Center for Technology Innovation at BROOKINGS

AISA
AI SAFETY ASIA

REPORT
AUGUST 2025

AI SAFETY GOVERNANCE, THE SOUTHEAST ASIAN WAY

Lyantoniette Chua, Philip Tham, Chinasa T. Okolo

AI SAFETY GOVERNANCE, THE SOUTHEAST ASIAN WAY

AUGUST 2025

LYANTONIETTE CHUA is a co-founder of AI Safety Asia (AISA), where she is the executive director for strategic futures & global affairs.

PHILIP THAM is the managing editor of this report and an analyst for AI Safety Asia (AISA).

CHINASA T. OKOLO is a fellow in the Governance Studies program at the Brookings Institution.

ABOUT AI SAFETY ASIA

Established in 2024, AI Safety Asia (AISA) is a global non-profit dedicated to building Asia as a globally-leading safe and responsible AI innovator. We strive to minimise the risks of AI adoption while supporting societies to adopt AI safely. At the core of our mission is the integrated, diverse and collaborative nature of our work. We believe that achieving safe and governed AI begins with building bridges—between generations, disciplines, and regions through actionable steps across our three interrelated program pillars that convene, research and build capacity to shape the future of AI governance in Asia. Beginning in Southeast Asia, we support the development of policy, technical, and research capabilities across the Asia-Pacific region to help ensure AI is governed safely and responsibly.

READING THIS REPORT

The report is organized in various sections. The main insights can be found in **Section 2.1: Regional snapshot** (which synthesizes commonalities across the 10 ASEAN member states and Timor-Leste), **Section 3: Recommendations for Southeast Asia,** and **Section 4: Lessons for the wider world.**

Specific country profiles are in **Section 2: Regional and country reports** and are meant to be modular and stand alone. Observers keen to learn more about a specific country can simply refer to each profile in isolation from other sections. Similarly, reading all country reports in the section is not necessary to gather the key insights from this report.

Key definitions and/or insights from each section are also consolidated in blue boxes across the report. These summaries are meant to be concise, with explanations embedded in the main text of the report.

ACKNOWLEDGMENTS AND EDITORIAL TEAM

This report is a joint effort between AI Safety Asia and the Brookings Institution Center for Technology Innovation. The coauthors are Lyantoniette Chua (Chief Editor and co-founder, AISA), Philip Tham (Managing Editor, Lead Analyst, AISA), and Chinasa T. Okolo (Publishing Director, Fellow, Brookings Institution).

Contributors

We gratefully acknowledge the expertise and time shared by the contributors listed below, arranged in alphabetical order by last name, including panelists from AISA's 2024-2025 roundtable series. While their perspectives informed this work, the views and interpretations presented do not necessarily represent those of the individuals or their organizations.

- **Zar Motik Adisuryo:** Analyst, AI Safety Asia
- **Hoda Alkhzaimi:** Co-Chair, Global Future Council for Cyber Security, World Economic Forum
- **Rina Chandran:** Southeast Asia Editor, Rest of World
- **Siriwat William Chhem:** Co-Founder and President of AI Forum; Advisor at BowerGroupAsia
- **Massimiliano Colonna:** Director of Communications, Governance Studies, The Brookings Institution
- **Michelle Du:** Research Intern, Center for Technology Innovation, The Brookings Institution
- **Justo Fernandes:** Director of Information System and Application Development, Timor Leste
- **Quentin Flament:** Analyst, AI Safety Asia
- **Mohan Kankanhalli:** Director, National University of Singapore AI Institute
- **Esben Kran:** Founder, Apart Research; Co-Founder, European Network for AI Safety
- **Jam Kraprayoon:** Researcher, Institute for AI Policy and Strategy
- **Nitya Kuthiala:** Analyst, AI Safety Asia

- **Lee Wan Sie:** Cluster Director, AI Governance & Safety, IMDA
- **Barani Maung Maung:** Tech Policy and Safety Expert; Online Safety Policy Lead, Ofcom
- **Alex Moltzau:** Policy Officer, European AI Office, European Commission
- **Micheal Nelson:** Senior Fellow, Asia Program, Carnegie Endowment for International Peace
- **Danh Nguyen:** Global Head of Legal at GreenFuture.tech, Legal-Tech Expert at LegalTech.org
- **Isyrah Fahmi Osman:** Head of Research and Development, Information Technology Protective Security Services, Brunei Darussalam
- **Adelle Patten:** Visual Design and Website Manager, Governance Studies, The Brookings Institution
- **Elmer C. Peramo:** Senior Science Research Specialist, Department of Science and Technology – Advanced Science and Technology Institute, Philippines
- **Benjamin Prudhomme:** Vice President, Policy, Safety and Global Affairs, Leadership Team, Mila – Quebec AI Institute
- **Alex Read:** Chief Technical Specialist, UNDP Lao PDR
- **Alejandro Reyes:** Scholar-in-Residence, Asia Society Hong Kong Center; Adjunct Professor, University of Hong Kong
- **Hammam Riza:** President, Kolaborasi Riset dan Inovasi Industri Kecerdasan Artifisial; Delegate to UK AI Summit for Indonesia
- **Supheakmungkol Sarin:** Executive Director, AI Safety Asia
- **Josie Stewart:** Research and Communications Assistant, Center for Technology Innovation, The Brookings Institution
- **Raymond Sun:** Lawyer, Herbert Smith Freehills; Director, Data Science and AI Association of Australia; Developer, Global AI Regulation Tracker
- **Bunhong Taing:** Director of Policy and Legal Affairs, Digital Government Committee, Cambodia
- **Jun-E Tan:** Senior Research Associate, Khazanah Research Institute
- **Robert Trager:** Co-Director, Oxford Martin AI Governance Initiative; International Governance Lead, Centre for the Governance of AI; Senior Research Fellow, Blavatnik School of Government, University of Oxford

- **Edward Tsoi:** Co-Founder, AI Safety Asia
- **Nicol Turner Lee:** Senior Fellow, Governance Studies and Director, Center for Technology Innovation, The Brookings Institution
- **Abhishek Vats:** Sustainability Director, ASEAN Youth Organization
- **Judy Wang:** Research Projects Assistant, Governance Studies, The Brookings Institution
- **Yonah Welker:** Public Technologist; Former Tech Envoy (EU Commission projects), Ministry of AI
- **Gita Wirjawan:** Former Minister of Trade, Indonesia; Chairman and Founder, Ancora Group and Ancora Foundation

Reviewers

We would also like to thank our reviewers for their invaluable contributions:

- **Thorin Bristow:** AI Governance Associate, One World Trust
- **Nick Caputo:** Researcher, Oxford Martin AI Governance Initiative
- **Chloe Chadwick:** DPhil Researcher, Oxford Internet Institute, University of Oxford
- **Alessandro Salvatore Ciceri:** Junior Program Officer, AI Governance, International Telecommunication Union, AI for Good
- **Luise Eder:** Doctoral Researcher, University of Oxford; AI Governance Initiative, AI Governance Alliance
- **Sherly Haristya:** Adjunct Lecturer, LSPR Institute of Communication and Business, Jakarta, Indonesia
- **Mari Izumika:** Research Fellow, AI Alignment Network
- **Anonyo Mitra:** Capacity Building Lead, AISA
- **Lee Shi Hao:** Former Insight and Technical Lead, AISA
- **Jason Tucker:** Researcher, Institute for Futures Studies, Sweden; Adjunct Associate Professor, AI Policy Lab at Umeå University

EXECUTIVE SUMMARY

Artificial Intelligence (AI) safety governance today is largely shaped by global powers like the United States, China, and Europe. Yet, as AI systems grow more powerful and governance models begin to harden, Southeast Asia cannot afford to be a bystander. With its digitally connected populations, rapid technological adoption, and strategic geopolitical positioning, the region has much at stake and much to contribute. While the region has recently taken strides in building a robust AI safety ecosystem, more work still needs to be done. At the same time, Southeast Asia's developmental trajectory in this field offers a timely case for how diverse and developing Global Majority regions can draw from international best practices, while tailoring solutions to their own realities. This report makes the case for why Southeast Asia's voice must be amplified in global AI safety conversations now, before policy frameworks and international norms are set without its input. To illustrate this point, the authors detail the local operational environment, present recommendations, and share lessons for global partners seeking to advance this space.

The stakes are high. As AI capabilities advance rapidly, Southeast Asian countries must contend with frontier risks, governance gaps, and intensifying great-power competition. This report provides a snapshot of national and regional AI safety initiatives, updated as of June 2025, to show how countries are dealing with these problems. Under this point, the report identifies four key themes characterizing the current state-of-play of AI safety policy:

1. A universal recognition of AI's significance, but nuanced differences in perspectives.
2. A pragmatic approach to AI and digital policy development that leverages the benefit of hindsight and strikes a balance between pro-business innovation and safety.
3. Diversity as a double-edged sword for regional collaboration.
4. Persistent shared challenges confronting national-level AI risk management, such as a lack of quality data sets, poor cybersecurity, and capacity constraints.

Rooted in this understanding of the operational context, the report outlines a path forward. The authors identify 10 actionable recommendations for ASEAN and national governments.

- Among ASEAN partners, the emphasis should be on harmonizing approaches rather than enforcing regulatory uniformity, while managing great-power dynamics with pragmatism. Strengthening cooperation in areas such as capacity building, research, and catastrophic risk management is also essential.
- Governments at the national level should consider pursuing practical, high-impact measures—such as improving policy implementation, enhancing data collection, addressing frontier AI risks, and streamlining institutional processes—while actively engaging with international partners.

Finally, the report situates Southeast Asia within the global landscape, positioning the region as a reference point for other Global Majority contexts. AI safety governance the Southeast Asia way is marked by localized and inclusive governance, regional cooperation, and open-source tools. Key features include a focus on multilingual large language models, culturally aligned open-source safety tools, and ASEAN-led coordination. Taken together, these elements offer timely insights for a world racing to govern AI safely and equitably.

GLOSSARY OF ACRONYMS[1]

ADGMIN: ASEAN Digital Ministers' Meeting, a yearly meeting of ASEAN Digital Ministers that started in 2021.

ADGSOM: ASEAN Digital Senior Officials' Meeting, a coordinating meeting of senior officials focusing on digital affairs that assists ADGMIN and oversees the WG-AI.

AGI: Artificial General Intelligence, defined by the UN High-Level Panel on Artificial General Intelligence, AGI refers to "systems capable of equaling or surpassing human intelligence in diverse cognitive tasks."[2]

AI: Artificial Intelligence, defined in the ASEAN Guide on AI Governance and Ethics as the discipline of making analytical machines intelligent, with some forms of AI being able to adapt on its own by learning through use.[3]

ASEAN: Association of Southeast Asian Nations, the main regional organization in Southeast Asia, comprising 10 countries: Brunei, Cambodia, Indonesia, Laos, Malaysia, Myanmar, the Philippines, Singapore, Thailand, and Vietnam. Timor-Leste, the last Southeast Asian country, has applied to join ASEAN and its membership will be finalized by October 2025.

ISO/IEC JTC 1/SC 42: The International Organization for Standardization (ISO) and the International Electrotechnical Commission (IEC) Joint Technical Committee 1 for Information Technology, Sub Committee 42 on Artificial Intelligence is a subcommittee that develops and facilitates the development of international standards, technical reports, and technical specifications regarding AI.

LLM: Large Language Models (like ChatGPT, DeepSeek, Claude, etc.), generative models that leverage large datasets to generate natural language output.

R&D: Research and development

---

1 National-level acronyms are excluded in this list for brevity.

2 "Governance of the Transition to Artificial General Intelligence (AGI) Urgent Considerations for the UN General Assembly" (Council of Presidents of the United Nations General Assembly (UNCPGA), May 2025), https://uncpga.world/wp-content/uploads/2025/05/AGI-UN-CPGA-Report-pdf-2.pdf.

3 "ASEAN Guide on AI Governance and Ethics" (Association of Southeast Asian Nations, 2024), https://asean.org/wp-content/uploads/2024/02/ASEAN-Guide-on-AI-Governance-and-Ethics_beautified_201223_v2.pdf.

RAIR: ASEAN Responsible AI Roadmap (2025–2030) is a policy paper released in March 2025 to provide governments with guidance on prioritizing and operationalizing responsible AI in an integrated and interoperable manner. While released, it has not yet been endorsed and implemented by ASEAN.

RAM: Readiness Assessment Methodology, a toolkit by UNESCO for countries to assess their preparedness for AI development and governance

UNESCO: United Nations Educational, Scientific, and Cultural Organization

WG-AI: Working Group on AI, an ASEAN-level working group under the ADGSOM responsible for coordinating ASEAN's AI governance efforts—led by Singapore.

CONTENT

CONTENT

1.0. INTRODUCTION: 1.1. BACKGROUND

> **The global rules for Artificial Intelligence (AI)** are being written now, but they are being shaped predominantly by major players like the United States, China, and Europe by virtue of their concentration in compute, talent, and finance.[4]

When it comes specifically to AI safety[5]—broadly speaking, the management of risks associated with AI—these same countries shape the narrative. This leaves Southeast Asia[6]—a largely Global Majority,[7] highly digitally connected, and populous region—underrepresented in the global discourse. With over 700 million people (many of whom are young and online), internet penetration has doubled over the past decade to over 73 percent in 2024,[8] while the region's digital economy is projected to grow to almost $1 trillion by 2030.[9] This rapid digitalization

4 Hélène Draux, "Research on Artificial Intelligence – the Global Divides," Digital Science, Research on Artificial Intelligence – the Global Divides (blog), January 4, 2024, https://www.digital-science.com/blog/2024/01/research-on-artificial-intelligence-the-global-divides/.

5 See Section 1.3 Defining AI safety and its governance for more explanation on why we use "AI safety" as opposed to other terms such as secure, responsible, or trustworthy AI.

6 There are 11 countries in Southeast Asia, which are, in alphabetical order, Brunei, Cambodia, Indonesia, Laos, Malaysia, Myanmar, the Philippines, Singapore, Thailand, Timor-Leste, and Vietnam. All the countries, except Timor-Leste, are current member states of ASEAN. Timor-Leste is a prospective member of the organization and will join by October 2025.

7 "Global Majority" here refers to parts of the world traditionally considered as "developing" or the Global South, but highlights the fact that these regions make up the majority of the global population.

8 "Internet usage in Southeast Asia - statistics & facts," Statista, accessed July 18, 2025, https://www.statista.com/topics/9093/internet-usage-in-southeast-asia/#topicOverview

9 Joo-Ok Lee, "How ASEAN Is Building Trust in Its Digital Economy," World Economic Forum, How ASEAN Is Building Trust in Its Digital Economy (blog), January 12, 2024, https://www.weforum.org/stories/2024/01/asean-building-trust-digital-economy/.

underscores the need for inclusive and robust AI safety governance to protect the region's people from risks associated with the emerging technology. As Southeast Asia continues to integrate digital technologies into various sectors from health care to financial services, ensuring that AI systems are developed and deployed responsibly becomes paramount. This report seeks to correct the representational imbalance in global AI safety circles and spotlight progress made across Southeast Asia.

The report focuses on three key areas. First, the authors provide a snapshot of salient points in the AI safety policy landscape across the region and in each individual country.[10] Second, rooted in this understanding of the operating environment, the report proposes orienting ideas for the region and national governments to develop more robust governance frameworks. Third, the report embeds the region in a broader global context. It should be seen as a concise primer on the governance landscape toward AI safety in Southeast Asia for global observers, providing a set of considerations for Southeast Asian policymakers on the regional and national levels, and a playbook for decisionmakers of other Global Majority regions to draw from based on developments up to March 2025.

10 Note that this section does not aim to be a comprehensive overview of all policies related to governance, policies, and safety in AI for each country/organization. Rather, it is merely a summary of salient points raised during our roundtable series. See Appendix A for a comprehensive list of speakers, and Appendix B for an overview of our methodology.

1.2. METHODOLOGY

> **Insights were pulled from six roundtable discussions** organized by AI Safety Asia (AISA) over half a year from September 2024 to March 2025 involving over 30 panelists and more than 1,000 participants.

Crucially, all 10 ASEAN member states and Timor-Leste (a prospective member of the organization) were represented during the roundtables. Panelists represented a diversity of stakeholders, ranging from various levels of government (from former ministers to current directors of agencies in charge of AI policy), the private sector, academia, and civil society. The authors then supplemented these insights with desk research to provide more context. A more in-depth discussion of this methodology is provided in Appendix B. While this report draws on insights shared by speakers during the roundtable series, the recommendations and interpretations presented reflect the authors' analysis and are not necessarily the views of the speakers or their affiliated organizations. The content aims to build upon and be inspired by the discussions, rather than represent direct endorsements by individual participants.

1.3. WHY THIS REPORT, AND WHY NOW

AI developments are currently happening at lightning speed. In a May 2025 report, the UN High-Level Expert Panel on Artificial General Intelligence warned that AI systems are "rapidly advancing toward artificial general intelligence," with industry leaders and experts anticipating that artificial general intelligence (AGI) could emerge "within this decade".[11] Meanwhile, global approaches to AI safety governance are also becoming increasingly hardened—whether innovation- and industry-led in the U.S., state-driven in China, or centered on consumer rights in the European Union.

The inclusion of Southeast Asia in global AI governance discourse matters now more than ever because it presents a timely case of how diverse and developing regions can draw from global best practices while tailoring solutions to their own political, cultural, and economic realities. In doing so, Southeast Asia offers a pragmatic, locally-grounded model that may resonate with other parts of the Global Majority facing similar challenges.

The report is also a call to action and identifies critical gaps that must be addressed to ensure inclusive, effective, and future-ready policies. For policymakers within and beyond the region, the authors provide a unique lens to consider how AI safety governance can be both globally informed and locally responsive.

11 "Governance of the Transition to Artificial General Intelligence (AGI) Urgent Considerations for the UN General Assembly" (Council of Presidents of the United Nations General Assembly (UNCPGA), May 2025), https://uncpga.world/wp-content/uploads/2025/05/AGI-UN-CPGA-Report-pdf-2.pdf.

1.4. DEFINING AI SAFETY, ITS GOVERNANCE, AND OTHER CONCEPTS

AI safety is a nebulous concept.
While key definitions remain contested, it can be better understood compared against related terms.

At risk of oversimplification, AI security refers specifically to the protection of AI systems from external threats—such as cyberattacks or adversarial manipulation—and tends to have a narrower, more technical focus. In contrast, AI safety concerns the design, deployment, governance, control, and use of AI systems in ways that prevent unintended harm in use cases. Some entities and experts have also argued for a more holistic view of AI safety that incorporates leveraging the benefits of AI for human societies and the environment, as well as alignment with human values.[12] While there is overlap between the two—particularly in ensuring system robustness—AI safety has a broader scope, including not only technical reliability and control but also normative ethical concerns such as fairness, transparency, accountability, and the promotion of socially beneficial outcomes. Some of these concerns are also addressed under the overlapping terms responsible AI or trustworthy AI, which emphasize ethical governance, human rights, and public trust. That said, it should be acknowledged that "AI safety" is not a neutral or universally defined term, but a concept that carries various interpretations and implications depending on technical, cultural, and geopolitical contexts. For uniformity, the report adopts a deliberately grounded, policy-relevant interpretation of the concept focusing on institutional readiness, risk management, and developmental goals.

