounded in scientific evidence on psychological and behavioural harm, informational asymmetry, and algorithms' ability to amplify risks. In addition, Vietnam's decision to include national security, public order, and public safety as central criteria for identifying unacceptable risk represents an important adjustment. This approach aligns with the characteristics of Vietnam's legal and political environment and ensures a balance between adopting international best practices and safeguarding essential public interests.

According to Clause 2, Article 7 of the Vietnamese Draft AI Law, high risk AI systems are those that may cause harm to life, health, rights, lawful interests of organisations or individuals, or other important public interests.

In other words, the central criterion for classifying a system as high risk is the severity of its impact on legally protected interests, including life, health, fundamental rights, public interests, public order, and public safety. This approach aligns with the European Union's legal framework, even though the draft law does not specify the method for assessing severity.

However, similar to the LDTI 2025, the Draft AI Law does not define the threshold at which a risk becomes "severe" enough to fall within the high-risk category. The absence of quantification or explicit evaluation criteria makes the risk threshold theoretically flexible but lacking in practical predictability. This may result in inconsistencies in risk assessment across enforcement bodies or across different sectors.

Furthermore, the draft still relies on qualitative and difficult-to-measure terms, most notably the expression "serious harm". This raises an important question. For instance, should an error in an AI driven recruitment decision that causes an applicant to lose an employment opportunity be considered "serious", or is this concept reserved only for situations involving physical harm or threats to life. These gaps demonstrate the need to implement guidelines to ensure transparency and to anticipate how the law will be applied.

In the Draft AI Law, the concept of an AI system with significant impact has been removed. According to Clause 2, Article 43 of LDTI 2025, a significant impact AI system is defined as an AI system used for multiple purposes, with a large number of users, a large number of parameters, and a large volume of data. The basis for identifying such systems lies in these indicators: multipurpose use, large user base, parameter volume, and data scale.

The approach to this concept in the LDTI 2025 appears to have drawn inspiration from the European Union's treatment of general-purpose AI systems, also known as GPAI. However, there is a key difference between the approach of the EU and that of Vietnam. Under the EU framework, general-purpose AI systems are considered to have a significant impact when they exceed a cumulative computing capacity of ten to the power of twenty-five (10^25). In addition, the annex provides detailed criteria for assessing risk, including the number of users, the number of parameters, and the volume of computational resources. In contrast, Vietnam's legal framework simply refers to large numbers of users, parameters, and data volume, without defining how these are to be quantified or what specific thresholds must be met [28]. From the perspective of regulatory logic based on risk, it may be observed that high risk AI systems are classified according to legal and societal consequences. By contrast, AI systems with significant impact are identified through technical indicators such as parameter count, data volume, and scale of user base. The absence of a unified reference system or risk taxonomy makes the classification scheme appear fragmented, difficult to apply consistently, and particularly vulnerable to conflicts when an AI system exhibits overlapping characteristics across multiple categories. Although the provision on significant impact of AI systems in the LDTI 2025 contains certain limitations as discussed above, it should not be entirely removed. If the regulatory framework retains only the two categories of high risk and unacceptable risk, a governance gap will emerge for AI systems that are not dangerous enough to warrant strict restrictions yet still capable of producing clear and measurable effects on human rights, the information environment, or the market. In addition, a tiered risk structure that includes a significant impact category, as adopted in the EU AI Act, helps ensure alignment between Vietnamese law and global standards and facilitates compliance for businesses operating in an international environment. More importantly, this intermediate category preserves the balance between innovation and the protection of rights, preventing both excessive regulatory tightening and complete regulatory absence. Therefore, although the concept requires further refinement, the category of significant impact AI systems

remains essential and should not be removed. Instead, improvements should be made to address its internal shortcomings.

Clauses 3 and 4 of Article 7 of the Vietnamese Draft AI Law provide definitions of medium-risk and low-risk AI systems. A medium-risk AI system is one that may cause confusion, manipulation, or deception for users because they are unable to recognise the AI nature of the system or the content it generates. A low risk AI system is defined as any system that does not fall within the categories specified in the preceding clauses. This approach does not differ significantly from the definitions adopted in the EU AI Act.

