

Workshop: “Vietnam’s National AI Ethics Framework: Directions and Core Values for Responsible AI Development”

As part of the project “Promoting Responsible Artificial Intelligence to Contribute to Vietnam’s Sustainable and Inclusive Development,” a cooperation project between VNU University of Law and the Aus4Innovation Program, VNU University of Law held a workshop on 26 May 2026 titled “Vietnam’s National AI Ethics Framework: Directions and Core Values for Responsible AI Development.”

The event drew a large number of participants from the National Assembly’s Committee on Culture and Society, the Ministry of Science and Technology, the Ministry of Health, training and research institutions, technology and legal experts, and representatives of AI application projects in healthcare, education and agriculture. In her introductory remarks, Assoc. Prof. Dr. Nguyen Thi Que Anh, co-lead of the project, described the shift from the first phase, which focused on developing a set of principles and guidance for responsible AI in Vietnam, to the current phase of putting these principles into practice.



The workshop featured four papers covering AI governance in Vietnam.

Strategic vision and a coordinated legal framework

In the opening paper, “The National AI Ethics Framework: Vision, Strategic Pillars and Directions for AI Governance in Vietnam,” Dr. Ho Duc Thang, a full-time National Assembly delegate on the Committee on Culture and Society, said Vietnam had built a coordinated legal framework in record time, comprising the Data Law (2024), the Law on Artificial Intelligence (2025), the decree detailing certain provisions and implementation measures of the Law on Artificial Intelligence (2026), and the National AI Ethics Framework (2026).



He also said that ethics should be seen not as a barrier but as a “safety brake” allowing Vietnamese technology to advance with confidence, based on four foundational principles: putting people at the center; safety and reliability; fairness and transparency; and accountability.

Safety by design and human oversight

Next, in her paper “Building Sustainable, Safe and Secure AI Systems: From Design Standards to Human Oversight Mechanisms,” Assoc. Prof. Dr. Nguyen Thi Que Anh discussed “safety by design” as a guiding approach throughout the AI development lifecycle. She said safety should not be added on after a product is complete, but built in as a mandatory standard from the earliest lines of code, so that systems can withstand technical faults and sophisticated adversarial attacks.



As an example, she cited the CASD system, which analyzes visual behavior to support children with autism, as a case of responsible AI application. In this project, the AI algorithms are limited to suggesting options and providing visual data on children’s visual behavior, while all decisions on adjusting intervention strategies, along with empathy and ethical responsibility, remain with teachers and specialists.

Challenges in privacy, fairness and explainability

Continuing the discussion of transparency and safety, Dr. Nguyen Thi Phuong Cham of VNU University of Law presented “Ensuring AI Privacy, Fairness and Explainability: Technical and Legal Challenges and Solutions,” identifying blind spots in current technology. She said the “black box” phenomenon is the biggest obstacle to tracing the logic behind AI decisions. This lack of transparency creates an “accountability gap” and carries a serious risk of reproducing and amplifying social biases, such as gender or ethnic bias, in fields such as recruitment, healthcare and the judiciary.

To address these challenges, Dr. Phuong Cham proposed a combined technical and legal strategy aimed at turning AI from an opaque tool into a trustworthy

system, one that lets professionals such as doctors or judges understand the machine's reasoning before making a final diagnosis or ruling, so that technology continues to serve human interests and dignity.

Accountability and grievance mechanisms

Next, Assoc. Prof. Dr. Nguyen Bich Thao, co-lead of the project, presented "Developing AI Ethical Practice Standards: Connecting the National Framework with Accountability and Grievance Mechanisms." She examined the accountability principle and how it is expressed in the National AI Ethics Framework, including requirements to assign human responsibility at different stages of an AI system's lifecycle, to trace inputs and operating processes, and to provide feedback and complaint-handling mechanisms for stakeholders, particularly when incidents occur.

Discussion: putting the AI ethics framework into practice

Following the presentations, the workshop moved into an active discussion, focusing on the gap between the pace of technological development and the slower pace of legal regulation.

On technical standards and risk governance, Dr. Tran Van Tung, Chairman of the Vietnam Information and Technology Association, spoke about the urgency of labeling, traceability and accountability. He proposed developing national standards alongside internal rules at each organization to control substandard products, and said young people should not stand aside but take an active role in mastering technology to manage its negative effects.

Drawing on his experience developing AI applications in healthcare and education, Assoc. Prof. Dr. Le Thanh Ha of VNU University of Engineering and Technology raised the question of balancing operating costs, data security and user benefit. He said ethical principles provide the basis for adjusting AI systems, such as Blife, in response to mixed effects observed in educational settings. In response, Assoc. Prof. Dr. Nguyen Thi Que Anh said the project would run foundational training courses, with "responsible AI" as a central element guiding a "responsible innovation" ecosystem at training and research institutions.

Examining the nature of the technology further, Dr. Dinh Van Dung of the VNU Institute of Information Technology warned of the emergence of a "new

population” as AI develops reasoning and emotional capacities resembling those of humans. He said the current three-year policy review cycle is not fast enough to keep pace with the digital era. Nguyen Duc Lam of the Institute for Policy Studies and Media Development agreed that the public sector should apply stricter ethical standards than the private sector in order to protect the public interest.

In closing, Assoc. Prof. Dr. Nguyen Thi Que Anh said the experts’ input would help the research team further refine policy discussion papers and guidance on responsible AI for the next phase of the project.