This report focuses on regional AI safety initiatives—the gamut of initiatives by stakeholders like multilateral organizations, governments, the private sector, academia, or civil society adopted with the intention of promoting AI safety. Such initiatives include, among other things, AI safety governance, which refers to intergovernmental organization or governmental policies[13] aiming to regulate the design, deployment, adoption, and use of AI systems to mitigate

12 There are admittedly live philosophical questions on issues such as the human-centricity of AI safety governance. This report, however, does not take a hard stance on these issues for pragmatic reasons, as we primarily aim to examine concrete policies and initiatives already taken or omitted by stakeholders in the region.

13 Note that the term "policies" is used loosely here to refer to intergovernmental organization or government-issued guidelines that guide decisionmaking. Legislative acts, executive

its risks. Examples of AI safety governance in this report include hard laws, soft regulatory frameworks, voluntary guidelines, and ethics principles. Nongovernance AI safety initiatives, including research and development on AI risk management and AI projects like contextualized and multilingual large language models (LLMs) to ensure equity and inclusive development are also discussed.

Of note, initiatives that primarily approach AI from an economic standpoint—such as offering subsidies to AI developers to set up operations—are not considered by this report to fall under AI safety governance, even if they may have indirect downstream effects of cultivating AI expertise and strengthening safety governance capacities over time.

There is also a distinction between "traditional" and "frontier" AI systems in the report. In line with the ASEAN Expanded Guide on AI Governance and Ethics—Generative AI (2025), traditional AI (sometimes also called predictive or diagnostic AI) is defined as systems that are trained on data to complete specific tasks such as classification or prediction.[14] These systems generally focus on specific tasks, operate within a narrow scope, and follow often hard-coded rules. Frontier AI systems, as outlined in the Bletchley Declaration (2023), refer to highly capable models with broader capabilities that can conduct a wide variety of tasks on top of the traditional AI capabilities, and that match or exceed the capabilities present in today's most advanced models.[15] Included in this category are systems like cutting-edge generative AI models, agentic AI, and AGI.[16]

Similarly, the report distinguishes between "catastrophic" and "societal" risks. Catastrophic risks refer to low-probability but high-impact risks. In the context of AI safety, catastrophic AI risks could include AI-enabled human extinction or the irreversible collapse of human civilization, possibly through the loss of human control, unintended consequences, or goal misalignment. It is often discussed in relation to advanced AI capabilities. In contrast, social risks are more "normal" nonexistential risks, often with social, ethical, and economic considerations.[17] AI-enabled job displacement, misinformation and disinformation, and inter-community tension

---

orders, frameworks, voluntary guidelines, strategies, etc. are collectively considered to be policy in this report.

14 "Expanded ASEAN Guide on AI Governance and Ethics–Generative AI" (ASEAN, January 2025), https://asean.org/book/expanded-asean-guide-on-ai-governance-and-ethics-generative-ai/.

15 "The Bletchley Declaration by Countries Attending the AI Safety Summit, 1–2 November 2023," GOV.UK, November 2023, https://www.gov.uk/government/publications/ai-safety-summit-2023-the-bletchley-declaration/the-bletchley-declaration-by-countries-attending-the-ai-safety-summit-1-2-november-2023.

16 Admittedly,"frontier" AI is a moving goal-post as AI development advances rapidly. For instance, while generative AI LLMs like OpenAI's GPT-3.5 might have been considered "frontier" when first released in 2022, these models are relatively commonplace and non-frontier anymore.

17 Atoosa Kasirzadeh, "Two Types of AI Existential Risk: Decisive and Accumulative," January 17, 2025, https://arxiv.org/html/2401.07836v3.

belong to this category.

Key definitions (summarized from above paragraphs for ease of reference)

- **AI safety** concerns the design, deployment, governance, control, and use of AI systems in ways that prevent unintended harm. As opposed to AI security, which focuses narrowly on technical reliability, it incorporates broader normative ethical concerns. The authors adopt a grounded, policy-relevant interpretation of the concept, focusing on institutional readiness, risk management, and developmental goals.

- **AI safety initiative** is a broad catch-all term used to refer to the gamut of initiatives by stakeholders (multilateral organizations, governments, the private sector, academia, civil society, etc.) taken with the intention of promoting AI safety. It includes AI safety governance, which refers to policies aimed at regulating the design, deployment, adoption, and use of AI systems to mitigate its risks. Nongovernance AI safety initiatives include research and development (R&D) work and multilingual AI projects.

- **"Traditional AI"** has a narrow scope, is often hard-coded, and produces classifications and/or predictions as a primary output.

- **"Frontier AI"** is a moving goal post referring to highly capable models with broader, more general capabilities that match or exceed the state-of-the-art.

- **"Catastrophic"** risks are low-probability but high-impact risks that could cause existential threats, such as (but not restricted to) the irreversible collapse of human civilization.

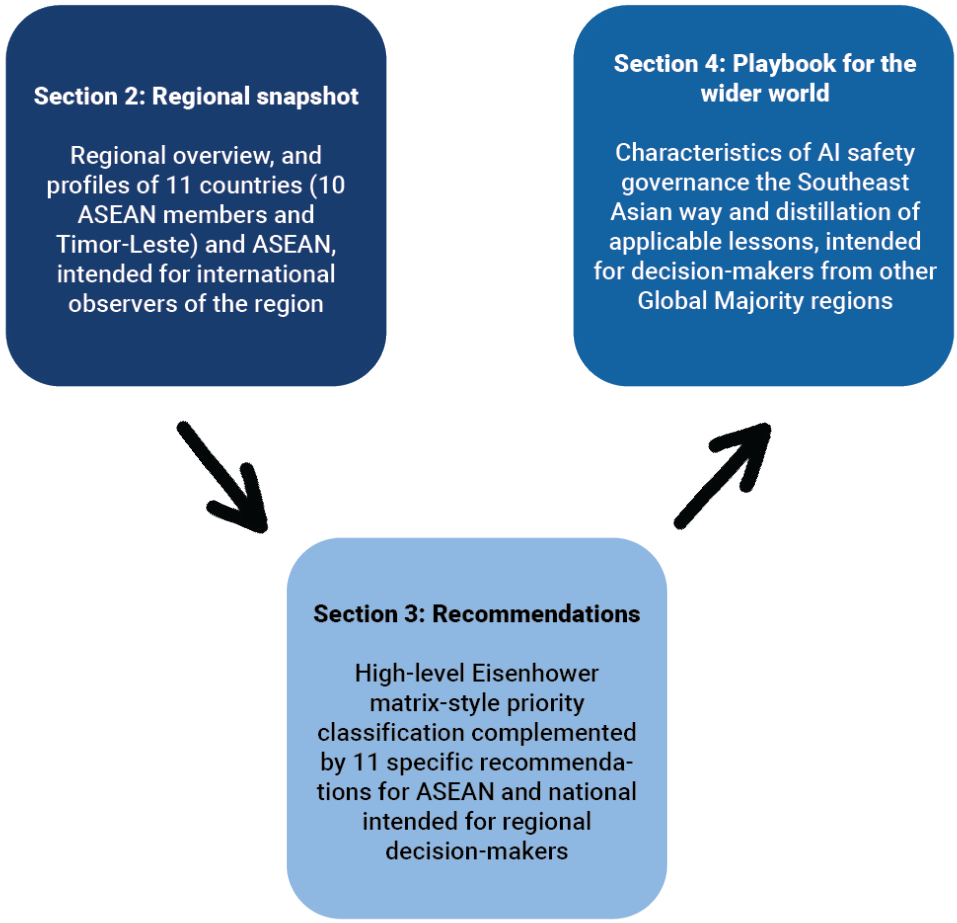
- **"Societal"** risks are more "normal," higher probability risks that do not typically raise existential concerns, with these including AI-enabled job displacement and misinformation.

1.5. ORGANIZATION OF THE REPORT

The report is organized in six sections. The main insights can be found in Section 2.1: Regional snapshot (which synthesizes commonalities across the 10 ASEAN member states and Timor-Leste), Section 3: Recommendations for Southeast Asia, and Section 4: Lessons for the wider world. Specific country profiles in Section 2: Regional and country reports are modular and stand alone. Each section can be read in isolation from others, and reading all country reports is not necessary to gather key insights from the report.

FIGURE 1

Organization of the report



On Section 2: Region and country reports

This section captures a snapshot of AI safety governance initiatives across the region and then specifically in the ten ASEAN member states, Timor-Leste, and finally ASEAN itself, updated up to June 2025. At the country level, we outline the current domestic AI safety initiative landscape, the country's international engagement in this field, notable case studies, and key insights drawn from our roundtable discussions. Importantly, this snapshot is not meant to be comprehensive—it offers a primer for what panelists at AISA's roundtable discussions deemed to be the most significant aspects of policy, aimed at global observers seeking to understand regional and country-specific developments.

On Section 3: Recommendations for Southeast Asia: 10 policy directives for 2025–2030

Policy recommendations build on the regional state-of-play outlined in Section 2. This section proposes using an Eisenhower matrix to classify high-level priorities, followed by 10 orienting ideas—four for ASEAN (2025–2028) and six for national governments (2025–2030).

On Section 4: Lessons for the wider world

Southeast Asia's AI governance is shaped by localized, pragmatic approaches, regional cooperation, inclusive multi-stakeholder processes, and open-source safety tools. While not without challenges, its evolving framework offers valuable lessons—especially for Global Majority regions with similar contexts.

2.0. REGION AND COUNTRY REPORTS

> In this section, the authors provide a snapshot of governance levels toward safe, secure, and responsible AI across Southeast Asia, identify four key themes characterizing the current state-of-play for regional AI policy on the domestic level, and outline the region's footprint in the international arena.

Following this overview, the report delves into country-specific profiles in Sections 2.2 to 2.12, providing a primer on governance initiatives across the 10 ASEAN member states, Timor-Leste (a prospective ASEAN member slated to join the organization in October 2025), and then for the regional organization itself in Section 2.13.

For each country profile, a summary and timeline of key developments are provided, as well as a lay-of-the-land for domestic developments and international engagements, insights from the roundtable sessions, and case studies of AI safety initiatives. Content in this section is based on salient points raised by speakers at our roundtable sessions. Note that the views shared by speakers do not necessarily reflect that of organizations they are affiliated with. A more in-depth discussion on the roundtable data collection methodology is presented in Appendix B. The authors also include quotes from the roundtable session to underscore the importance of key points, some of which have been edited for length and clarity.

2.1. REGIONAL SNAPSHOT

The AI safety governance landscape among Southeast Asian countries is diverse. Case in point, Singapore has been ranked by the Global Index on Responsible AI[18] in 2024 as 11th in the world and second in Asia and Oceania, while Laos and Myanmar rank in the bottom 15 percent.[19] This section seeks to capture broad trends and characteristics of the AI policy space across the region extracted from the specific country profiles in the succeeding section.

Table 1 offers a summary of Southeast Asian countries and how they fare in their policy approaches. In this overview, comprehensive hard law on AI refers to regulations covering multiple stages in the AI life cycle (development, deployment, adoption) that are legally binding, such as the EU AI Act. Comprehensive soft law on AI refers to regulations covering multiple stages in the AI life cycle that are voluntary and not legally binding, such as AI Governance and Ethics guides. AI governance research center refers to a research hub or center that focuses specifically or primarily on AI safety governance or responsible AI.

18 The index country level scores are a result of a comprehensive assessment of country commitments to the responsible development of AI with a human rights focus, across 19 thematic areas, complemented with robust secondary data sources.

19 "Global Index [on] Responsible AI 2024 1st Edition Report" (Global Index [on] Responsible AI, n.d.), https://girai-report-2024-corrected-edition.tiiny.site/.

TABLE 1

Government-led AI policymaking in Southeast Asian countries

Implementation | Not yet addressed | In progress/pending

Country	National AI strategy	Comprehensive hard law	Comprehensive soft law (voluntary guides, principles, etc.)	Generative AI regulation/ frameworks	AI governance research center	AI governance institutional structure
Singapore	✓ (2019, updated 2023)	×	✓ (MGF, 2020)	✓ (MGF-Gen AI, 2024)	✓ (DTC)	Smart Nation initiative under MDDI
Indonesia	✓ (2020)	~ (Draft legislation, unclear implementation date)	✓ (Circular Letter No. 9/2023)	×	×	Komdigi, BRIN (*)
Malaysia	✓ (2021)	×	✓ (AIGE Guidelines, 2024)	×	×	NAIO
Thailand	✓ (2022)	~ (Draft Royal Decree, unclear implementation date)	✓ (Digital Thailand Ethics Guidelines, 2021)	✓ (Generative AI Governance Guidelines for Organisations, 2024)	✓ (VISTAI)	ETDA (AIGC), ONDE, NSTDA
Vietnam	✓ (2021)	~ (Draft Digital Tech Law, expected 2026)	✓ (Decision No. 1290/2024)	×	~ (National AI innovation centres, unknown status)	MoST, MIC, MPS (*)
Philippines	✓ (2021)	~ (4 pending bills)	~ (Patchwork of sectoral regulations, but no overarching guide)	×	✓ (CAIR)	DTI, DOST, DICT (*), proposed AI offices
Cambodia	~ (2025, Draft National AI Strategy)	×	×	×	×	MPTC (*)
Laos	~ (2025 planned)	×	×	×	×	MoTC(*)
Myanmar	×	×	×	×	×	MoTC(*)
Brunei	×	×	✓ (Draft AIGE Guide, 2025)	×	×	AITI (AIGE Working Group)
Timor-Leste	×	×	×	×	×	TIC TIMOR (*)

NOTE: Light blue indicates implementation, orange indicates not yet addressed, and dark orange indicates in progress/pending. (*) means agency does not focus on AI policy, but has previously handled some aspect of it.

Of the 10 ASEAN member states, the findings of Table 1 suggest that six countries (Singapore, Indonesia, Malaysia, Thailand, Vietnam, and the Philippines) are relatively more mature in their AI safety governance policy trajectories. These countries have published national AI strategies and have adopted (or are close to adopting) comprehensive soft regulatory frameworks for AI. The four remaining ASEAN members (Laos, Myanmar, Brunei, and Cambodia) and Timor-Leste have less developed regulatory frameworks, as they have yet to publish national AI strategies and have taken limited steps toward regulatory adoption. Among the six more mature countries, Singapore stands out given the plethora of resources it has already invested into AI governance initiatives. In contrast to its neighbors, it has already adopted regulation on generative AI, has an internationally recognized AI governance research center, and has a government entity specifically focusing on AI policy.[20] Indonesia, Malaysia, Thailand, Vietnam, and the Philippines are slowly adopting similar approaches, having taken significant steps in 2024 and 2025 to discuss and implement comprehensive AI policy. However, only some of these countries have considered recent developments in generative AI (Thailand) and/or taken at least initial strides in building up government-backed AI safety R&D capacity (Thailand, Philippines, and Vietnam).

Drawing from the country-specific profiles below,[21] four key themes can be extracted around the current state-of-play of regional AI safety policy that include:

1. **A universal recognition of AI's significance, but nuanced differences in perspectives.** On the whole, all Southeast Asian countries are aware of AI's significant impact on the economy, society, and politics. All countries have, unsurprisingly, expressed a desire to harness this emerging technology for economic benefit. Yet, each country views the significance of AI from a slightly different perspective. For instance:

    - Malaysia sees AI as an economic leveler, seeking to improve its AI readiness and expertise to move into high-income economy status. The fervor surrounding its recent boom in data center construction bears testament to this.

    - The Philippines is more acutely aware of the socioeconomic consequences of job displacement that might arise due to its unique job market profile that has a heavy emphasis on business process outsourcing (BPO).[22]

---

20 The SEA AI Policy Observatory, a hub for tracking AI policies in Southeast Asia developed by AISA, has documented at least 45 pieces of AI-related policy documents adopted by Singapore since 2020. Indonesia (18), Malaysia (12), the Philippines (23), Thailand (17), and Vietnam (16) lag behind by number of AI policy documents. Note that these numbers include ASEAN-level AI documents as well as non-AI safety related policies. These numbers are merely meant to illustrate the governance gap in the region. "SEA AI Observatory," accessed July 2, 2025, https://seaobservatory.com/.

21 See profiles from Sections 2.2 to 2.13 for more detailed information and citations.

22 "The Global BPO Sector: The Impact of Outsourcing to the Philippines," Movate (blog), April 15, 2024, https://www.movate.com/articles/outsourcing-bpo-services-to-the-philippines-the-global-leaders/.

- Vietnam's challenging geopolitical situation compels the government to focus on national data protection and cybersecurity as a key pillar of AI safety governance.

- Countries like Myanmar and Timor-Leste are aware of the importance of AI but have limited governance capacity. The former is in the midst of a civil war while the latter is a new country focused on the provision of basic infrastructure and services. These countries' AI conversations tend to be centered around the application of AI to solve existing developmental preoccupations, as well as immediate and specific concerns like misinformation, deep fakes, and cybercrime.

2. **A pragmatic approach to policy development in the Southeast Asian region.** Southeast Asian countries take a pragmatic approach to AI safety governance, and their policy positions are best understood in the context of their unique operating environments.

    - Some Southeast Asian countries acknowledge the need to quickly formulate policy, often leveraging the benefit of hindsight and recontextualize policy frameworks from abroad.

        - Thailand's AI Draft Royal Decree and Vietnam's Draft Law on Digital Technology Industry borrow from the EU AI Act's risk-based approach.

        - Indonesia and the Philippines have used UNESCO's AI Readiness Assessment Methodology (RAM) framework[23] to develop policy tools to enhance their AI safety governance landscapes. Cambodia, Laos, Malaysia, Thailand, and Vietnam are in the process of completing the RAM process, while Timor-Leste is preparing to start.

        - Brunei's upcoming cybersecurity and AI Governance and Ethics guidelines borrow from ASEAN and U.K. documents.

    - Countries are also aware of the need to strike a balance between fostering pro-business innovation and ensuring safety, though there is a growing realization that these two objectives are not mutually exclusive.[24] Yet in the face of limited governance bandwidth, a priority on economic development, and the perceived relative certainty of AI's economic potential vis-à-vis the unsettled "frontier"-ness of AI safety science, AI policy tends to be cast within broader digital economy initiatives while safety policy takes a backseat. In a similar vein, while countries have been willing to adopt voluntary AI governance and ethics guidelines, they remain more reticent about omnibus

23 The RAM is a tool created by UNESCO that was mandated by its Recommendation on the Ethics of Artificial Intelligence in 2021. It is designed to help countries evaluate their preparedness to develop and govern AI.