However, the current legal framework does not provide specific criteria or a mechanism for identifying the components of this category, particularly a list of high risk AI systems similar to the approach adopted in the European Union. The absence of such guidance leads to several practical consequences. When facing difficulties in determining the risk level of an AI system, competent authorities or relevant actors tend to classify the system into the remaining category through a process of elimination. Although this method appears simple from a legislative drafting perspective, it creates a range of legal issues that must be addressed. These include determining the authority responsible for classification, developing procedures and evaluation criteria, and clarifying the obligations of entities involved in the development, provision, and operation of AI systems.

Therefore, the current provisions in the LDTI 2025 and the Draft AI Law unintentionally create a grey zone within the legal framework, resulting in risk classification that lacks consistency, predictability, and legal certainty. This situation may hinder both the effectiveness of state management and the healthy development of the AI market.

4.2. Some Policy Recommendations for Vietnamese Law

Based on the analysis of the AIA and the existing shortcomings in Vietnam's legal framework for regulating AI, the author proposes several recommendations to improve Vietnamese law, with a view towards establishing a modern and coherent regulatory system.

First, Vietnam should adopt a dedicated law governing AI, rather than regulating AI in a limited manner under the LDTI 2025, which is essential in the current Vietnamese context. This recommendation is grounded in the following three arguments:

i) In accordance with the principle of codification in modern legislative technique, technologies that exert far-reaching influence on the structure of society, such as AI, should be governed by a specialised legal instrument at the statutory level. Such an instrument would serve to establish a comprehensive conceptual framework, define the scope of application, set out governance principles, determine risk classification, clarify the rights and obligations of relevant parties, and provide oversight mechanisms. The mere inclusion of a few provisions in the LDTI 2025, which already covers a broad range of fields such as digital devices, software, and the semiconductor industry, is insufficient to address the complex, sensitive, and interdisciplinary issues associated with AI. These include biometric identification, automated decision making, machine ethics, and implications for human rights [30].

ii) The risk-based governance approach is considered a progressive regulatory model. However, if this approach is not enshrined directly in the statutory text and instead delegated to subordinate legislation such as decrees or circulars, it would violate the constitutional principle of separation of powers. According to Clause 2, Article 14 of the Constitution 2013, only the National Assembly has the authority to limit or affect the human rights and civil liberties of individuals. Risk classification of AI systems entails obligations of transparency, restrictions on system functionality, bans on deployment, mandatory product withdrawal, and even criminal liability. All of these measures directly impact rights such

as property ownership, freedom of enterprise, and privacy. Delegating the authority to define classification criteria, scope of application, and sanctions entirely to executive agencies places regulatory power in the hands of bodies outside the legislative branch, without a mechanism for legislative oversight. This contravenes the rule-of-law principle of "foreseeability" in legal regulation.

iii) From a legal technical perspective, the model of multi-layered regulation requires that the highest layer, namely the statutory law, must establish the legal architecture [29]. This includes the guiding principles, categories of risk, types of legal obligations, regulated stakeholder groups, limitations on state power, and mechanisms for adjudication. Subordinate legal instruments should serve only as implementation guidance and technical updates. Examples include the list of prohibited AI systems, documentation requirements for risk assessment, or source code standards. In the absence of this foundational layer, the system will lack predictability, be vulnerable to arbitrary policy shifts, and create institutional risks.

iv) In the context of Vietnam's push for digital transformation and its participation in next-generation trade agreements that include provisions on technology and data, such as the CPTPP and the DEPA, a dedicated AI law is not only an internal requirement of the national legal system. It is also a necessary condition for ensuring the country's capacity to coordinate nationally in response to rapidly emerging international standards. Without a specific law, Vietnam will lose strategic standing in shaping digital sovereignty strategies and weaken its capacity to protect vulnerable groups, especially consumers and those at risk in AI-driven environments.

Second, Vietnam should introduce legal provisions that categorically prohibit certain AI systems based on human rights protection and assessments of societal tolerance thresholds. At present, LDTI 2025 does not provide for a category of AI systems that must be absolutely prohibited. This omission fails to reflect the human-centred approach that Vietnam aims to adopt fully. Certain types of AI systems, such as technologies that manipulate behaviour without user awareness, social scoring based on inferred personal data, or biometric categorisation that may reinforce structural bias, not only pose high risks but also infringe upon moral boundaries and human dignity. Ordinary monitoring mechanisms cannot adequately mitigate these harms.