24 Local experts at our Singapore, Indonesia, and Malaysia panel discussions mentioned that there was growing awareness of how AI safety provisions could be seen as a crucial prerequisite to responsible AI development rather than as an impediment to innovation.

hard-law instruments. This mix of hard and soft law reflects sensible approaches to managing the two priorities but might complicate ASEAN's goal of regional regulatory harmonization.

3. **Diversity as a double-edged sword for regional collaboration.** Given the nuanced differences in perspectives of AI, operating contexts, and existing regulatory frameworks, each country has distinct priorities and discrete objectives for AI safety governance. Differing national priorities might result in gridlock and impasse on the regional level, as ASEAN is a consensus-based organization.[25] Yet potentially powerful synergies exist. For instance, while Vietnam's data protection framework relies on a consent-centric approach, Singapore's possible role as a data management sandbox could offer insights into useful regulatory strategies.

4. **Persistent shared challenges confront national-level AI risk management.** National AI safety governance in Southeast Asia faces a unique set of region-specific challenges.

While most nations grapple with these difficulties, some are especially vulnerable to particular issues.

- **Lack of quality data sets.** Good-quality, standardized statistics for informed, data-driven, and evidence-based decisionmaking is absent or collected unevenly in several countries such as Indonesia.[26] Various national languages of Southeast Asian countries, like Khmer in Cambodia, Lao in Laos, and Burmese in Myanmar, are low-resource languages, complicating the ability of developers to create AI assistants in these languages. Beyond these national languages, data on smaller, regional languages is even more difficult to access.[27]

- **Poor cybersecurity.** Indonesia, for example, faces significant challenges in protecting its data.[28] Brunei and Cambodia have recently taken steps toward personal data pro-

---

25 Ongoing negotiations on the Digital Economy Framework Agreement (negotiations of which are slated to end by 2025 with provisions related to AI standards harmonization), have reportedly been slowed by varying levels of understanding and readiness among member states. "Demystifying DEFA: Policy and Political Drivers Across ASEAN," Southeast Asia Public Policy Institute, August 26, 2024, https://seapublicpolicy.org/work/demystifying-defa-policy-and-political-drivers-across-asean/.

26 "Data standardization poses challenge for Satu Data Indonesia plan," The Jakarta Post, August 6, 2020, sec. Business, https://www.thejakartapost.com/news/2020/08/06/data-standardization-poses-challenge-for-satu-data-indonesia-plan.html.

27 "Speaking in Code: Contextualizing Large Language Models in Southeast Asia," Carnegie Endowment for International Peace, January 6, 2025, https://carnegieendowment.org/research/2025/01/speaking-in-code-contextualizing-large-language-models-in-southeast-asia?lang=en.

28 Yanuar Nugroho, "Indonesia's National Data Centre Ransomware Attack: A Digital Governance Failure?," FULCRUM (blog), August 8, 2024, https://fulcrum.sg/indonesias-nation-

tection in the past years.[29] Myanmar and Timor-Leste lack consolidated data protection laws altogether.

- **Constrained budgetary and governance capacity.** Developing countries in the region have other pressing priorities like basic infrastructure provision, leaving AI safety governance on the backburner despite a recognition of its importance.

- **Limited technical capacity and human capital.** Low wages and the lack of talent development pipelines stymie the growth of AI talent pools in countries like the Philippines.[30] Such experts are necessary to create thriving AI ecosystems of competent regulators, designers, deployers, and adopters.

- **Infrastructure gaps.** Malaysia, Thailand, and Singapore are racing ahead with data center construction and have solid digital infrastructure.[31] Other countries, like Indonesia, face significant disparities in digital and electrical infrastructure between rural and urban areas.[32] Meanwhile, Myanmar, Laos, and Timor-Leste face infrastructural challenges across the entire country.

- **Regulatory fragmentation.** Malaysia, Thailand, and Indonesia have AI regulatory structures that require greater definition. The lack of centralization, combined with overlapping mandates among agencies, might lead to policy duplication and inefficiencies. Personality-based politicking might also stand as a barrier to efficient AI safety governance. The Philippines has a complex patchwork of soft-law regulation that provides uneven coverage across AI use cases.

- **Less-than-desired level of implementation despite ambitious policy targets.** Indonesia's Satu Data initiative is an example of a well-intentioned and much-needed project to increase data collection that ran out of steam due to uneven implementation.[33] Other countries, such as Thailand, are susceptible to changes in political administration,

---

al-data-centre-ransomware-attack-a-digital-governance-failure/.

29 "Brunei enacts new law giving citizens control over personal data," The Scoop, March 8, 2025, https://thescoop.co/2025/03/08/brunei-enacts-new-law-giving-citizens-control-over-personal-data/.

30 "IT talent development and retentional," Philippine Institute for Development Studies, February 24, 2025, https://www.pids.gov.ph/details/news/in-the-news/it-talent-development-and-retention.

31 "Malaysia, Thailand, and Japan set pace on APAC data centre construction per capita through end of decade," Cushman & Wakefield, June 12, 2025, https://www.cushmanwakefield.com/en/singapore/news/2025/06/malaysia-thailand-japan-set-pace-on-apac-data-centre-construction-per-capita-through-end-of-decade.

32 Mona Siahaan, "Electrification Rate in Indonesia in 2023, by Province," Statista, accessed July 2, 2025, https://www.statista.com/statistics/865193/indonesia-electrification-rate-by-region/.

33 "Data standardization poses challenge for Satu Data Indonesia plan."

raising concerns about the policy continuity.

Southeast Asian countries also have differing levels of participation in global summits, as illustrated in Table 2, consolidating the footprint of Southeast Asian countries in major international AI governance forums.

TABLE 2

Footprint of Southeast Asian countries in international AI governance[34]

Country	2023 U.K. AI Safety Summit	2024 Seoul AI Summit	2025 Paris AI Action Summit	ISO/IEC JTC AI Committee
Singapore	✓	✓	✓	✓
Indonesia	✓	×	✓	×
Malaysia	×	×	×	✓
Thailand	×	×	✓	×
Vietnam	×	×	~ (did not sign agreement)	×
Philippines	✓	✓	~ (participated in side events)	✓
Laos	×	✓	×	×
Myanmar	×	×	×	×
Brunei	×	×	×	×
Cambodia	×	×	✓	×
Timor-Leste	×	×	×	×

NOTE: Light blue indicates presence, orange indicates absence, and dark orange indicates caveat.

34 Admittedly, judging international participation in AI safety governance discourse by only looking at participation in four forums might run the risk of oversimplifying the situation. These forums were chosen as rough barometers of international engagement as (i) the global AI summits are the more publicized forums on the topic, and (ii) the ISO and IEC are the main international standard setting organizations. Wael William Diab and Mike Mullane, "How the ISO and IEC Are Developing International Standards for the Responsible Adoption of AI," UNESCO, How the ISO and IEC Are Developing International Standards for the Responsible Adoption of AI (blog), August 2, 2024, https://www.unesco.org/en/articles/how-iso-and-iec-are-developing-international-standards-responsible-adoption-ai, Gustav R. Grob, "Importance of ISO and IEC International Energy Standards and a New Total Approach to Energy Statistics and Forecasting," Applied Energy, Energex 2002 - Energy Policies, Wael William Diab and Mike Mullane, "How the ISO and IEC Are Developing International Standards for the Responsible Adoption of AI," UNESCO, How the ISO and IEC Are Developing International Standards for the Responsible Adoption of AI (blog), August 2, 2024, https://www.unesco.org/en/articles/how-iso-and-iec-are-developing-international-standards-responsible-adoption-ai; Gustav R. Grob, "Importance of ISO and IEC International Energy Standards and a New Total Approach to Energy Statistics and Forecasting," Applied Energy, Energex 2002 - Energy Policies and Economics and Rational Use of Energy of Energy Topics VI and VII, 76, no. 1 (September 1, 2003): 39–54, https://doi.org/10.1016/S0306-2619(03)00045-X.

On international AI safety governance affairs, with the notable exception of Singapore, Southeast Asian countries have a small footprint. Participation in the 2023, 2024, and 2025 iterations of global AI summits in the U.K., Seoul, and Paris acts as a barometer of the level of international engagement. While engagement does seem to be slowly improving—Indonesia, Thailand, Vietnam, Cambodia, and Timor-Leste (in addition to Singapore) attended the 2025 AI Action Summit in Paris—international participation has tended to be inconsistent in the past (see the country profiles on the international engagements of Indonesia and the Philippines).

Yet, the lack of participation does not mean an unawareness of the importance of participating in the global discourse on AI regulation. Vietnam, for example, has been aware of the need to cooperate with international organizations on standards setting. Multiple government officials in our roundtable discussions have also indicated the importance of having their country be represented at international AI forums. The absence of participation by these countries may not necessarily indicate unwillingness but rather stem from factors such as limited administrative capacity or a focus on more pressing developmental priorities.

Regional engagement on AI safety governance at ASEAN seems to be a more active arena for Southeast Asian countries to participate in multilateral collaboration.

- **Malaysia**, the 2025 ASEAN Chair, has been vocal about its desire to push for initiatives such as an ASEAN AI Safety Network (ASEAN AI SAFE) and an ASEAN Guide for Sustainable Data Centre Development.[35]

- The **Philippines** has also signaled its intention to propose a regional AI regulatory framework based on its domestic draft legislation during its 2026 chairmanship.[36]

- Meanwhile, **Singapore** leads the ASEAN Working Group on AI Governance, the primary entity tasked with addressing AI governance issues in the organization and has been a key proponent of major regional policy documents like the initial and extended versions of the Guide on AI Governance and Ethics.[37]

35 "Secretary-General of ASEAN Meets with Minister of Digital of Malaysia," ASEAN Main Portal (blog), January 16, 2025, https://asean.org/secretary-general-of-asean-meets-with-minister-of-digital-of-malaysia/.

36 "Philippines to Propose ASEAN AI Regulatory Framework, House Speaker Says," Reuters, January 17, 2024, sec. Technology, https://www.reuters.com/technology/philippines-propose-asean-ai-regulatory-framework-house-speaker-says-2024-01-17/.

37 Singapore Concludes Fruitful Chairmanship of the ASEAN Digital Ministers Meeting (Ministry of Digital Development and Information, 2025), https://www.mddi.gov.sg/newsroom/singapore-concludes-fruitful-chairmanship-of-adgmin/.

KEY TAKEAWAYS: REGIONAL STATE OF PLAY

- Six countries (Singapore, Indonesia, Malaysia, Thailand, Vietnam, and the Philippines) are more mature in their AI safety governance policy trajectory. Of this group, Singapore stands out as a very mature AI ecosystem, while the remaining five are fast followers. Laos, Myanmar, Brunei, Cambodia, and Timor-Leste lag behind.
- The regional state-of-play of AI safety governance can be characterized by:
    - A universal recognition of AI significance, but nuanced differences in perspectives
    - A pragmatic approach to policy development that leverages hindsight, localizes policy frameworks from abroad, and balances safety and innovation in a nuanced manner
    - Diversity as a double-edged sword for regional collaboration
    - Persistent shared challenges confronting national-level AI risk management.
- The region has a small footprint in international AI safety governance affairs. Yet, the lack of participation does not indicate an unawareness of the importance of participating. Regional engagement on AI safety governance at ASEAN seems to be a more active arena in this field.

2.2. BRUNEI

After providing a snapshot of the regional state-of-play, the following sections summarize country-specific AI safety initiatives on the domestic and international levels. These country profiles are not meant to be comprehensive, but reflect salient points raised by speakers at AISA's roundtable series. Detailed information on the speakers and the roundtable data collection methodology can be found in Appendices A and B respectively.

2.2. BRUNEI

See roundtable session here (February 2025)

Although a latecomer to AI governance, Brunei has made cautious yet meaningful strides in developing a foundational framework. Its policy approach is distinctly pragmatic, leveraging hindsight from international best practices. This strategy allows Brunei to develop policy efficiently while conserving government resources. Led by the Authority for Info-communications Technology Industry (AITI), the current focus is on creating voluntary, nonbinding guidelines. However, Brunei's presence in global and regional AI governance discussions remains limited.

1. Timeline highlights

- **2024:**
    - In May, the AITI formed the AI Governance and Ethics (AIGE) Working Group bringing together 25 experts from government, academia, and industry.[38]
    - In July, the AIGE Working Group released a Draft Guide on AI Governance and Ethics for public comment.[39]

2. Governance landscape

Strategic policies

- None announced. Brunei has not yet announced a national AI strategy or a specific economic blueprint for AI.

Regulatory and legal instruments

38 "AITI-AI Governance and Ethics," Authority for Info-communications Technology Industry, accessed July 2, 2025, http://www.aiti.gov.bn/regulatory/ai-governance-and-ethics/.

39 "Draft Guide on Artificial Intelligence (AI) Governance and Ethics for Brunei Darussalam" (Authority for Info-communications Technology Industry of Brunei Darussalam, July 9, 2024), https://www.aiti.gov.bn/media/eysd0thx/draft-guide-on-ai-governance-ethics-for-public-consultation_9july2024.pdf.

- Guide on AI Governance and Ethics (2025). Released in April 2025, this is the country's primary AI governance document to date.[40] It is a nonbinding, voluntary guide that proposes seven general principles for responsible AI along with a governance framework for organizations to implement the principles, borrowing from ASEAN-level documents. Seventeen stakeholders from across industry, academia, and civil society provided input to the draft, reflecting a growing national discourse on AI governance.

- Cybersecurity legislation (planned). Panelists mentioned that Brunei was currently working on cybersecurity legislation that aims to adapt best practices from the region and wider world. For instance, U.K.'s "security-by-design" principles for AI cybersecurity were mentioned to have been a best practice that authorities were considering.

Key institutions and research hubs

- Authority for Info-communications Technology Industry (AITI). The primary government agency spearheading AI safety and governance initiatives.

- AI Governance and Ethics (AIGE) Working Group. The multi-stakeholder body established by AITI, composed of experts from government, academia, and industry, to draft the national AI governance framework.

3. International engagement

- Limited participation. Brunei's involvement in international AI governance is limited. It did not participate in the major global AI Summits in the U.K. (2023), Seoul (2024), or Paris (2025). Its involvement in regional AI governance policy is similarly muted.

4. Key insights and challenges

- Pragmatic latecomer advantage. Brunei strategically uses its position as a latecomer to its advantage. By observing and adapting established best practices, it can create policy frameworks in a highly resource-efficient and rapid manner, avoiding the trial-and-error

"There is value first **looking at baseline standards across the world,** then **comparing that to what we have in the region,** and from there, trying to **develop our own AI standards and guidelines** that fit our national objectives."

Isyrah Fahmi Osman, Head of R&D, Cyber Security Brunei, Bandar Seri Begawan Roundtable

40 Logan Carnicelli, Mega Valentina, Christopher Zoromski, "Brunei Releases Voluntary AI Guidelines To Enhance Responsibility and Trust," US-ASEAN Business Council, April 30, 2025, https://www.usasean.org/article/brunei-releases-voluntary-ai-guidelines-enhance-responsibility-and-trust.

costs faced by first-movers. For instance, the Draft Guide on AI Governance and Ethics was developed by the AIGE Working Group within three months and bears similarities to the seven principles of responsible AI development, deployment, and use published in the ASEAN AI Guide on AI Governance and Ethics. Panelists noted that such an approach was not forsaking national sovereignty in policy development, but rather a strategic adaptation that tailors global best practices to the local context.

- Challenge of limited influence. The direct consequence of its size and latecomer status is a limited ability to shape global AI norms. Its focus remains on domestic application and compliance rather than international standard setting.

2.3. CAMBODIA

2.3. CAMBODIA[41]

See roundtable session here (March 2025)

Cambodia is gradually developing its AI safety and governance framework, driven by a pragmatic recognition of its role as a technology consumer rather than a developer.[42] Spearheaded by the Ministry of Post and Telecommunications (MPTC), the government is focused on establishing foundational policies for cybersecurity and data protection as prerequisites for safe AI adoption. While formal regulation is still in early stages, nongovernmental actors and international partnerships are playing a crucial role in building local capacity, developing Khmer-language AI tools, and increasing Cambodia's presence in regional and global governance discussions.

1. Timeline highlights

- **2021:** The Cambodia Digital Economy and Society Policy Framework (2021–2035) is released, identifying the utility of AI for public benefit and acknowledging technology risks.[43]
- **2022–2023:**
    - The MPTC releases the Draft Law on Cybersecurity and Draft Law on Personal Data Protection for review.[44]
    - A Ministry of Industry, Science, Technology, and Innovation (MISTI) report on the "AI Landscape in Cambodia" recommends a multi-ministry approach.[45]
- **2024:**

41 In addition to the authors of this paper, Quentin Flament also contributed to this section.

42 "Consumer Electronics - Cambodia | Statista Market Forecast," Statista, accessed June 30, 2025, http://frontend.xmo.prod.aws.statista.com/outlook/cmo/consumer-electronics/cambodia.

43 "Cambodia Digital Economy and Society Policy Framework 2021-2035" (Ministry of Post and Telecommunications, June 14, 2023), https://mptc.gov.kh/en/2023/06/cambodia-digital-economy-and-society-policy-framework-2021-2035/.

44 Kang Sothear, "MPTC Finalises More Draft Laws, Policies on Cybersecurity, Posts," Khmer Times, November 6, 2022, https://www.khmertimeskh.com/501180686/mptc-finalises-more-draft-laws-policies-on-cybersecurity-posts/.

45 "AI Landscape in Cambodia: Current Status and Future Trends" (Phnom Penh, Cambodia: Ministry of Industry Science, Technology & Innovation, May 2023), https://misti.gov.kh/public/file/202305301685426285.pdf.

- Cambodia launches a partnership with UNESCO to conduct its AI Readiness Assessment Methodology (RAM).[46]

- A forum cohosted by MISTI and the United Nations Economic and Social Commission for Asia and the Pacific (UNESCAP) marks the start of work on a national AI strategy.[47]

- **2025:**

  - AI Forum Cambodia signs a memorandum of understanding (MOU) with AI Singapore to collaborate on the development of an open-source Khmer LLM.[48]

  - Cambodia attends the Paris AI Action Summit.

2. Governance landscape

Strategic policies

- Cambodia Digital Economy and Society Policy Framework (2021–2035). The primary high-level policy document that outlines the government's approach. It calls for data-driven governance but stops short of specifying ethical standards for AI.[49]

- Draft National AI Strategy (2025). Released for input from stakeholders in June 2025, the Draft outlines six strategic priorities comprising 41 strategic measures, with the strategic priorities being human resource development, data and infrastructure, AI for digital government, sectoral AI adoption and development, ethical and responsible AI, and collaborative R&D. Currently in the drafting process with the help of UNESCAP, this strategy aims to create a formal governance framework that promotes economic development while safeguarding ethical priorities.[50] A national workshop to finalize

46 "Initiating AI Readiness Assessment: Shaping Cambodia's Future in Technology," UNESCO, Initiating AI Readiness Assessment: Shaping Cambodia's Future in Technology (blog), November 19, 2024, https://www.unesco.org/en/articles/initiating-ai-readiness-assessment-shaping-cambodias-future-technology.