However, before enacting such prohibitions, the legislature must not rely solely on technical logic or international ethical standards. The National Assembly should proactively conduct investigations, evaluations, and public consultations to determine the community's tolerance threshold for prohibited conduct. This is a necessary step to ensure the social acceptability of legal norms, particularly in a context where technology is evolving rapidly and social values are shifting. A technology may be condemned in one jurisdiction but accepted in another, depending on differences in legal culture, public perception, and foundational societal values. Assessing societal tolerance thresholds not only enables lawmakers to delineate the boundary between absolute prohibition and strict control, but also facilitates the establishment of transparent and accountable criteria for determining which technologies should be placed on the list of banned systems.

Third, the Draft AI Law should consider reinstating the concept of AI with significant impact, while ensuring that this term does not conflict or overlap with the definition of high-risk AI so as to preserve coherence and consistency in legislative drafting. When reintroducing this concept into the legal framework, it is essential to clarify the logical relationship between the classification groups. The key question is whether AI with significant impact encompasses high-risk AI, excludes it, or exists as a parallel category with different legal obligations. Without a clearly structured classification system, ambiguity will arise and the application of the law will be impeded.

A consistent classification framework should be built on a common reference axis, meaning that the types of impact must be

measured by criteria that are equivalent or can be translated into one another. If one category is defined by quantitative indicators such as the number of users or the scale of data, while another is based on qualitative indicators such as the degree of interference with human rights, it will be extremely difficult to develop a unified, fair, and verifiable enforcement mechanism for all relevant actors.

Therefore, Vietnam should review its classification structure for AI systems and harmonise the core concepts in a manner that is coherent and operational in practice. This approach not only prevents conceptual overlap but also reduces the compliance burden for enterprises, while facilitating the enforcement, supervision, and sanctioning functions of regulatory authorities. Finally, with respect to high-risk AI systems, Vietnam should issue an annex that specifies the detailed list, technical criteria, and corresponding compliance obligations. This approach ensures transparency and practical feasibility in implementation and provides a unified foundation for enterprises and regulators in assessing, deploying, and supervising high-risk AI systems.

Fourth, Vietnam should establish a national focal agency for AI to ensure regulatory coherence, effective coordination, and enforceability in the classification of AI risks. This role should be assigned to the Ministry of Science, Technology and Innovation. A key principle in the regulation of high technology is institutional consistency in the formulation of classification criteria and the implementation of regulatory mechanisms. Under LDTI 2025, the authority to define and classify high-risk and large-impact AI systems is delegated to individual ministries and sectors, based on their respective areas of governance. This reflects a traditional administrative decentralisation approach. However, AI is not a linear technology confined to a single sector. It functions as an interdisciplinary, cross-sectoral digital infrastructure that simultaneously affects domains such as healthcare, justice, transportation, commerce, education, and national security.

The absence of a national-level agency with the authority to establish and coordinate the AI risk classification system may lead to the following problems:

i) Each ministry or sector may develop its own criteria for risk assessment based on its technical capacity, policy priorities, or subjective perception of acceptable risk. This could result in a situation where the same AI system is classified as high risk in one sector but considered non-threatening in another. Such inconsistency in implementation creates difficulties for multi-sector businesses and undermines uniformity in enforcement.

ii) This fragmentation will undermine the integrity of the national AI risk governance system. A core requirement of risk-based regulation is the existence of a unified reference system that enables consistent risk assessment, comparison, and timely updates. If ministries and sectors independently define and interpret risk without a central coordinating authority, Vietnam will be unable to establish a consolidated risk database. It will also be incapable of measuring the overall technological risk landscape and therefore unable to develop accurate long-term policy planning.

iii) The risk classification system functions as the legal backbone of all subsequent regulatory mechanisms, including compliance obligations, conformity assessment, licensing, and legal liability. When this backbone is fragmented or lacks standardisation at the central level, the entire AI legal ecosystem becomes misaligned.

To address these three challenges, Vietnam should consider assigning the authority to define and coordinate AI risk classification to the Ministry of Science, Technology and Innovation. This agency would be responsible for the following:

- Developing unified national criteria and a risk classification system for AI.
- Receiving data and classification proposals from specialised ministries and sectors.

- Conducting balanced assessments of benefits, risks, and compliance costs.
- Issuing an official list of AI system categories according to risk levels.
- Updating classifications periodically based on technological developments and societal impact.

5. Conclusion

In response to the rapid development of AI, Vietnam has shown initiative in adopting and learning from the European Union's risk-based governance model in its legislative efforts. This marks a positive step that reflects Vietnam's determination to align its legal framework with the digital age. However, the current legal design in Vietnam still presents several shortcomings. Classification criteria remain unclear and lack quantifiable indicators, leading to potential conceptual overlap and regulatory inconsistencies across categories.