47 Ben Sokhean, "MISTI, UN-ESCAP Enhance Private Sector Engagement in STI," Khmer Times, May 12, 2024, https://www.khmertimeskh.com/501487246/misti-un-escap-enhance-private-sector-engagement-in-sti/.

48 Nop Sreymao, "Khmer Language Gets AI-Infusion," Khmer Times, January 21, 2025, https://www.khmertimeskh.com/501627225/khmer-language-gets-ai-infusion/.

49 "Cambodia Digital Economy and Society Policy Framework 2021–2035."

50 Rajah, Tann Asia-Heng Chhay, and Tiv Sophonnora, "Announcement on the Progress of Draft National Artificial Intelligence Strategy and the Opening of Public Consultation," Lexology, June 19, 2025, https://www.lexology.com/library/detail.aspx?g=5efbb800-2588-49ce-a2f6-cdb72a50c729.

the Draft was scheduled for July 2025.[51]

Regulatory and legal instruments

- None specific to AI. Cambodia currently has no AI-specific hard or soft regulations.
- Key draft legislation. The Draft Law on Cybersecurity (2022) and the Draft Law on Personal Data Protection (2023) are currently under review and are regarded as crucial foundations for the development of future comprehensive AI policies.

Key institutions and research hubs

- Ministry of Post and Telecommunications (MPTC). The lead government agency spearheading AI policy, supported by the Digital Government Committee as the primary interministerial mechanism for AI governance.
- Cambodia Academy of Digital Technology (CADT). A crucial government-backed institution for fostering AI talent, organizing regional conferences, and guiding policy development.
- AI Forum Cambodia. A key nongovernmental organization taking the lead in developing an open-source Khmer LLM and fostering local skill development. It signed a memorandum of understanding with AI Singapore in 2025 on LLM development.

3. International engagement

- Increasing multilateral presence. While absent from the 2023 and 2024 AI summits, Cambodia attended the 2025 Paris AI Action Summit and is a member of the Hiroshima AI Process Friends Group, signaling a growing commitment to international governance.
- Regional participation. Cambodia actively participates in ASEAN-level bodies, including the ASEAN WG-AI and ASEAN Digital Senior Officials' Meeting (ADGSOM).
- Developmental partnerships. Collaborates closely with international bodies like UNESCO (on the RAM process), UN-ESCAP (on AI readiness), and partners like AI Singapore (on LLM development) to build domestic capacity.

4. Case studies

- Khmer-language tools (Translatekh, Sarika). The MPTC has launched several AI tools to foster digital transformation. Translatekh provides translation between Khmer and English, while Sarika offers text-to-audio functionality. These tools, along with an award-winning

51 "Announcement on the Progress of Draft National Artificial Intelligence Strategy and the Opening of Public Consultation"

Khmer Braille translation project, aim to democratize access to information.

- Open-source Khmer LLM. A collaborative project between AI Forum Cambodia and AI Singapore, this initiative is a direct response to the challenge of limited local-language data. Its goal is to create an LLM that accurately reflects Cambodia's sociocultural context and ensures equitable access to AI for Khmer speakers.

5. Roundtable top insights

- Limited Khmer data. A significant technical barrier is the scarcity of high-quality data in Khmer to train AI systems. This "low-resource language" issue risks entrenching biases and making it difficult to develop culturally relevant AI models. Cognizant of this challenge, Cambodian organizations have already started collaborating with international partners like AI Singapore to develop and adopt best practices.[52]

- Role as a technology consumer. Cambodia recognizes its limited role as a developer in the global AI ecosystem. This pragmatism informs its regulatory approach, which focuses on governing the use of AI within its borders rather than imposing heavy burdens on designers and developers.

- Pragmatic focus on near-term risks. The roundtable discussion highlighted a consensus that for a technology consumer like Cambodia, the most pressing concerns are the "social risks of AI." These are less technical but more difficult to define and manage, requiring a focus on local context and practical governance solutions.

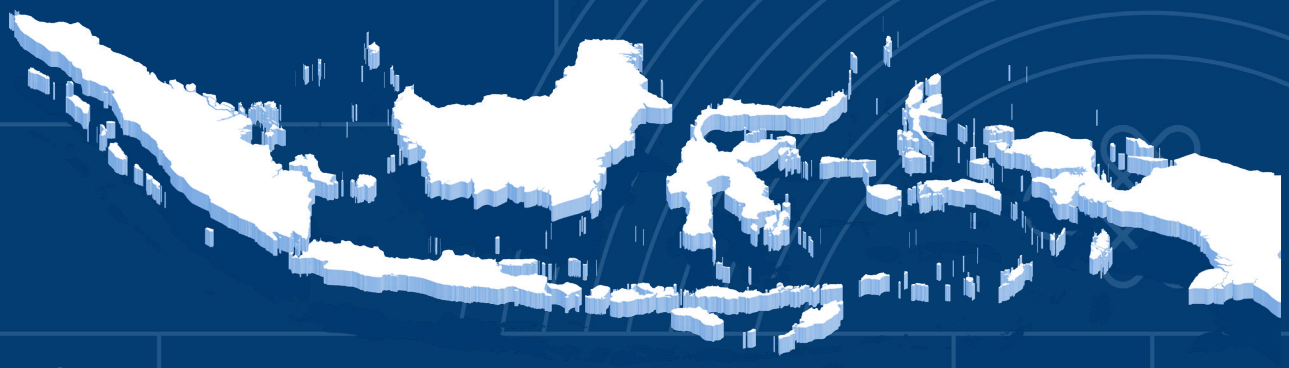
- Prioritize policy implementation. Panelists emphasized the need to move beyond high-level strategy to tangible, results-driven implementation, particularly for clear, achievable solutions like public education programs and job reskilling initiatives.

"The private sector focuses on the more technical side of risk, like reliability, validity, robustness, and resilience. As consumers [rather than developers or deployers], the risk areas that are of higher concern to us are the **social risks of AI**... These are risks that are less technicaly but are more difficult to define concisely."

Bunhong Taing, Director for Policy & Legal Affairs, Vice-Chair of AI Policy Sub-Working Group, Ministry of Post and Telecommunications, Cambodia, Phnom Penh Roundatable

52 Sreymao, "Khmer Language Gets AI-Infusion."

2.4. INDONESIA



2.4. INDONESIA[53]

See roundtable session here (September 2024)

With its large, youthful population and a vibrant tech sector,[54] Indonesia is well-positioned to benefit from AI. However, significant challenges mark its path. The government has a national AI strategy and has issued initial soft-law guidelines, with signals that more comprehensive, omnibus-style regulation is forthcoming. Yet, critical and persistent issues with data management, severe cybersecurity vulnerabilities, infrastructure gaps between urban and rural areas, and the risk of AI-exacerbated exploitation in its large gig economy persist. While international engagement has been inconsistent, the country is beginning to take steps, such as completing a UNESCO readiness assessment, to build a more robust governance framework.

1. Timeline highlights

- **2020:** The National Strategy for AI 2020–2045 is articulated.[55]
- **2022:** The Personal Data Protection Law is enacted.[56]
- **2023:**

53 In addition to the authors of this paper, Nitya Kuthiala also contributed to this section.

54 Indonesia boasts a population of over 280 million people, over 1200 distinct ethnic groups, and more than 694 local languages and dialects (in addition to Bahasa Indonesia as the national language) spread across 38 provinces. It also has approximately 21,000 startups—the sixth highest volume in the world. Shofa, Jayanty Nada. "Indonesia ranks sixth among countries with most startups." Jakarta Globe, June 21, 2023, https://jakartaglobe.id/tech/indonesia-ranks-sixth-among-countries-with-most-startups. Further application of AI in the country's financial services, retail, and logistics sector is expected to add USD$ 366 billion to its GDP in the next decade. "Indonesian Gov't to Complete AI Roadmap by June 2025," Tech in Asia, June 3, 2025, sec. Artificial Intelligence, https://www.techinasia.com/news/indonesian-govt-to-complete-ai-roadmap-by-june-2025; Jayanty Nada Shofa, "Indonesia Ranks Sixth Among Countries with Most Startups," Jakarta Globe, June 21, 2023, https://jakartaglobe.id/tech/indonesia-ranks-sixth-among-countries-with-most-startups.

55 "Indonesia Artificial Intelligence," Asia Society Policy Institute, July 12, 2022, https://asiasociety.org/policy-institute/raising-standards-data-ai-southeast-asia/ai/indonesia.

56 "Indonesia's Comprehensive Personal Data Protection Law Guide - Indonesia Guide | Doing Business in Indonesia," ASEAN Briefing, accessed July 2, 2025, https://www.asean-briefing.com/doing-business-guide/indonesia/company-establishment/personal-data-protection-law.

- Participates in the U.K. AI Safety Summit.[57]

- The Ministry of Communication and Digital Affairs (Komdigi) issues Circular Letter No. 9 on AI Ethical Guidelines.[58]

- **2024:** The Indonesia AI Readiness Assessment Report (in collaboration with UNESCO) is completed.[59]

- **2024–2025 (Planned):** Government signals intent to introduce new, comprehensive AI regulations.

- **2025:** Participates in the Paris AI Action Summit and signs the outcome declaration.[60]

2. Governance landscape

Strategic policies

- National Strategy for AI, 2020–2045. Articulated by the National Research and Innovation Agency (BRIN) in 2020, this strategy aims to guide AI development toward the twin goals of making Indonesia a significant global AI player and ensuring AI is ethical and inclusive. It identified "Ethics and Policy" as a foundational focus area, recommending the formation of a data ethics board to oversee AI development, and for the government to establish regulations and standards for AI innovation.[61]

Regulatory and legal instruments

- No comprehensive AI law: Indonesia currently lacks omnibus legislation for AI, although there has been talk of new comprehensive AI legislation slated for 2024–2025 that would tie the existing regulatory framework together.[62] Governance has so far been guided by

57 "The Bletchley Declaration by Countries Attending the AI Safety Summit, 1–2 November 2023."
58 Michelle Virgiany, Sakurayuki, and Naila Amatulla, "Ethical Guidelines on Use of Artificial Intelligence (AI) in Indonesia," Hiswara Bunjamin & Tandjung, February 7, 2024, https://www.hbtlaw.com/insights/2024-02/ethical-guidelines-use-artificial-intelligence-ai-indonesia.
59 "Indonesia: Artificial Intelligence Readiness Assessment Report" (UNESCO, 2024), https://unesdoc.unesco.org/ark:/48223/pf0000392317.
60 "Statement on Inclusive and Sustainable Artificial Intelligence for People and the Planet.," February 11, 2025, https://www.elysee.fr/en/emmanuel-macron/2025/02/11/statement-on-inclusive-and-sustainable-artificial-intelligence-for-people-and-the-planet.
61 "Indonesia Artificial Intelligence."
62 "Indonesia: Kominfo Announces Government Approach for AI Governance | News," DataGuidance, March 7, 2024, https://www.dataguidance.com/news/indonesia-kominfo-announces-government-approach-ai.

sector-specific soft law.

- Circular Letter No. 9 on AI Ethical Guidelines (2023). Issued by the Ministry of Communication and Digital Affairs, this is the primary initial guide for the AI ecosystem. It highlights values like inclusivity, safety, and accountability and introduces approaches for AI regulation, including data protection and strong oversight.

- Financial Services Authority (OJK) guidance. The OJK has also issued sector-specific guidance on responsible and trustworthy AI for the financial industry.[63]

Key Institutions & Research Hubs

- Ministry of Communication and Digital Affairs (Komdigi): The lead ministry issuing AI guidelines and working with UNESCO on readiness assessments.

- National Research and Innovation Agency (BRIN): The agency that formulated the national AI strategy.

3. International engagement

- Limited and inconsistent. Indonesia's international engagement is rather limited and inconsistent. It participated in the 2023 U.K. and 2025 Paris AI summits but was absent from the 2024 Seoul summit.

- Regional role. On the regional level, it has been active in leading ASEAN discussions on digital economic integration but has not yet presented strong positions on emerging technology governance.

- Developmental partnerships. Indonesia is the first country in Southeast Asia to have completed a UNESCO Readiness Assessment Methodology (RAM) evaluation.[64] Completed in late 2024, the post-RAM report highlighted:

    - Inequalities that could be perpetuated by the increasing adoption of AI, such as the gender gap and regional differences

    - The existence of a research funding gap

63 Virgiany, Sakurayuki, and Amatulla, "Ethical Guidelines on Use of Artificial Intelligence (AI) in Indonesia."

64 "Indonesia: Artificial Intelligence Readiness Assessment Report." "UNESCO and KOMINFO Completed AI Readiness Assessment: Is Indonesia Ready for AI?" UNESCO, October 9, 2024, https://www.unesco.org/en/articles/unesco-and-kominfo-completed-ai-readiness-assessment-indonesia-ready-ai?.com.

- The existence of a talent gap. It also recommended the establishment of a national agency to coordinate AI-related policies and standards.

4. Case studies

- Surabaya's AI-powered traffic management system. The city of Surabaya successfully implemented an AI system that uses computer vision and machine learning to analyze real-time CCTV data.[65] By automatically adjusting traffic light signals, the system significantly reduces congestion, improves air quality, and enhances public safety, serving as a model for AI's contribution to environmental sustainability.

5. Roundtable top insights

- Poor standard of data collection and data paradox. Indonesia has massive amounts of data, but its quality and standardization are poor. Former President Joko Widodo's Satu Data (One Data) Initiative launched in 2019 sought to standardize data across regions,[66] but results have been modest.[67] Further, care should be taken to ensure representation of Indonesia's many indigenous languages to strengthen language preservation and prevent discrimination against minority groups. Lackluster cybersecurity is another major issue,[68] highlighted by a major cyberattack in June 2024 that impacted 282 government agencies with only 2 percent of the breached data having been recovered thus far.[69]

---

65 "TrafiCamTM Sensors Help Meet Indonesia's Ambitious Traffic Management Plans in Jakarta and Surabaya," Teledyne Flir (blog), June 24, 2017, https://www.flir.com/discover/traffic/roads-tunnels/traficam-sensors-help-meet-indonesias-ambitious-traffic-management-plans-in-jakarta-and-surabaya/?srsltid=AfmBOorcfKbVCY1HU-9yvcoAWpB6H_glkfO2BP-8DJ6OgnPE7_Rsc113m.

66 "Implement One Data Indonesia Policy (ID0113)," Open Government Partnership, accessed July 2, 2025, https://www.opengovpartnership.org/members/indonesia/commitments/ID0113/.

67 "Data standardization poses challenge for Satu Data Indonesia plan," The Jakarta Post, August 6, 2020, sec. Business, https://www.thejakartapost.com/news/2020/08/06/data-standardization-poses-challenge-for-satu-data-indonesia-plan.html. Ausma Bernot, Dian Tjondronegoro, Bahtiar Rifai, Rakibul Hasan, Alan Wee-Chung Liew, Tom Verhelst, & Milind Tiwari, "Institutional Dimensions in Open Government Data: A Deep Dive Into Indonesia's Satu Data Initiative and Its Implications for Developing Countries," Public Performance & Management Review 47(6), 2024, https://www.tandfonline.com/doi/full/10.1080/15309576.2024.2377609.

68 Kevin C. Desouza and Marc E. Barda Picavet, "Realising Trustworthy and Inclusive Artificial Intelligence for Indonesia" (IBM Center for the Business of Government, August 2024), https://www.businessofgovernment.org/sites/default/files/Realising%20Trustworthy%20and%20Inclusive%20Artificial%20Intelligence%20for%20Indonesia.pdf.

69 Yanuar Nugroho, "Indonesia's National Data Centre Ransomware Attack: A Digital Governance Failure?," FULCRUM (blog), August 8, 2024, https://fulcrum.sg/indonesias-national-data-centre-ransomware-attack-a-digital-governance-failure/.

- Gig economy worker exploitation. Panelists noted the effects of AI on worker exploitation in Indonesia's large gig economy as a key area of concern. The integration of AI into operations could worsen existing issues by imposing unrealistic productivity targets through opaque systems, increasing the vulnerability of workers.

- Limited resources and infrastructure gaps. Indonesia faces a constrained fiscal space, limiting funding for AI governance initiatives. This is exacerbated by stark infrastructure disparities. While 4G coverage is high, actual internet usage and computer access remain low, especially in rural areas where panelists pointed out that populations might not be able to afford broadband services. The electrification rate in Highland Papua (14.06 percent) compared to Jakarta (100 percent) illustrates the deep regional inequalities that limit equitable AI adoption.[70] Limited access to the internet, compute, and electricity locks rural groups out from enjoying the benefits of AI adoption, entrenching regional inequalities.

"It is our responsibility to ensure AI is a **tool for good**, guided by **ethical principles** and a **commitment to safety.**"

Gita Wirjawan, former Minister of Trade, Indonesia

70 Mona Siahaan, "Electrification Rate in Indonesia in 2023, by Province," Statista, accessed July 2, 2025, https://www.statista.com/statistics/865193/indonesia-electrification-rate-by-region/.

2.5. LAOS

2.5. LAOS

See roundtable session here (February 2025)

Laos's engagement with AI safety and governance remains nascent, with government efforts focused on foundational priorities like building basic digital infrastructure and literacy. AI policy is underdeveloped and currently embedded within broader digital economy goals rather than standalone risk management frameworks. Facing limited domestic capacity, the country's path forward is defined by a pragmatic, step-by-step approach, heavily reliant on international collaboration to build expertise and avoid being left behind in the regional digital landscape.

1. Timeline highlights

- **2021:** The National Digital Economy Development Vision, Strategy, and Plan is released, outlining ambitions for technology adoption but without specific AI safety provisions.[71]
- **2022:** The Digital Maturity Assessment, supported by the UN Development Program (UNDP), is published, concluding that Laos is in the early stages of digital transformation.[72]
- **2023:** A memorandum of understanding is signed with China's Chongqing College of Electronic Engineering to cooperate on AI and Big Data research.[73]
- **2025:** Near-conclusion of UNESCO RAM and national AI Strategy is expected to be developed.[74]

2. Governance landscape

Strategic policies

- National Digital Economy Development Vision, Strategy, and Plan (2021). The primary

71 "National Digital Economy Development Vision (2021–2040)" (Ministry of Technology and Communications, December 2021), https://mtc.gov.la/index.php?r=site%2Fdetail&id=1084.

72 Pradeep Mukherko, Rajarshi Chatterjee, Amarabati Sen, Harsh Kapoor, Aditi Sambhar, "Digital Maturity Assessment - Lao PDR," UNDP, July 2022, https://www.undp.org/sites/g/files/zskgke326/files/2022-08/UNDP_LaoPDR_DMA_2022.pdf.

73 Jonathan Meadley, "Laos And China Sign MoU to Research Development of Artificial Intelligence," Laotian Times, March 1, 2023, https://laotiantimes.com/2023/03/01/laos-and-china-sign-mou-to-research-development-of-artificial-intelligence/.

74 Roundtable source.

high-level document mentioning AI, which frames the technology as a tool for economic development and public safety but neglects governance or ethical risks.[75]

- Planned AI strategy (2025). The government's first intended foray into creating a dedicated national strategy for AI.

Regulatory and legal instruments

- None. The country has not introduced any hard or soft law specifically for AI safety governance. Broader legal frameworks for data privacy and cybersecurity are also not comprehensive.

Key institutions and research hubs

- Ministry of Technology and Communications (MOTC): The central government body leading digital policy and defining the country's technological ambitions.

3. International engagement

- Limited multilateral participation. Laos has not participated in the major global AI Summits like those in the U.K., South Korea, or Paris, nor is it a member of the ISO/IEC AI standards subcommittee.

- Developmental partnerships. China stands as a key bilateral partner in research and education program, having signed a MOU on AI and Big Data with the country in 2024.[76] Meanwhile, there is also close collaboration with UN agencies like the UNDP and UNESCO for capacity building. Laos' participation in the UNESCO RAM process provided an internationally backed, systematic evaluation of the country's preparedness for ethical AI. The findings are expected to serve as a foundational blueprint for the planned 2025 national AI strategy and other future policies.

4. Roundtable top insights

- Risk of lagging behind. With limited regulatory capacity, a lack of specialized expertise, and few jobs complementary to AI,[77] Laos faces significant challenges in both harnessing AI's benefits and managing its risks. While the government acknowledges immediate threats

---

75 "National Digital Economy Development Vision (2021–2040)."

76 Meadley, "Laos And China Sign MoU to Research Development of Artificial Intelligence."

77 Tristan Hennig and Shujaat Khan, "How Artificial Intelligence Will Affect Asia's Economies," International Monetary Fund, How Artificial Intelligence Will Affect Asia's Economies (blog), January 5, 2025, https://www.imf.org/en/Blogs/Articles/2025/01/05/how-artificial-intelligence-will-affect-asias-economies.

like misinformation, broader systemic risks related to data privacy, and AI-driven economic inequality, these issues remain largely unaddressed due to constrained governance capacity.

- Foundational deficits. Low digital literacy across the population and underdeveloped digital infrastructure remain primary bottlenecks,[78] making it difficult to implement or even discuss advanced AI policies.

- The need to adopt a gradual, step-by-step approach. Panelists strongly advocated for a sequenced approach to technology policy. The immediate priorities would be to build minimum basic digital literacy and infrastructure, focus on curating quality datasets, develop a local talent pool, and pursue more ambitious goals such as building data centers or expanding compute capacity.

- Leverage international partnerships as a lifeline. Given domestic constraints, international collaboration is not just beneficial but essential. Laos can use templates from organizations like ASEAN and leverage the expertise of partners like the UN to bridge its capacity gaps, augment its policymaking process, and access best practices for responsible AI governance.

"It's a multi-step process. Once there's a minimum level of **digital infrastructure** in place, the next effort has to be around **compute capacity** and **good quality data sets,** and then **developing skills and talent.**"

Alex Read, Chief Technical Specialist, UNDP Lao PDR, Vientiane Roundtable

78 "Country Report Lao People's Democratic Republic: Lao Social Indicatory Survey II 2017" (United Nations Children Fund (UNICEF) Government of Lao People's Democratic Republic, March 2020).

2.6. MALAYSIA

2.6. MALAYSIA[79]

See roundtable session here (January 2025)

Malaysia is rapidly emerging as a key player in the regional AI landscape, driven by strong ambitions to become a global AI hub, a recent surge in foreign direct investment in its data center industry, and its abundant natural resources. The government accelerated its AI governance efforts in 2024, launching a suite of policies and institutions aimed at balancing economic growth with ethical considerations. While its international engagement has historically been limited, Malaysia is leveraging its 2025 ASEAN chairmanship to champion regional AI initiatives.[80] The potential for bureaucratic complexity and the need for caution when engaging with Big Tech stand as salient insights from the AISA roundtable discussion.

1. Timeline highlights

- **March 2021:** The Ministry of Science, Technology, and Innovation (MOSTI) releases the National AI Roadmap 2021–2025.[81]
- **July 2021:** The National Fourth Industrial Revolution (4IR) Policy is published, framing AI within a broader economic context.[82]
- **May 2024:** The AI Nexus initiative, Talent Roadmap, and AI Consortium are launched as collaborative platforms between academia, government, and industry.[83]
- **September 2024:** MOSTI releases the National Guidelines on AIGE.[84]

---

79 In addition to the authors of this paper, Nitya Kuthiala also contributed to this section.
80 "ASEAN Chairmanship 2025 – Malaysia," Association of Southeast Asian Nations (blog), December 31, 2024, https://asean.org/asean-chairmanship-2025-malaysia/.
81 "Malaysia National Artificial Intelligence Roadmap 2021–2025" (Malaysian Science, Technology Information Centre, December 1, 2023), https://mastic.mosti.gov.my/publication/artificial-intelligence-roadmap-2021-2025/.
82 Malaysia, ed., "National Fourth Industrial Revolution (4IR) Policy" (Economic Planning Unit, Prime Minister's Department, 2019), https://ekonomi.gov.my/sites/default/files/2021-07/National-4IR-Policy.pdf.
83 Yiswaree Palansamy, "Malaysia Artificial Intelligence Nexus 2024: Pioneering Malaysia's Building of an Ecosystem to Accelerate AI Leadership," Malaymail, May 23, 2024, https://www.malaymail.com/news/malaysia/2024/05/23/malaysia-artificial-intelligence-nexus-2024-pioneering-malaysias-building-of-an-ecosystem-to-accelerate-ai-leadership/136026.
84 "The National Guidelines on AI Governance & Ethics" (Malaysian Science, Technology Information Centre, September 24, 2024), https://mastic.mosti.gov.my.

- **December 2024:** The National AI Office (NAIO) is launched under the Ministry of Digital.[85]

- 2025: Malaysia assumes the rotating ASEAN chairmanship with a stated intention to prioritize AI-related initiatives.[86]

2. Governance landscape

Strategic policies

- National AI Roadmap (2021–2025). The foundational document outlining Malaysia's AI ambitions. It is aligned with UNESCO principles and recognizes seven pillars for responsible AI, including fairness, security, and transparency.[87] The roadmap also signals an intent to adopt a "national AI Ethics Charter" to "[enshrine] the goal of elevating human happiness and quality of life."

- National 4IR Policy (2021). A broader economic strategy that acknowledges the need to manage risks from emerging technologies like AI.

Regulatory and legal instruments

- National Guidelines on AI Governance and Ethics (AIGE, 2024). A set of nonbinding, voluntary guidelines for stakeholders, the AIGE provides a shared responsibility framework, stresses the importance of national and Islamic precepts, and calls for stakeholders to adopt the seven principles from the roadmap.

- Sector-specific regulations. Guidelines for AI use in the public sector are also being developed by the Malaysia Digital Economy Corporation (MDEC) and the National Digital Department (JDN). These guidelines are slated for release in 2025.[88]

- Planned policies (2025 and beyond). The new NAIO is tasked with developing an AI Technology Action Plan (2026–2030), an AI adoption regulatory framework, and a formal AI code of ethics.

---

85 "The National AI Office (NAIO)," MyDIGITAL, accessed July 2, 2025, https://www.mydigital.gov.my/initiatives/the-national-ai-office-naio/.

86 "Malaysia Assumes ASEAN Chairmanship, Aims To Strengthen Regional Cooperation," Office Portal of Ministry of Finance (blog), January 1, 2025, https://www.mof.gov.my/portal/en/news/press-citations/malaysia-assumes-asean-chairmanship-aims-to-strengthen-regional-cooperation.

87 "The National Guidelines on AI Governance & Ethics."

88 "Launch Ceremony of the Public Sector Artificial Intelligence (AI) Adaptation Guidelines" (Ministry of Digital, February 27, 2025), https://www.digital.gov.my/api/file/file/27022025_PRESS%20RELEASE_LAUNCH%20CEREMONY%20OF%20THE%20PUBLIC%20SECTOR%20ARTIFICIAL%20INTELLIGENCE%20(AI)%20ADAPTATION%20GUIDELINES.pdf.

Key Institutions & Research Hubs

- National AI Office (NAIO): Launched in late 2024 as a centralized entity under the Ministry of Digital to coordinate national AI governance, its exact regulatory role is still being defined as analysts are unsure if the office will take on a coordinator or regulatory role.

- Ministry of Science, Technology, and Innovation (MOSTI). Historically the primary ministry for AI policy, it is responsible for the AI Roadmap and AIGE.

- National Technology and Innovation Sandbox (NTIS). Established in 2020, it launched an AI Sandbox in 2024 in collaboration with NVIDIA to provide a controlled environment for AI companies to test innovations with eased regulatory requirements.

- AI Policy and Skilling Lab (2024). A partnership between NAIO and Google to convene experts, policymakers, and stakeholders to develop policy recommendations for secure AI.

3. International engagement

- Limited Summit Participation: Malaysia has not participated in the global AI summits in the U.K., South Korea, or France.

- Active in technical and standards bodies. It is a member of the ISO/IEC JTC 1 on Artificial Intelligence[89] and has contributed to discussions at the Global Partnership on Artificial Intelligence (GPAI) and UNESCO.[90]

- ASEAN leadership (2025). As the 2025 ASEAN chair, Malaysia has expressed its intent to prioritize regional AI initiatives. It has called for ASEAN to adopt AI to combat cybercrime and has initiated discussions on AI use in the defense sector.[91] It has also proposed the ASEAN AI Safety Network as a platform to facilitate AI safety research (see Section 2.13 on ASEAN).

4. Case study

- National AI Technology and Innovation Sandbox (NTIS). The NTIS, and specifically its AI Sandbox, represents a key public-private partnership model for fostering innovation. By

89 "ISO/IEC JTC 1/SC 42 Participation," International Organization for Standardization, accessed July 2, 2025, https://www.iso.org/committee/6794475.html?view=participation.
90 Si Ying Thian, "Malaysia Aims to Be a Global Leader in Responsible AI Innovation," GovInsider (blog), accessed July 2, 2025, https://govinsider.asia/intl-en/article/malaysia-aims-to-be-a-global-leader-in-responsible-ai-innovation.
91 "Accelerating Regional Economic Growth, Building Safer Digital Environment and Enhancing Regional Connectivity Key Takeaways for Malaysia at ADGMIN 2025," Ministry of Digital Malaysia (blog), February 17, 2025, https://digital.gov.my.

easing regulatory hurdles, it provides AI startups a controlled environment to test their products, accelerating the transition from R&D to market readiness. While not explicitly a safety initiative, its structure supports the responsible development and deployment of AI by providing access to essential resources and a framework for governance. The target is to support 900 AI startups by 2026.[92]

5. Roundtable top insights

- Regulatory fragmentation and complexity. The recent and rapid launch of multiple new institutions and initiatives creates a risk of overlapping mandates, bureaucratic inefficiency, and internal politicking, which could undermine the development of a coherent governance framework. Panelists emphasized that while the flurry of new initiatives is a positive sign, their success hinges on effective coordination. A key challenge will be to minimize political turf wars and ensure a cohesive, whole-of-government approach to avoid fragmented and inefficient policymaking.

- Navigating collaboration with Big Tech. A massive influx of foreign investment from Big Tech for data center construction presents both an opportunity and a challenge. While it boosts the economy,[93] roundtable participants noted that Malaysia must manage the significant environmental impact (energy and water consumption) and navigate narratives around "AI and data sovereignty" to ensure its national interests are protected. Given its nascent experience in tech governance, Malaysia should engage these firms with caution to ensure its long-term national and environmental interests are not compromised.

- Economic opportunity vs. AI safety trade-off. The government has so far successfully balanced attracting investment with advancing its governance framework. However, with limited resources, there is a persistent risk that AI safety initiatives could be sidelined in favor of more immediate economic priorities.

"There's a remarkable number of things happening given that [Malaysia] is such a small developing country. **If we're able to coordinate them, we can accomplish many things.** We need to **not work in silos.**"

Dr. Jun-E Tan, Senior Research Associate, Khazanah Research Institute, Kuala Lumpur Roundtable

92 "AI Sandbox Programme Expected to Create 900 Startups and 13,000 New Talents by 2026," The Star, April 18, 2024, https://mranti.my/happenings/news/ai-sandbox-programme-expected-to-create-900-startups-and-13-000-new-talents-by-2026.

93 Dylan Butts, "Malaysia Is Emerging as a Data Center Powerhouse amid Booming Demand from AI," CNBC, June 17, 2024, sec. Technology, https://www.cnbc.com/2024/06/17/malaysia-emerges-as-asian-data-center-powerhouse-amid-booming-demand.html.

2.7. MYANMAR

2.7. MYANMAR[94]

See roundtable session here (March 2025)

Myanmar's development of AI safety governance is virtually nonexistent due to ongoing civil conflict and limited state capacity. Existing national policies view AI through a narrow economic lens, with little attention to safety or ethical risks. The current military government has enacted laws that raise concerns about state overreach and the use of AI for surveillance and suppression. With a minimal international footprint, the country was characterized by panelists as an AI "risk-critical" context. Panelists at the AISA roundtable highlighted that any progress in the local AI safety ecosystem is primarily driven by grassroots initiatives and can be supported by regional frameworks from organizations like ASEAN.

1. Timeline highlights

- **2019**
    - A Draft Cyber Law is released by the civilian government for feedback and review (subsequently abandoned in 2021).[95]
    - The Safe City AI-enabled surveillance project is launched in Mandalay.
- **2021:** A military coup overthrows the civilian government.
- **2024:** The Myanmar Digital Economy Roadmap 2030 and E-Governance Master Plan 2030 are released by the Ministry of Transport and Communications (MOTC).[96]
- **2025:** The Cybersecurity Law 2025 is enacted by the military government.[97]

---

94 In addition to the authors of this paper, Quentin Flament also contributed to this section.

95 "Myanmar Cyber Legal and Policy Framework: Policies Related to e-Government, e-Commerce, and Cyber Security," January 25, 2019, https://www.myanmar-responsiblebusiness.org/pdf/2019-policies-related-to-eGovernment-eCommerce-cyberSecurity.pdf?v=2021.

96 "Myanmar E-Governance Master Plan 2030" (Ministry of Transport and Communication, July 2024), https://motc.gov.mm/sites/default/files/Myanmar%20e-Governance%20Master%20Plan%202030%20(%E1%80%99%E1%80%B0%E1%80%80%E1%80%BC%E1%80%99%E1%80%BA%E1%80%B8).pdf.

97 "Cybersecurity Law" (Republic of the Union of Myanmar State Administration Council, January 1, 2025), https://www.lincolnmyanmar.com/wp-content/uploads/2025/01/Cybersecurity-Law.pdf.

2. Governance landscape

Strategic policies

- Digital Economy Roadmap 2030 and E-Governance Master Plan 2030. These documents, released by the MOTC, frame technology primarily as a tool for economic development. They notably lack any mention of AI safety or ethical governance principles.[98]

Regulatory and legal instruments

- Cybersecurity Law 2025. Enacted by the current government, this law addresses cybersecurity risk management and digital platform regulation. However, analysts have raised concerns that it grants "overbroad powers to the junta,"[99] potentially enabling excessive state control.

- No AI-specific regulation. Myanmar has no dedicated hard or soft laws for AI safety governance.

Key institutions and research hubs

- Ministry of Transport and Communications (MOTC). The lead ministry for digital development, although no agency focusing on AI policy exists.

- Grassroots initiatives. In the absence of state-led efforts, local entities are beginning to explore AI. For example, My Me My Mine Media Agency organized AI Yangon 2025, the country's largest AI expo, to foster technology development in sectors like agriculture and trade and showcase local research.

3. International engagement

- Limited participation. Myanmar's involvement in international AI governance is minimal. It has not participated in key global forums like the U.K., Seoul, or Paris AI Safety Summits. Its limited participation is partially explained by the junta's diplomatic sidelining post-coup.

- Regional engagement. Government officials have quietly attended ASEAN workshops on emerging technologies, though they have not been active in regional AI discussions.

4. Cautionary case study

98 "Myanmar Digital Economy Roadmap" (Digital Economy Development Committee, September 2018), https://myanmar.gov.mm/documents/20143/9096339/2019-02-07+DEDC+RoadMap+for+Websites.pdf/; "Myanmar E-Governance Master Plan 2030."
99 "Myanmar Digital Economy Roadmap."

- Safe City project. The project uses AI-powered surveillance technology for facial and license plate recognition. While officially promoted for public safety, analysts and critics have raised serious concerns that the technology is used by the military government to identify and suppress dissidents.[100] This serves as a stark, cautionary case study on the dual-use nature of AI and the heightened risks of its deployment in a politically fragile, "risk-critical" environment.

5. Roundtable top insights

- "Risk-critical" context. Panelists characterized Myanmar as an AI "risk-critical" country, where violent conflict, political instability, and fragile institutions create a high-risk environment. The population is highly vulnerable to AI-related risks like misinformation[101] and AI-powered surveillance, a situation exacerbated by low levels of digital literacy.

- Dual-use technology concerns. AI technology is actively used by the government for surveillance. The Safe City project, which uses facial and license plate recognition AI, was promoted as a measure for public safety but has been criticized by analysts as a tool for identifying and suppressing dissidents.[102]

- Cyclical barriers to policy development. Panelists noted that Burmese researchers and policymakers face a cyclical challenge. They lack access to "internationally recognized and credible" forums due to political, financial, and linguistic barriers. This leads to limited attention on AI safety as a policy issue, resulting in a lack of data collection. Without data, governance gaps cannot be accurately identified, which in turn limits the incentive and investment needed to address them.

- Grassroots and regional lifelines. With limited central government action, panelists highlighted that any progress in AI safety governance depends on external support. Grassroots movements (like the AI Yangon Expo) and the ASEAN community can act as lifelines. At the

100 "Myanmar: Facial Recognition System Threatens Rights," Human Rights Watch (blog), March 12, 2021, https://www.hrw.org/news/2021/03/12/myanmar-facial-recognition-system-threatens-rights.

101 For example, profiteers capitalized on the magnitude 7.7 earthquakes that struck Myanmar in March 2025 to farm social media engagement through generating clickbait content. Several viral videos carried fabricated depictions of post-earthquake damage, misrepresenting the geographic scope and severity of the disaster. Keith Paolo Catibog Landicho and Karryl Kim Sagun Trajano, "Disasters and Disinformation: AI and the Myanmar 7.7 Magnitude Earthquake," S. Rajaratnam School of International Studies, accessed July 2, 2025, https://rsis.edu.sg/rsis-publication/idss/ip25055-disasters-and-disinformation-ai-and-the-myanmar-7-7-magnitude-earthquake/?doing_wp_cron=1751047175.7470688819885253906250.