A comparison with the EU's legislative experience reveals that, while Vietnam has embraced several risk-based elements in its classification approach, its legal framework has yet to establish a sufficiently robust and coherent regulatory structure. On this basis, the author concludes that Vietnam should refine its legal approach to AI by adopting a dedicated AI law. This law should fill the current legal gaps in AI regulation, address overlapping conceptual frameworks, and assign final authority and coordination responsibilities to a specialised, competent national agency. Such reforms are necessary to ensure regulatory coherence and systematisation in addressing AI-related issues.

Ultimately, this will enable Vietnam's legal framework to approximate international best practices in AI governance. It also affirms the principle that law should not hinder scientific and technological progress, but rather guide and enable innovation within a structured and rights-respecting regulatory environment.

Acknowledgments

This research is funded by Vietnam National University HoChiMinh City (VNU-HCM) under grant number DM2024-34-02.

References

[1] Vietnam Lawyer Journal, The Year 2024 Marks a Comprehensive Revolution of Artificial Intelligence. https://lsvn.vn/nam-2024-danh-dau-cuoc-cach-mang-toan-dien-cua-tri-tue-nhan-tao-a151967.html, 2024 (accessed on: June 01st, 2025).

[2] Vietnam Law and Legal Forum, Managing AI Development: Legal Challenges and Responsibilities. https://vietnamlawmagazine.vn/managing-ai-development-legal-challenges-and-responsibilities-72073.html, 2024 (accessed on: May 23rd, 2025).

[3] EU Artificial Intelligence Act Portal, Key Issues Risk-Based Approach. https://www.euaiact.com/key-issue/3#:~:text=The%20EU%20AI%20Act%20introduces,classifies%20risk%20into%20four%20categories, 2024 (accessed on: May 23rd, 2025).

[4] P. M. Hanh, The European Union's Artificial Intelligence Act: Some Contents and Legal Issues, Journal of Democracy and Law, 2(413) (2024) https://danchuphapluat.vn/stores/customer_file/dcpl/102024/10/10-pham-minh-hanh-413-094150.pdf (accessed on: May 23rd, 2025).

[5] Organisation for Economic Co-operation and Development, Recommendation of the Council on Artificial Intelligence (OECD Legal Instrument OECD/LEGAL/0449). https://legalinstruments.oecd.org/en/instruments/OECD-LEGAL-0449, 2019 (accessed on: May 23rd, 2025).

[6] European Commission, White Paper on Artificial Intelligence (2020), COM(2020), pp. 10 - 19.

[7] S. Wachter, B. Mittelstadt, C. Russell, Why Fairness Cannot Be Automated: Bridging the Gap Between EU Non-Discrimination Law and AI, Computer Law & Security Review 41 (2021) 105567. http://dx.doi.org/10.2139/ssrn.3547922 (accessed on: May 23rd, 2025).

[8] National Institute of Standards and Technology, AI Risk Management Framework (AI RMF 1.0) (NIST AI 100-1). https://nvlpubs.nist.gov/nistpubs/ai/nist.ai.100-1.pdf, 2023 (accessed on: May 23rd, 2025).

[9] Akpuokwe CU, Adeniyi AO, Bakare SS, Legal Challenges of Artificial Intelligence and Robotics: A Comprehensive Review, Computer Science & IT Research Journal 5(3) (2024), http://www.fepbl.com/index.php/csitrj (accessed on: May 23rd, 2025).

[10] Bathaee, Y., The Artificial Intelligence Black Box and the Failure of Intent and Causation, Harvard J. Law Technol. 31, 2018, pp. 889-938.

[11] Ben Chester Cheong, Transparency and Accountability in AI Systems: Safeguarding Wellbeing in the Age of Algorithmic Decision-Making, Frontiers in Human Dynamics 6 (2024) https://doi.org/10.3389/fhumd.2024.1421273.

[12] Emmanouil Papagiannidis, Patrick Mikalef, Kieran Conboy, Responsible Artificial Intelligence Governance: A Review and Research Framework, The Journal of Strategic Information Systems 34(2) (2025) 101885. https://doi.org/10.1016/j.jsis.2024.101885.

[13] Institute for Policy and Media Development, Artificial Intelligence Policy in Vietnam: Balancing Development Promotion and Risk Control (IPS 2024), pp. 73.