102 "Fears of 'Digital Dictatorship' as Myanmar Deploys Artificial Intelligence," The Straits Times, March 19, 2021, https://www.straitstimes.com/asia/se-asia/fears-of-digital-dictatorship-as-myanmar-deploys-artificial-intelligence.

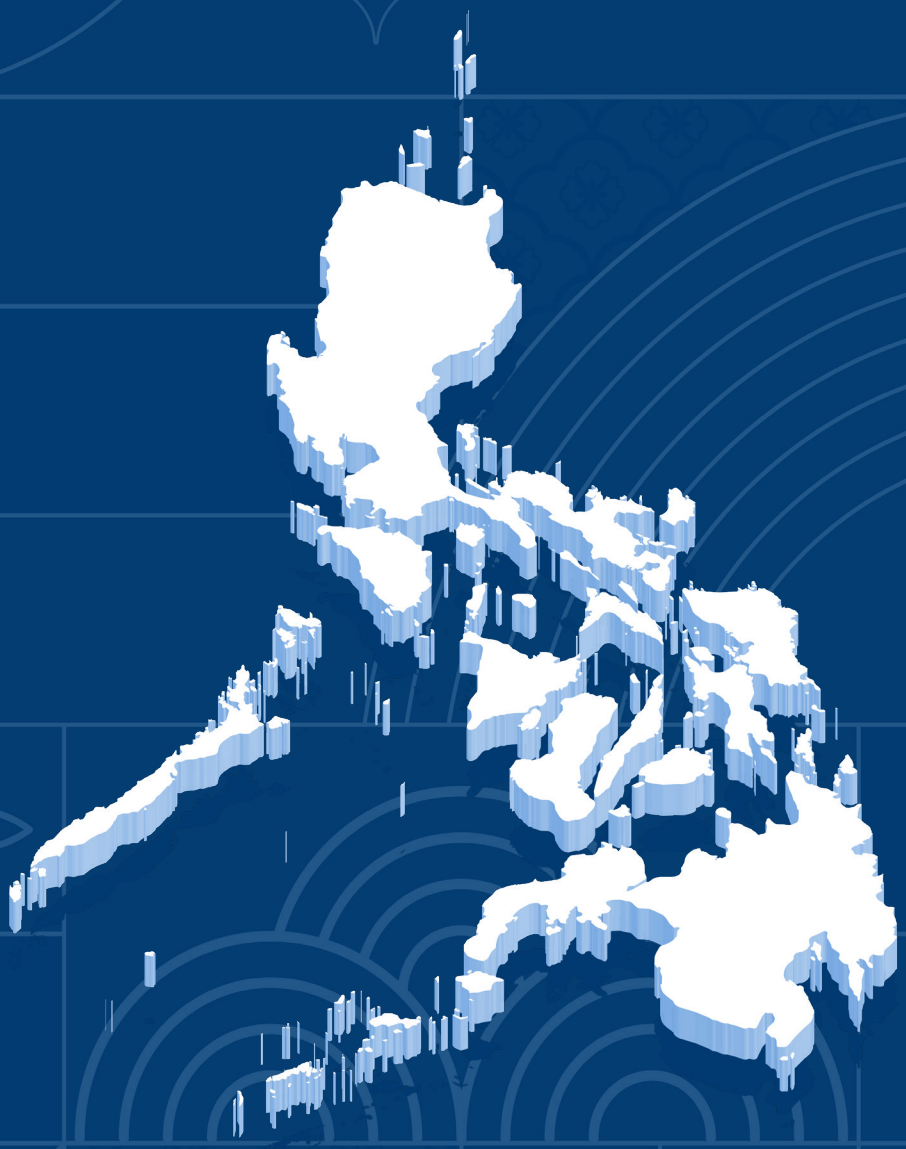
same time, regional frameworks, such as the ASEAN Guide on AI Governance and Ethics,[103] can offer practical templates and valuable insights to buoy Burmese AI safety governance efforts, providing a path forward where domestic state-led initiatives are absent.

"The obstacles that [Myanmar] faces are **cyclical and systemic.** Researchers lack access to internationally recognised spaces. If there's no internationally recognised researched, there's little attention paid to this area, which then leads to insufficient data to explore the area, so no one would invest in the problem because there's no data to prove that this problem exists."

Barani Maung Maung, Tech policy and safety expert, Myanmar Roundtable

103 "ASEAN Guide on AI Governance and Ethics" (Association of Southeast Asian Nations, 2024), https://asean.org/wp-content/uploads/2024/02/ASEAN-Guide-on-AI-Governance-and-Ethics_beautified_201223_v2.pdf.

2.8. PHILIPPINES



2.8. PHILIPPINES

See roundtable session here (November 2024)

The Philippine economy's heavy reliance on the BPO sector[104] makes it uniquely vulnerable to AI-driven job displacement. Cognizant of this risk, Filipino legislators have shown a notable openness to hard law, with several comprehensive bills filed in Congress to regulate AI. This approach contrasts with the softer, voluntary guidelines often preferred by its regional neighbors. Non-state actors are also actively contributing to the governance discourse. Key challenges identified in roundtable discussions include the need to address labor market dynamics in both the formal and informal sectors, refine talent development pipelines, and improve the efficiency of multi-stakeholder policy formulation. On the international front, the Philippines' engagement has been sporadic but shows intent to take on a regional leadership role during its 2026 ASEAN chairmanship.[105]

1. Timeline highlights

- **2021:** The Department of Trade and Industry (DTI) launches the National Artificial Intelligence Strategy Roadmap (NAISR).[106]
- **2023:**
    - HB7396 (AI Development and Regulation Act) is filed in the House of Representatives.[107]

104 BPO and related services constitute a significant 9 percent of Filipino GDP, employing around 1.3 million Filipinos. Across the labor market, an Oxford Economics and Cisco study found that up to 1.1 million jobs will "disappear from the labor market" in the Philippines which might exacerbate pre-existing socioeconomic inequalities. "Outsourcing BPO Services to the Philippines"; "Technology and the Future of ASEAN Jobs: The Philippines" (Oxford Economics, Cisco, 2018), https://www.cisco.com/c/dam/m/en_sg/ai-asean-jobs/files/assets/common/downloads/page0047.pdf.

105 "Myanmar Won't Be Allowed to Lead Association of Southeast Asian Nations in 2026, in Blow to Generals," CNBC, September 6, 2023, sec. Asia Politics, https://www.cnbc.com/2023/09/06/myanmar-wont-be-allowed-to-lead-asean-in-2026-in-blow-to-generals.html.

106 "What's Up at DTI" (Department of Trade and Industry, May 31, 2021), https://dti-webfiles.s3-ap-southeast-1.amazonaws.com/e-library/Main+Publications/What's+Up/2021/What's+Up+No.+13.pdf.

107 "Philippines: Introduced Artificial Intelligence Development and Regulation Act of the Philippines (HB7396) Including Business Registration Requirement," Digital Policy Alert, March 1, 2023, https://digitalpolicyalert.org/event/22797-introduced-artificial-intelligence-development-and-regulation-act-of-the-philippines-hb7396-including-business-registration-requireme-

- The University of the Philippines System launches its Principles for Responsible Artificial Intelligence.[108]
- Participates in the U.K. AI Safety Summit and signs the Bletchley Declaration.

- **2024:**
  - Participates in the South Korea AI Safety Summit and adopts the Seoul Ministerial Statement.
  - DTI launches NAISR 2.0 and establishes the Center for AI Research (CAIR).[109]
- **2025:** Conclusion of UNESCO RAM.[110]
- **2026:** ASEAN rotating chairmanship.

2. Governance landscape

Strategic policies

- National Artificial Intelligence Strategy Roadmap (NAISR, 2021). Updated in 2024 and released by the Department of Trade and Industry (DTI), this roadmap serves as the primary framework to boost the country's AI readiness through policies focused on investment, infrastructure, and implementation. It acknowledges the need to balance innovation with responsible AI regulation, with the ethical consideration of the technology being included in the roadmap.[111]

Regulatory and legal instruments

- Multiple proposed hard laws. In contrast to its neighbors, the Philippines has four broad AI regulation bills under consideration in Congress. Notably, areas like misinformation, data

---

nt.

108 "2023 - UP Principles for Responsible Artificial Intelligence," UPLB Legal Office (blog), 2023, https://legal.uplb.edu.ph/up-policies/2023-up-principles-for-responsible-artificial-intelligence/.

109 "DTI Launches National AI Strategy Roadmap 2.0 and Center for AI Research, Positioning the Philippines as a Center of Excellence in AI R&D," Department of Trade and Industry Philippines (blog), July 3, 2024, https://www.dti.gov.ph/news/dti-launches-national-ai-strategy-roadmap-2-0-center-ai-research-positioning-philippines-center-excellence-ai-rampd/.

110 "Philippines: Artificial Intelligence Readiness Assessment Report," UNESCO, 2025, https://unesdoc.unesco.org/ark:/48223/pf0000393860?posInSet=1&queryId=cb72b22d-9dd3-44cd-9090-c4c89328a09c.

111 "Philippines: Artificial Intelligence Readiness Assessment Report," UNESCO.

protection, and cybersecurity are under-addressed in these bills.

- HB7396 AI Development and Regulation Act. Promotes the safe development of AI and alignment with ethical principles, human rights, and public interest.[112]

- HB7913 Artificial Intelligence Regulation Act. Lays out guiding principles for AI development and use and a bill of rights for Filipinos in the digital age.[113]

- HB10944 Artificial Intelligence Act. Calls for the creation of a Philippine Artificial Intelligence Board tasked with developing and researching AI systems.[114]

- HB10385 AI Regulation Act. Recommends the establishment of an AI Bureau to ensure technical development, protect workers' rights, and prevent job displacement.[115]

- Patchwork of soft-law regulations. The current AI safety governance architecture in the Philippines is largely fragmented and sectoral. While there is no single comprehensive framework, sector-specific initiatives do exist. The Department of Information and Communications Technology (DICT) published a Joint Memorandum Circular in 2024 outlining guidelines for ethical and trustworthy AI use in government.[116] The Commission on Elections has also introduced rules governing the use of AI in campaigns in 2024.[117]

Key institutions and research hubs

- Department of Trade and Industry (DTI). The lead agency for the national AI strategy.

- Department of Information and Communications Technology (DICT). The primary policy-making, implementing, and administrative body of the Philippine government for the Information and Communications Technology (ICT) agenda, also heavily involved in AI policy.

---

112 "AI Development and Regulation Act," Pub. L. No. House Bill No. 7396, (2023), https://ldr.senate.gov.ph/bills/house-bill-no-7396-19th-congress.

113 Keith Micah "Att.Mike" D.L. Tan, "Artificial Intelligence (AI) Regulation Act," Pub. L. No. House Bill No. 7913 (2023), https://docs.congress.hrep.online/legisdocs/basic_19/HB07913.pdf.

114 Mary Mitzi Cajayon-Uy, "Artificial Intelligence Act," Pub. L. No. House Bill No. 10944 (2024), https://docs.congress.hrep.online/legisdocs/basic_19/HB10944.pdf.

115 Bryan B. Revilla, Lani Mercado-Revilla, and Ramon Jolo Revilla III, "AI Regulation Act," Pub. L. No. House Bill No. 10385 (2024), https://docs.congress.hrep.online/legisdocs/basic_19/HB10385.pdf.

116 "Philippines: Artificial Intelligence Readiness Assessment Report," UNESCO.

117 "Philippines: AI and social media guidelines for the 2025 elections issued by the COMELEC," Baker McKenzie, October 4, 2024, https://insightplus.bakermckenzie.com/bm/data-technology/philippines-ai-and-social-media-guidelines-for-the-2025-elections-issued-by-the-comelec.

- Department of Science and Technology (DOST). Focuses on AI policy from a R&D perspective.

- Center for AI Research (CAIR). Established in July 2024 by the DTI, it aims to conduct research on AI for economic growth and plans to partner with AI Singapore to explore regional solutions. In 2024, DTI said it expected CAIR to pursue income-generating projects and become self-sustaining in three years.[118] However, no concrete timeline has been laid out for CAIR or its projects to receive approval from the National Economic and Development Authority.

- Private/academic sector. Academic institutions like the University of the Philippines have produced foundational research, while nonprofits and private corporations actively advocate for policy solutions and support risk-mitigation projects.

3. International engagement

- Sporadic participation. The Philippines has participated in the 2023 U.K. AI Safety Summit and the 2024 Seoul AI Summit. While it did not attend the 2025 Paris AI Action Summit, it participated in ministerial-level UNESCO discussions on the sidelines of the summit,[119] and its Consulate General in San Francisco attended an AI Action Summit debriefing in the U.S.[120]

- ISO/IEC JTC 1 on AI. The country is a member of the standard setting committee.

- Regional ambitions. Prominent politicians, such as Speaker of the House of Representatives Martin Romualdez, have announced their intention to propose a regional AI regulatory framework based on its domestic draft legislation during its ASEAN chairmanship in 2026.[121] Such a regulatory framework might strike a different tone from ASEAN's current soft-law and business-friendly approach toward governing AI, exemplified by the organization's Guide on AI Governance and Ethics (see the ASEAN initiatives section of this report).

---

118 Lade Jean Kabagani, PH Seeks ASEAN Support for Proposed AI Regulatory Framework, Philippine News Agency, January 17, 2024, https://www.pna.gov.ph/articles/1228244.

119 "Philippines participates in UNESCO discussions on AI readiness," Department of Foreign Affairs, Republic of the Philippines, February 25, 2025, https://dfa.gov.ph/dfa-news/news-from-our-foreign-service-postsupdate/36166-philippines-participates-in-unesco-discussions-on-ai-readiness.

120 Philippine Consulate General in San Francisco, "The event convened leading technology executives, policy experts, and the consular corps in the San Francisco Bay Area to discuss…" Facebook, March 31, 2025, https://www.facebook.com/PHinSF/photos/the-event-convened-leading-technology-executives-policy-experts-and-the-consular/1077026827784418/.

121 Reuters, Philippines to Propose ASEAN AI Regulatory Framework, House Speaker Says, January 17, 2024, https://www.reuters.com/technology/philippines-propose-asean-ai-regulatory-framework-house-speaker-says-2024-01-17/.

4. Case study

- DOST ASTI ALaM. The Department of Science and Technology's (DOST) Automated Labelling Machine (ALaM) Project is a key initiative under the Philippine Sky Artificial Intelligence Program. It aims to create an optimized workflow for developing machine learning models for applications such as analyzing satellite images to classify land use and assessing the structural strength of buildings. The project seeks to spread the benefits of AI technology beyond urban centers and across the rural-urban divide, serving as a model for inclusive AI deployment.

- The Ambit. The Ambit is a local grassroots initiative started in the Philippines that shows how civic networks can meaningfully shape international AI governance. In just 15 months, and with no formal funding, The Ambit has mobilized domain key stakeholders, fostered policy dialogue, and built regional leadership in responsible AI. This impact is reflected in the Philippines' rank of 31st out of 183 countries in the 2024 Global Index on Responsible AI by the Global Center on AI Governance. The report highlights the country's strong performance in the Non-State Actors pillar and its recognition as the second-brightest spot globally for protecting human rights in AI.

- Whitebox Research. This is a nonprofit based in Manila working toward developing more AI interpretability and safety researchers, particularly in or near Southeast Asia, funded by the Long-Term Future Fund and Manifund.

5. Roundtable top insights

- Labor market dynamics. The BPO sector is highly susceptible to automation by AI, creating a significant risk of job displacement. Protections must be extended beyond the formal economy to the vast informal and gig economy, where workers are vulnerable to exploitation via AI-driven systems.

- Talent development and retention. A key challenge is the lack or underutilization of policies and incentives to train and retain AI talent. Low wages relative to other economies stunt the growth of a robust local AI governance and research ecosystem.[122]

- Need for coordinated and streamlined governance mechanisms. Panelists highlighted the need to streamline multi-stakeholder deliberative processes on AI safety governance. They cited examples of processes that were either too rushed for meaningful input (a 10-day open comment period) or too inefficient and unstructured to produce synthesized outcomes (an overly long congressional hearing).

122 Reflecting on the pay differential during our roundtable session, one AI safety sector worker shared that while an internship as a job candidate at a U.S. or U.K.-based NGO might pay USD $26 per hour, a job at a big consulting firm in Manila would only pay USD$650 per month, breaking down to around USD$4 per hour (assuming a 40 hour work week).

- Risk of policy duplication and regulatory fragmentation. The four draft bills under call for the creation of multiple agencies poised to take charge of AI regulation. Other agencies and ministries, such as the DOST, DTI, and the DICT, already have some mandate in drafting AI policy in the Philippines. The existing sector-specific patchwork of AI safety soft-law regulations also provides uneven coverage across AI use cases.

"By building a **strong pipeline of talent**, we can equip the workforce to handle the complexities of modern AI systems that cut across several industries."

Elmer C. Peramo, Senior Science Research Specialist, Department of Science and Technology, Philippines, Manila Roundtable

2.9. SINGAPORE

2.9. SINGAPORE

See roundtable session here (December 2024)

Resource-rich and forward-looking, Singapore positions itself as a global and regional hub for trusted AI. Its domestic playbook is characterized by a balance between enabling innovation and managing risks, preferring agile, tech-agnostic guidelines over rigid laws. This is exemplified by the 2024 update to its Model AI Governance Framework to cover generative AI.[123] A university-based AI Safety Institute anchors its research ambitions. Internationally, Singapore is an active participant in shaping global norms, chairing the ASEAN working group on AI governance and contributing open-source evaluation tools to the global community.

1. Timeline highlights

- **2017:** AI Singapore (AISG) established to bring together research institutions and companies.[124]
- **2019:** National AI Strategy (NAIS) launched.[125]
- **2020:** Model Governance Framework (MGF) for traditional AI, released in 2019, updated in 2024.[126]
- **2021:** Ministry of Health releases AI in Health care Guidelines (AIHGle).[127]
- **2022:** Digital Trust Centre (DTC) established at Nanyang Technological University (NTU).[128]

123 "Model AI Governance Framework for Generative AI: Fostering a Trusted Ecosystem" (Infocomm Media Development Authority, May 30, 2024), https://aiverifyfoundation.sg/wp-content/uploads/2024/05/Model-AI-Governance-Framework-for-Generative-AI-May-2024-1-1.pdf.
124 "AI Singapore," May 3, 2022, https://aisingapore.org/.
125 "National Artificial Intelligence Strategy: Advancing Our Smart Nation Journey" (Smart Nation Singapore, 2019), https://file.go.gov.sg/nais2019.pdf.
126 "Model Artificial Intelligence Governance Framework: Second Edition" (Infocomm Media Development Authority, January 2020), https://www.pdpc.gov.sg/-/media/files/pdpc/pdf-files/resource-for-organisation/ai/sgmodelaigovframework2.pdf.
127 "Artificial Intelligence in Healthcare Guidelines" (Ministry of Health Singapore, October 2021), https://isomer-user-content.by.gov.sg/3/9c0db09d-104c-48af-87c9-17e01695c67c/1-0-artificial-in-healthcare-guidelines-(aihgle)_publishedoct21.pdf.
128 "Digital Trust Centre (DTC) About Us," Nanyang Technological University Singapore, accessed July 2, 2025, https://www.ntu.edu.sg/dtc/about-us.

- **2023:** NAIS updated to version 2.0 to address generative AI and establish a "Trusted Environment."[129]

- **2024:**

  - Model AI Governance Framework for Generative AI (MGF-Gen AI) released.[130]

  - Cybersecurity Agency (CSA) releases Guidelines on Securing AI Systems.[131]

  - Elections (Integrity of Online Advertising) (Amendment) Bill passed, effectively banning deepfakes in election content.[132]

  - DTC at NTU is officially designated as Singapore's AI Safety Institute (AISI).[133]

  - Singapore participates in the San Francisco launch of the International Network of AISIs.

- **2025:**

  - Release of Joint Testing Report in collaboration with Japan under AISI framework to evaluate guardrails on non-English LLMs.[134]

  - Publication of Singapore AI Safety Red Teaming Challenge Evaluation Report 2025 to understand LLM guardrail performance in multilingual contexts.[135]

---

129 "NAIS 2.0 Singapore National AI Strategy: AI for the Public Good for Singapore and the World," 2023, https://file.go.gov.sg/nais2023.pdf.

130 "Model AI Governance Framework 2024 - Press Release," Infocomm Media Development Authority (blog), January 16, 2024, https://www.imda.gov.sg/resources/press-releases-factsheets-and-speeches/press-releases/2024/public-consult-model-ai-governance-framework-genai.

131 "Guidelines on Securing AI Systems" (Cyber Security Agency of Singapore, October 2024), https://isomer-user-content.by.gov.sg/36/e05d8194-91c4-4314-87d4-0c0e013598fc/Guidelines%20on%20Securing%20AI%20Systems.pdf.

132 "Elections (Integrity of Online Advertising) (Amendment) Act 2024," Pub. L. No. Bill No. 29/2024 (2024), https://sso.agc.gov.sg/Bills-Supp/29-2024/Published/20240909?DocDate=20240909.

133 "Digital Trust Centre (DTC) About Us."

134 "SG Announces New AI Safety Initiatives at the Global AI Action Summit in France," Infocomm Media Development Authority (blog), February 11, 2025, https://www.imda.gov.sg/resources/press-releases-factsheets-and-speeches/press-releases/2025/singapore-ai-safety-initiatives-global-ai-summit-france.

135 "Singapore AI Safety Red Teaming Challenge Evaluation Report" (Infocomm Media Development Authority, February 2025), https://www.imda.gov.sg/-/media/imda/files/about/

2. National governance landscape

Strategic policies

- National AI Strategy (NAIS 2.0, 2023). Updated from the 2019 version, it identifies building a "Trusted Environment" wherein AI is developed and deployed in a "safe, trustworthy, and responsible manner."

Regulatory and legal instruments

- Model AI Governance Framework for Generative AI (MGF-Gen AI, 2024). A set of voluntary, pro-innovation guidelines outlining nine dimensions for trusted AI, including incident reporting, data management, and testing. The MGF-Gen AI was finalized with feedback from experts and has been praised by industry leaders.[136]

- Personal Data Protection Act (PDPA, 2012). The core data protection law governing data collection and use, supported by advisory guidelines on using personal data in AI systems. Advisory guidelines on personal data use in AI systems were also released in 2013 by the commission established to administer and enforce the PDPA.[137]

- Elections (Integrity of Online Advertising) (Amendment) Act (2024). A rare instance of hard-law specific to an AI application, it prohibits digitally manipulated content in online election advertising.[138]

- Sectoral/industry-specific guidelines. Various government agencies have issued optional but recommended guidelines on specific AI use cases. For instance, the Ministry of Health released the AI in Health Care Guidelines (2021)[139] while the Monetary Authority of Singapore released Principles to Promote Fairness, Ethics, Accountability, and Transparency in the Use of AI and Data Analytics (2018).[140]

Key institutions and research hubs

- National AI Group. This group is housed under the Ministry of Digital Development and

---

emerging-tech-and-research/artificial-intelligence/singapore-ai-safety-red-teaming-challenge-evaluation-report.pdf.

136 "Model AI Governance Framework for Generative AI."

137 "Personal Data Protection Act 2012," accessed July 2, 2025, https://sso.agc.gov.sg/Act/PDPA2012.

138 Elections (Integrity of Online Advertising) (Amendment) Act 2024.

139 "AIHGIe."

140 "Principles to Promote Fairness, Ethics, Accountability and Transparency (FEAT) in the Use of Artificial Intelligence and Data Analytics in Singapore's Financial Sector" (Monetary Authority of Singapore, 2018), https://www.mas.gov.sg/~/media/MAS/News%20and%20Publications/Monographs%20and%20Information%20Papers/FEAT%20Principles%20Final.pdf.

Information (MDDI) that specifically develops and coordinates AI policies under the Smart Nation whole-of-government national initiative.

- Digital Trust Centre (DTC). Housed within the Nanyang Technological University, the DTC was designated as Singapore's AI Safety Institute in 2024 and spearheads R&D in trust technologies and AI safety, distinguishing itself from government-housed AISIs in other countries.[141]

- AI Verify Foundation. A subsidiary of the Infocomm Media Development Authority (IMDA), it develops open-source governance tools, including Project Moonshot, an evaluation toolkit for LLMs (see case studies).

- AI Singapore (AISG). A national platform that drives AI innovation and talent, best known for developing the SEA-LION family of open-source LLMs for Southeast Asian languages (see case studies).

3. International engagement

- Shaping global norms. Singapore is an active participant in the AI Summit series and a founding member of the International Network of AISIs, collaborating with the U.S., U.K., and EU on safety testing. To date, it is the sole Southeast Asian AISI represented in the network.

- Regional leadership. It chairs the ASEAN Working Group on AI Governance (WG-AI) and has been a key proponent of regional policies like the ASEAN Guide on AIGE and its Expanded Guide on Gen AI.[142]

- Small states diplomacy. It co-developed the AI Playbook for Small States with Rwanda through the Digital Forum of Small States (FOSS) to shape discourse relevant to their unique context.

4. Case studies

- Project SEA-LION (Southeast Asian Languages in One Network). An open-source family of LLMs anchored by AI Singapore, specifically trained on regional languages (e.g., Thai, Vietnamese, Bahasa Indonesia), SEA-LION aims to improve utility and reduce bias for Southeast Asian contexts, serving as a prime example of regional cooperation and contextualized AI.

---

141 The US' AISI is housed within the National Institute of Standards and Technology (NIST), a government agency (source). Similarly, the U.K.'s AISI is a directorate of the U.K. Department for Science, Innovation, and Technology (source).

142 "ASEAN Guide on AI Governance and Ethics"; "Expanded ASEAN Guide on AI Governance and Ethics-Generative AI."

- Project Moonshot. An open-source LLM evaluation toolkit designed to integrate and simplify benchmarking, red-teaming, and testing. It provides a structured way for developers and system owners to manage safety risks, promoting the goal of trustworthy and safe AI.

5. Roundtable top insights

- Data sovereignty vs. data hub ambitions. Given its robust data laws and well-annotated datasets, Singapore aims to become a "trusted global data hub." Yet, this is challenged by regional data localization policies. Its small population generates systems data, making cross-border data flows essential for training robust, multicultural AI models. Panelists stressed that while Singapore's data protection laws are robust, a possible next frontier could be innovating in data enablement. This includes creating concepts like "data unions," where individuals can choose to pool and monetize insights from their data, and leveraging high-quality, well-annotated data to become a thought leader in trusted data sharing.

- Generative AI-enabled scams. As a major financial hub, Singapore is "uniquely vulnerable" to AI-enabled scams, a risk amplified by its open economy and rapid push toward a cashless society and digital banking. Reports from 2024 indicate a significant year-on-year increase in deep fake attacks, posing a direct threat to its digital economy.[143]

- Position as a regional R&D Hub. Singapore is poised to be the regional R&D hub for AI safety. The key opportunity is to distinguish itself by contextualizing global research on issues like value alignment for the specific needs of Southeast Asian societies and other small states, as demonstrated by initiatives like SEA-LION and the Digital FOSS AI Playbook for Small States. For instance, in February 2025, Singapore and Japan collaborated under the International Network of AISI framework to evaluate guardrails on non-English LLMs.[144] Singapore also published the AI Safety Red Teaming Challenge Evaluation Report 2025 to understand LLM guardrail performance in multilingual contexts.[145]

"Being a small state is **not necessarily a disadvantage**, especially if you are **forward-looking**, doing **cutting-edge** research and engaged with all other players, locally and globally."

Prof. Mohan Kankanhalli, Director, NUS AI Institute

143 Fabian Koh, "Singapore Registers Asia-Pacific's Biggest Spike in Identity Fraud, Driven by Deepfake Surge - CNA," November 22, 2024, https://www.channelnewsasia.com/singapore/identity-fraud-deepfakes-scams-ai-4761836.

144 "SG Announces New AI Safety Initiatives at the Global AI Action Summit in France."

145 "Singapore AI Safety Red Teaming Challenge Evaluation Report."

2.10. THAILAND

2.10. THAILAND[146]

See roundtable session here (January 2025)

As one of the more active players in the region, Thailand's domestic regulatory landscape has been dynamic, with multiple agencies introducing guidelines and draft legislation over the past five years. Its approach to governance is marked by a mix of soft-law principles and proposed hard-law instruments, notably a draft Royal Decree inspired by the EU's risk-based model. The country is also proactive in promoting AI R&D and is positioning itself to become a niche leader in medical AI. However, key challenges remain, including regulatory fragmentation among government agencies and concerns about policy consistency across changing political administrations. While historically slow in international engagement, Thailand is showing signs of ramping up its activity on the global stage.

1. Timeline highlights

- **2019:** The VISTEC-depa Thailand AI Research Institute (now VISTAI) is founded.[147]
- **2021:** Cabinet approves the Digital Thailand AI Ethics Guidelines.[148]
- **2022:**
    - In July, the National AI Strategy and Action Plan (2022–2027) was approved by the cabinet.[149]
    - In August, the Generative AI Governance Guideline for Organizations was published by the Electronic Transactions Development Agency (ETDA) and Office of the National Digital Economy and Society Commission (ONDE) through the AI Governance Centre.[150]

146 In addition to the authors of this paper, Nitya Kuthiala also contributed to this section.

147 About VISAI," VISAI.ai, accessed July 2, 2025, https://visai.ai/about.

148 Stanati Netipatalachoochote and Ludovic Pailler, "Developing Artificial Intelligence Legislation in Thailand: Lessons from the European Union," Journal of Human Rights, Culture and Legal System 5, no. 1 (March 17, 2025): 1–32, https://doi.org/10.53955/jhcls.v5i1.424.

149 "The Cabinet Approved the (Draft) Thailand National AI Strategy and Action Plan (2022–2027)," NECTEC : National Electronics and Computer Technology Center (blog), July 30, 2022, https://www.nectec.or.th/en/about/news/cabinet-national-ai-strategy.html.

150 "Thailand: Issued DE and ETDA Generative AI Governance Guideline for Organisations Including Guidelines for the Ethical Application of Generative AI for Organisations," Digital Policy Alert, October 30, 2024, https://digitalpolicyalert.org.

- In October, the Draft Royal Decree on Business Operations that Use AI was introduced.[151]

- In November, the AI Governance Clinic (AIGC) was established.[152]

- **2025:**

  - Participates in the Paris AI Action Summit.

  - Hosts the third UNESCO Global Forum on the Ethics of AI.[153]

2. Governance landscape

Strategic policies

- National AI Strategy and Action Plan (2022–2027). This foundational document released in 2022 lays out the general thrust of Thai AI governance, aiming to build AI infrastructure and support economic growth while explicitly recognizing the importance of managing ethical, legal, and social implications as a key strategy.

Regulatory and legal instruments

- Draft Royal Decree on Business Operations that Use Artificial Intelligence Systems (2022). Inspired by the EU AI Act's risk-based approach, it proposes three risk categories ("Unacceptable," "High," and "Limited") with increasingly strict requirements and includes punitive measures like criminal liability.[154]

- Draft Act on the Promotion and Support for National AI Innovations (2023). This draft law aims to promote responsible and innovative AI development by establishing mechanisms like a testing sandbox and an AI compensation fund for third-party damages.[155]

---

151 "Thailand: Published Draft Royal Decree on Business Operations That Use Artificial Intelligence Systems," Digital Policy Alert, October 18, 2022, https://digitalpolicyalert.org.

152 "ETDA to Launch Thailand's First AI Governance Clinic (AIGC)," NECTEC : National Electronics and Computer Technology Center (blog), November 7, 2022, https://www.nectec.or.th/en/about/news/etda-aigc-promote.html.

153 "Global Forum on the Ethics of AI," UNESCO, accessed July 2, 2025, https://www.unesco.org/en/forum-ethics-ai.

154 For instance, AI-enabled social scoring is categorized as an "Unacceptable Risk," subject to a blanket prohibition on its use. Chatbots are a "Limited Risk," requiring only a transparency obligation.

155 "AI Regulation in Thailand: Current State and Future Directions" (Baker McKenzie, November 2023), https://insightplus.bakermckenzie.com/bm/attachment_dw.action?attkey=-FRbANEucS95NMLRN47z%2BeeOgEFCt8EGQJsWJiCH2WAWuU9AaVDeFguGeARDEncDx-

- Generative AI Governance Guideline for Organizations (2024). This set of voluntary guidelines aims to encourage executives to establish governance frameworks for generative AI at the enterprise level. It also outlines benefits, limitations, and risks of generative AI.[156]

- Soft-law guidelines. Various nonbinding guidelines have been released, including the Digital Thailand AI Ethics Guidelines (2021) and the AI Governance Guidelines for Executives (2023) which is currently being trialed in the medical industry in partnership with Mahidol University.

Key institutions and research hubs

- No centralized agency. While there is no single coordinating body for AI policy, key responsibilities are distributed among:

- Electronic Transactions Development Agency (ETDA). Involved in drafting laws and houses the AI Governance Clinic.

- Office of the National Digital Economy and Society Commission (ONDE). Focuses on the digital economy and introduced the Draft Royal Decree.

- National Science and Technology Development Agency (NSTDA). Focuses on AI R&D and infrastructure projects.

- VISTEC-depa Thailand AI Research Institute (VISTAI). A key R&D institute that develops Thai-based LLMs, English-Thai translation systems, and partners with regional players like AI Singapore in projects like Project SEALD (Southeast Asian Languages in One Network Data), a multilingual LLM data collection endeavor.

3. International engagement

- Ramping up activity. Although Thailand did not participate in the 2023 U.K. or 2024 Seoul AI summits, it took part in the 2025 Paris AI Action Summit.

- Hosting UNESCO Global Forum on the Ethics of AI. Thailand hosted the third iteration of the forum in June 2025, the first time the forum will be held in the Asia-Pacific. This coincides with the planned release of its UNESCO AI Readiness Assessment Methodology report and signals a growing leadership role.

&nav=FRbANEucS95NMLRN47z%2BeeOgEFCt8EGQbuwypnpZjc4%3D&attdocparam=pB7HEsg%2FZ312Bk8OIuOIH1c%2BY4beLEAe9Q37ImwtvME%3D&fromContentView=1.

156 Jeremy Werner, "Thailand Unveils New Guidelines for Responsible Use of Generative AI in Organizations," BABL AI (blog), November 12, 2024, https://babl.ai/thailand-unveils-new-guidelines-for-responsible-use-of-generative-ai-in-organizations/.

"In Thailand, there's a lot of interest in medical AI, which is a **rich practical area for alignment research**. There's are a lot of complicated ethical questions around consent that could be explored."

Jam Kraprayoon, IAPS, Bangkok Roundtable

4. Case study

- LANTA HPC Cluster. In 2022, the NSTDA partnered with NVIDIA and Hewlett Packard, a multinational IT company, to build LANTA, a powerful and energy-efficient supercomputer.[157] LANTA plays a crucial role in risk mitigation, particularly for climate resilience. By analyzing vast amounts of climate data, it allows researchers to forecast weather with greater accuracy, enabling farmers to plan for drought-resistant crops. This serves as a real-life example of using AI to spread benefits equitably between urban and rural contexts.

5. Roundtable top insights

- Regulatory fragmentation. The proactivity of different agencies is a positive sign, but the lack of a comprehensive, macro-level coordinating body creates a significant risk of fragmentation. Panelists noted that ONDE and ETDA appear to be formulating policy in silos, raising questions about the interoperability of their respective draft laws.

- Policy consistency. Frequent changes in political administration have raised concerns about the consistency and continuity of Thai AI policy. Panelists stressed the need to depoliticize AI policy to ensure it can endure political turbulence.

- Capitalizing on a niche in medical AI. Thailand has cultivated a global reputation as a hub for medical tourism, with revenue from the industry being projected to reach $24.4 billion by 2027.[158] Panelists identified a clear opportunity for the country to capitalize on this niche by becoming a testing ground for both medical AI applications (like proactive health care) and the governance mechanisms needed to regulate them.

157 "NSTDA and HPE Form a Partnership to Build a New Supercomputer - NSTDA Eng," NSTDA (blog), December 14, 2021, https://www.nstda.or.th/en/news/news-years-2021/nstda-and-hpe-form-a-partnership-to-build-a-new-supercomputer.html.

158 Gumporn Supasettaysa, "Public-Private Partnerships in the Thailand Medical Tourism Industry" (Harvard Kennedy School Rajawali Foundation Institute for Asia, September 2023), https://rajawali.hks.harvard.edu/wp-content/uploads/sites/2/2023/10/331231_hvd_ash_thailand_medical_v2.pdf.

2.11. TIMOR-LESTE

2.11. TIMOR-LESTE[159]

See roundtable session here (March 2025)

As the youngest state in Southeast Asia, Timor-Leste's engagement with AI safety and security policy is at a nascent stage. Efforts are embedded within a broader national strategy, Timor DIGITAL 2032, which prioritizes foundational digital transformation, such as modernizing public services and expanding national connectivity. AI is acknowledged as a future enabler, but specific safety policies and ethical governance frameworks are absent. Facing systemic challenges like limited human capital and competing developmental priorities,[160] the country relies heavily on international partnerships to build institutional capacity.

1. Timeline highlights

- **2017:** The Information and Communication Technology Agency (TIC TIMOR) is established.[161]
- **2022:** Granted ASEAN Observer Status.[162]
- **2023:** The Timor DIGITAL 2032 strategy is launched.[163]
- **2024:** A draft law on cybercrime is released for review.[164]

159 In addition to the authors of this paper, Quentin Flament also contributed to this section.

160 A UNDP 2023 report estimated that 48.3 percent of the population was multidimensionally poor (taking health, education, and standard of living deprivations into account). For comparison, Indonesia and the Philippines had 3.9 percent and 4.3 percent of their populations classified as multidimensionally poor. "Multidimensional Poverty Index 2023: Timor-Leste" (UNDP, 2023), https://hdr.undp.org/sites/default/files/Country-Profiles/MPI/TLS.pdf.