[14] Luca Bertuzzi, AI Act: Leading MEPs Revise High-Risk Classification, Ignoring Negative Legal Opinion (EURACTIV, 23 October 2023) https://www.euractiv.com/section/artificial-intelligence/news/ai-act-leading-meps-revise-high-risk-classification-ignoring-negative-legal-opinion/ (accessed on: May 23rd, 2025).

[15] H. Fraser, Bello, J. M Villarino, Acceptable Risks in Europe's Proposed AI Act: Reasonable and Other Principles for Deciding How Much Risk Management Is Enough, European Journal of Risk Regulation (2023) doi:10.1017/err.2023.57.

[16] D. Golpayegani, H. J. Pandit, D. Lewis, To be High-Risk, or not to Be-Semantic Specifications and Implications of the AI Act's High-Risk AI Applications and Harmonised Standards' in Proceedings of the 2019 2nd International Conference on Control and Robot Technology (ACM 2023) doi:10.1145/3593013.3594050.

[17] Institute for Policy and Development of Communication, AI Policy in Vietnam: Balancing Development Promotion and Risk Control, IPS, 2025, pp. 13.

[18] N. Rangone, L. Megale, Risks Without Rights? The EU AI Act's Approach to AI in Law and Rule-Making, European Journal of Risk Regulation 1 (2023) doi:10.1017/err.2025.13.

[19] M. Ebers, Truly Risk-Based Regulation of Artificial Intelligence How to Implement the EU's AI Act, European Journal of Risk Regulation 1 (2024), doi:10.1017/err.2024.78.

[20] H. Fraser, Bello y J. M. Villarino, Acceptable Risks in Europe's Proposed AI Act: Reasonable and Other Principles for Deciding How Much Risk Management Is Enough, European Journal of Risk Regulation, 2023, doi:10.1017/err.2023.57.

[21] I. Ulnicane, Artificial Intelligence in the European Union: Policy, Ethics and Regulation trong Thomas Hoerber, Gabriel Weber and Ignazio Cabras, The Routledge Handbook of European Integrations (Routledge 2022) doi:10.4324/9780429054136.chl-19.

[22] L. Hogan, Mark Lasek-Markey, Towards a Human Rights-Based Approach to Ethical AI Governance in Europe, Philosophies 9(6) (2024) 181 doi:10.3390/philosophies9060181.

[23] Applied AI Institute for Europe, AI Act: Risk Classification of AI Systems from a Practical Perspective https://aai.frb.io/assets/files/AI-Act-Risk-Classification-Study-appliedAI-March-2023.pdf (accessed on: May 23rd, 2025).

[24] J. Schuett, Risk Management in the Artificial Intelligence Act', European Journal of Risk Regulation 15(2) (2024) pp. 367-385 doi:10.1017/err.2023.1.

[25] L. M. Sang, T. Đ. Thành, Artificial Intelligence and Legal Challenges, Vietnam Journal of Science and Technology No. 8 (2020), https://scholar.dlu.edu.vn/thuvienso/bitstream/DLU123456789/143734/1/50605-469-154434-1-10-20200909.pdf (accessed on: May 23rd, 2025).

[26] T. T. T. Dung, N. V. Lam, N. V. Duong, Legal Issues of Artificial Intelligence: International Experiences and Implications for Vietnam, Legislative Studies Review 04 (2025) pp. 50-63.

[27] C. Katzenbach, AI will fix this - The Technical, Discursive, and Political Turn to AI in Governing Communication, Big Data & Society 8 (2021) pp. 1-5. https://doi.org/10.1177/20539517211046182

[28] N. L. Phuong, Classifying AI Systems by Risk Level, May 23, 2025, https://daibieunhandan.vn/phan-loai-he-thong-ai-theo-muc-do-rui-ro-10373456.html (accessed on: May 23rd, 2025).

[29] Hansard, Legislation: Skeleton Bills and Delegated Powers - Motion to Take Note (House of Lords Debates, 6 January 2022). https://hansard.parliament.uk/Lords/2022-01-06/debates/D37A1DDC-E0FB-461F-AB09-3DDC0E6D9769/LegislationSkeletonBillsAndDelegatedPowers (accessed on: May 23rd, 2025).

[30] OECD, Multi-level Regulatory Governance (OECD Working Papers on Public Governance No 13, 2009) https://www.oecd.org/gov/regulatory-policy/42406309.pdf (accessed on: May 23rd, 2025).