161 "TIC TIMOR | Agência da Tecnologia Informação e Comunicação," accessed July 2, 2025, http://www.tic.gov.tl/en/.

162 Bangkok Post Public Company Limited, "Asean Leaders Agree to Admit Timor-Leste," Bangkok Post, May 29, 2025, https://www.bangkokpost.com/thailand/general/3036805/asean-leaders-agree-to-admit-timor-leste.

163 "Timor Digital 2032: Timor-Leste's Ten-Year Digital and ICT Development Strategy and Policy" (TIC TIMOR, 2023), https://anyflip.com/gvocj/zerd.

164 "Proposal for Law No..../2024 Cybercrime" (IX Constitutional Government Ministry of Justice, June 6, 2024), https://www.laohamutuk.org/Justice/defamation/cybercrime/240606PPL%20CibercrimeEn.pdf.

- **2025 (planned):** Admission as a full member of ASEAN in October.[165]

2. Governance landscape

Strategic policies

- Timor DIGITAL 2032 (2023). The primary national strategy guiding digital transformation. It acknowledges AI as a future enabler but underscores a critical gap by not yet including specific AI safety policies or ethical governance frameworks.[166]

- Regulatory and legal instruments

- No AI-specific regulation. Timor-Leste currently has no hard or soft laws specifically for AI.

- Foundational laws lacking or in draft. Fundamental legal instruments for data protection, cybersecurity, and ICT regulation are either in draft stages or entirely lacking.

Key institutions and research hubs

- Information and Communication Technology Agency (TIC TIMOR). The lead government agency for digital transformation and technology initiatives.

- Private/academic sector. The AI ecosystem is limited. Actors like Telkomcel and the Universidade Nacional Timor Lorosa'e (UNTL) have begun to explore AI through small-scale initiatives like startup competitions and workshops.

- Civil society. Organizations like NGO Forum Timor-Leste and La'o Hamutuk have emerged as important advocates for digital rights. They scrutinize draft legislation and are laying the essential groundwork for a future rights-based governance model.

3. International engagement

- Regional focus. Timor-Leste's primary focus is on integration into ASEAN. It has participated as an observer in the ASEAN Digital Ministers' Meeting (ADGMIN).

- Limited global presence. It did not participate in the 2023 U.K., 2024 Seoul, or 2025 Paris AI summits.

- Developmental partnerships. The country engages extensively with international partners like the UNDP, UNCTAD, UNESCO, and the World Bank to build up institutional capacity and basic infrastructure.

165 Limited, "Asean Leaders Agree to Admit Timor-Leste."
166 "Timor Digital 2032."

4. Roundtable top insights

- First things first. Before tackling complex AI governance, panelists said Timor-Leste must focus on foundational issues like building basic digital literacy, establishing robust data privacy and cybersecurity laws, and strengthening institutional trust. For instance, grassroots understanding of AI risks like misinformation and algorithmic bias is extremely low. This increases the vulnerability of both institutions and individuals to technological misuse and exploitation.

- Limited human capital. A major obstacle is the scarcity of specialized human capital in AI, cybersecurity, and data governance. This creates a risk of overreliance on foreign technologies and external consultants, which could undermine the development of context-specific and culturally appropriate AI applications.

- Inherently diverse cultural context. Panelists pointed out that Timor-Leste's linguistic and cultural diversity (e.g., Tétum and Portuguese-speaking communities) demands localized AI solutions. Generic models trained on Western datasets are unlikely to meet local needs, creating a risk of "digital marginalization" without targeted investment in local data collection and annotation.

"[Timor-Leste] has **limited digital infrastructure and AI expertise.** AI requires advanced computing infrastructure and secure systems. [Timor-Leste] has **limited access to such advanced infrastructure**, which will slow down the adoption of the technology, and make it hard to develop localised AI solutions."

"**Digital literacy is very low.** In remote or mountainous regions, a civil servant who knows how to use a computer might be hard to find."

Justo Fernandes, Director of Application and Information System Development, Department of Information Technology and Communication, Timor-Leste, Dili Roundtable

2.12. VIETNAM

2.12. VIETNAM

See roundtable session here (February 2025)

Vietnam's approach to AI safety governance has matured significantly since 2021, characterized by a pragmatic recontextualization of international frameworks for local needs. The government demonstrates a trajectory moving from soft, voluntary principles toward hard law, with a strong emphasis on data protection and cybersecurity. While domestic policy development is robust, Vietnam's engagement in international AI governance forums remains limited. Key challenges include a lack of integration with global systems, a disconnect between macro-level data protection policies and micro-level implementation, and practical barriers to AI adoption, such as a scarcity of high-quality Vietnamese-language datasets.

1. Timeline highlights

- **2021:** The National Strategy on Research, Development, and Application of AI to 2030 is published.[167]
- **2023:** The government issues Decree No. 13/2023/ND-CP, a comprehensive legal document on personal data protection.[168]
- **2024:**
    - In June, the Ministry of Science and Technology (MoST) released Decision No. 1290/QD-BKHCN, outlining voluntary principles for responsible AI.[169]
    - In July, the Draft Law on Digital Technology Industry was proposed.
    - In December, the private sector-led Ethical AI Committee was established by FPT and partners.[170]

---

167 "Decision on Issuing the National Strategy on Research, Development, and Application of Artificial Intelligence until the Year 2030" (Viet Government Portal, March 17, 2021), https://wp.oecd.ai/app/uploads/2021/12/Vietnam_National_Strategy_on_RD_and_Application_of_AI_2021–2030.pdf.

168 "On Personal Data Protection," Decree No. 13/2023/ND-CP § (2023), https://resources.finalsite.net/images/v1710749328/saigon/l5cqb4izgsc1uwugo98e/Decree13_2023_ND-CPPersonalDataProtectionENGLISH.pdf.

169 Nguyen Xuan Thuy and Ho My Ky Tan, "AI Regulations in Vietnam and Opportunities for Investors," LNT & Partners (blog), December 31, 2024, https://www.lntpartners.com/legal-briefing/ai-regulations-in-vietnam-and-opportunities-for-investors.

170 "FPT to Support Vietnam's Push for Ethical and Responsible AI Innovation," FPT Soft-

- 2025: Vietnam participates in the Paris AI Action Summit.

2. Governance landscape

Strategic policies

- National Strategy on Research, Development, and Application of AI to 2030 (2021). This foundational document outlines a blueprint for Vietnam to become an AI hub. While initially light on ethical governance, it recognized the need to balance data use with privacy protection. Yet, the strategy lacks reference to the need for an AI safety governance framework, which analysts have pointed out as indicative that "the topic of ethics in relation to AI has barely begun to percolate in the country".[171]

Regulatory and legal instruments

- Decision No. 1290/QD-BKHCN (2024). A set of nine voluntary principles for the responsible development of AI systems, creating a human-centric approach that balances economic benefits with ethical considerations.

- Draft Law on Digital Technology Industry (2024). This draft law is slated to enter into force in May 2025.[172] A significant move toward hard regulation, it defines key terms like AI, introduces a risk-based governance approach inspired by the EU AI Act, and assigns distinct responsibilities to AI developers, suppliers, and users.[173]

- Decree No. 13/2023/ND-CP. One of the first comprehensive legal documents on data protection in the country.[174] While not specifically targeting AI systems, it signified a first step toward alignment with international data protection standards like the EU's General Data Protection Regulation (GDPR). It emphasizes the importance of obtaining consent for data processing and sets rules for the transfer of Vietnamese personal data abroad. Panelists

---

ware (blog), December 5, 2024, https://fptsoftware.com/newsroom/news-and-press-releases/press-release/fpt-to-support-vietnamm-push-for-ethical-and-responsible-ai-innovation.

171 "Viet Nam Artificial Intelligence," Asia Society Policy Institute, July 11, 2022, https://asiasociety.org/policy-institute/raising-standards-data-ai-southeast-asia/ai/viet-nam.

172 "AI Regulations Come into Focus in Vietnam's Draft Law on Digital Technology Industry," Tilleke & Gibbins (blog), January 21, 2025, https://www.tilleke.com/insights/ai-regulations-come-into-focus-in-vietnams-draft-law-on-digital-technology-industry/.

173 On the risk categorization of AI systems, the Draft Law on Digital Technology Industry identifies three tiers—high risk (systems that pose a risk to health, safety, rights, and interests), high impact (systems with a broad scope, large user base, and using high computational resources), and standard systems. It also outlines prohibited AI use cases, like crime prediction and the untargeted collection of CCTV images. This framework of tiered risk system bears striking resemblance to the EU AI Act.

174 "Decree 13/2023/ND-CP on Personal Data Protection" (PwC, April 25, 2023), https://www.pwc.com/vn/en/publications/2023/newsbrief-decree-13-personal-data.pdf.

noted the regulatory regime still lacked policy tools for more frontier data management ideas like data compensation and transparency across the whole lifecycle of AI development.

- Draft Personal Data Protection Law (2024). Expected to enter into force in January 2026, it includes elements from Decree No. 13, but it remains unclear if this law stands to replace or exist alongside it.[175] Specific to AI, it permits individuals and organizations to research and develop systems using personal data provided there is informed consent and the ability to withdraw. It also crystallizes some voluntary guidelines on AI research and development surfaced in Decision No. 1290, compiling aspects of the guide, like data protection and security, into hard law.

Key institutions and research hubs

- Ministry of Science and Technology (MoST). It plays a crucial role in AI safety governance policy by introducing guidelines and promoting responsible AI practices. Specifically, it issued Decision 1290.

- Ministry of Information and Communications (MIC). This ministry also plays a key role in AI safety development. It is actively involved in drafting and implementing the Draft Law on Digital Technology Industry.

- Ministry of Public Security (MPS). In charge of the development of the Personal Data Protection Law.

- Unclear status of national AI innovation centers. While the National Strategy laid out plans to set up two national innovation centers for AI by 2025, it is unclear if the country is on track to achieve this goal.

"The priority of government seems to be governing data consent. Vietnam doesn't yet have advanced mechanisms to handle things like data transparency and accountability."

Danh Nguyen, Global Head of Legal at GreenFuture.tech, Legal-Tech Expert at LegalTech.org.

175 "Vietnam's Draft Law on Personal Data Protection: Key Highlights and Implications," ASL LAW (blog), October 21, 2024, https://aslgate.com/vietnams-draft-law-on-personal-data-protection-key-highlights-and-implications/.

3. International engagement

- Limited but increasing participation. Vietnam's engagement in international AI safety governance is limited. It did not participate in the 2023 or 2024 AI summits. It attended the 2025 Paris AI Action Summit but did not sign the outcome statement.

- Regional engagement. Vietnam is a proactive participant in the ASEAN Digital Ministers' Meeting (ADGMIN) but has not made specific proposals on regional AI regulation.

4. Case study

- FPT's Ethical AI Committee in Vietnam. In December 2024, global technology company FPT, in collaboration with the Vietnam Software & IT Services Association (VINASA) and AI experts like Professor Yoshua Bengio, announced the establishment of the Ethical AI Committee. The committee is tasked with ensuring the ethical development, deployment, and application of AI by setting safety standards, assessing risks, and promoting responsible AI practices. The committee is an example of private sector-led AI safety governance initiatives.

5. Roundtable top insights

- Data protection as a national security priority. Panelists noted that, given Vietnam's unique geopolitical context, the government views data protection through a national security lens, prioritizing data sovereignty and protection from external misuse.

- Gaps in practical implementation. A key roundtable theme was the disconnect between macro-level policy (like data protection laws) and the micro-level capacity of citizens and businesses to effectively use them, suggesting a need for more support and resources. Further, while Vietnam's regulatory policies show signs of learning from overseas developments (for instance, the EU's GDPR), panelists mentioned that its efforts to integrate with international systems could be redoubled. For example, its data protection regime is not yet fully integrated with any cross-border data transfer mechanism, which could create a competitive disadvantage.

- Fundamental barriers to adoption: Panelists highlighted persistent practical challenges to AI adoption, namely the linguistic barrier between English-based AI and the Vietnamese public, and the operational bottleneck caused by a scarcity of high-quality local data.

2.13. ASEAN INITIATIVES[176]

On the whole, ASEAN is cognizant of the potential outsized economic benefit of AI to the region, which could result in a 10–18 percent increase in GDP, valued at approximately $1 trillion USD by 2030.[177] Yet, AI governance and ethics policy is a relatively recent addition to the suite of action areas that ASEAN member states focus on. At the first ASEAN Digital Ministers' Meeting in 2021, the ASEAN Digital Masterplan 2025 was adopted.[178] This plan explicitly included an enabling action to "adopt regional policy to deliver best practice guidance on AI governance and ethics, IoT spectrum, and technology",[179] marking one of the earliest instances of ASEAN identifying the policy area as a priority. Notably, the pace of policy formulation and adoption has picked up in recent years as global discourse on emerging technology becomes increasingly ubiquitous.

1. Timeline highlights

- **2021:** The inaugural ASEAN Digital Ministers' Meeting (ADGMIN) introduced the ASEAN Digital Masterplan 2025 (ADM2025) which included a pledge to deliver regional guidance on AI governance and ethics by 2025.[180]

- **2023:** Launch of negotiations for the Digital Economy Framework Agreement (DEFA), which aims to establish mechanisms for regulatory cooperation on emerging technologies, including AI.[181]

- **2024:**

176 In addition to the authors of this paper, Quentin Flament also contributed to this section.

177 "11th ASEAN Economic Community Dialogue Discusses Governance to Unlock AI Opportunity in ASEAN," Association of Southeast Asian Nations (blog), June 4, 2024, https://asean.org/11th-asean-economic-community-dialogue-discusses-governance-to-unlock-ai-opportunity-in-asean/.

178 "ASEAN Digital Masterplan 2025" (Association of Southeast Asian Nations, September 2021), https://asean.org/wp-content/uploads/2021/09/ASEAN-Digital-Masterplan-EDITED.pdf.

179 Ibid.

180 "ASEAN Digital Masterplan 2025."

181 "ASEAN Launches World's First Regionwide Digital Economy Framework Agreement," ASEAN Main Portal (blog), September 3, 2023, https://asean.org/asean-launches-worlds-first-regionwide-digital-economy-framework-agreement/.

- In January,[182] the Fourth ADGMIN in Singapore saw the release of the ASEAN Guide on AIGE and the establishment of the ASEAN Working Group on Artificial Intelligence (WG-AI).[183]

- In May, the ASEAN Committee on Science and Technology (COSTI) and U.S. Agency for International Development (USAID) released the Discussion Paper on the Responsible Development and Use of Generative AI in ASEAN, bringing generative AI into focus.[184]

- **2025:**

  - The Fifth ADGMIN in Thailand launched the Expanded ASEAN Guide AI Governance and Ethics—Generative AI specifically targeting generative AI.[185]

  - The ASEAN Responsible AI Roadmap was published, providing actionable regional and national policy recommendations, although it has not yet been implemented.[186]

  - (Ongoing) As ASEAN chairman, Malaysia is leading efforts to establish the ASEAN AI Safety Network (AI SAFE) and a Guide for Sustainable Data Centre Development.[187]

2. Governance landscape

Strategic policies and regulatory instrument

- ASEAN AI Governance and Ethics Guide (AIGE Guide, 2024). The AIGE guide promotes seven ethical AI principles based on UN and national frameworks and is oriented toward private sector practicality.[188] Specifically, it includes (i) a practical guide for organizations that wish to design, develop, and deploy traditional AI technologies, and (ii) a list of national and regional level initiatives governments can consider implementing to design, develop, and

182 "1st ASEAN Digital Ministers' Meeting approves Singapore led initiatives," MDDI, 22 January 2021, https://www.mddi.gov.sg/newsroom/1st-asean-digital-ministers-meeting-approves-singapore-led-initiatives
183 "ASEAN Guide on AI Governance and Ethics."
184 "Discussion Paper on the Responsible Development and Use of Generative AI in ASEAN" (ASEAN, USAID, U.S. Department of State, May 2024), https://aiasiapacific.org/wp-content/uploads/2024/11/Final-Discussion-Paper-To-be-published-5-nov-2024.pdf.
185 "Expanded ASEAN Guide on AI Governance and Ethics-Generative AI."
186 "ASEAN Responsible AI Roadmap (2025-2030)" (Association of Southeast Asian Nations, February 2025), https://asean.org/wp-content/uploads/2025/02/ASEAN-Responsible-AI-Roadmap-Final.docx.pdf.
187 "Accelerating Regional Economic Growth, Building Safer Digital Environment and Enhancing Regional Connectivity Key Takeaways for Malaysia at ADGMIN 2025."
188 "ASEAN Guide on AI Governance and Ethics."

deploy AI systems responsibly. The guide has been characterized as leaning closer to the U.S. National Institute of Standards and Technology (NIST) AI risk management approach rather than the strict framework of the EU AI Act.[189]

- Discussion Paper on the Responsible Development and Use of Generative AI in ASEAN (2024). Published in May 2024 by ASEAN COSTI and USAID as a complement to the AIGE Guide, the paper identifies policy, legal, and regulatory gaps in generative AI implementation, and provides recommendations on how to address them on the regional and national levels.[190]

- Expanded ASEAN Guide on AI Governance and Ethics—Generative AI (2025). Published in January 2025 and building upon the AIGE Guide and Discussion Paper in 2024, it focuses on six specific risks of generative AI including misinformation, privacy breaches, and deep fakes.[191]

- ASEAN Responsible AI Roadmap (2025–2030) (2025). Released in March 2025, the roadmap includes regional ("cross-cutting") and national ("targeted") policy recommendations for countries of different readiness tiers of Advanced, Promising, and Emerging. It also includes a self-assessment readiness scoring framework along 40 criteria.[192] While the roadmap has been released, it has not yet been endorsed and implemented by ASEAN.

- Digital Economy Framework Agreement. DEFA negotiations started in 2023 and are slated to end in 2025. They have included provisions for regulatory cooperation on AI, particularly focused on economic integration and infrastructure development.

- ASEAN AI Safety Network (AI SAFE) As ASEAN Chairman, Malaysia is shaping future directions for AI safety governance on the regional level. AI SAFE has been proposed as a platform to facilitate AI safety research, promote safe AI development and adoption, and

189 Both the U.S. NIST AI Risk Management Framework (AI RMF) and the EU AI Act aim to promote trustworthy and responsible AI, recognizing core principles like transparency, explainability, and fairness in AI systems. Substantial differences exist between the two. The U.S. AI RMF is high-level and broad, voluntary, excludes legal penalties, and nonbinding. It focuses on flexible recommendations targeting risk management processes. The EU AI Act is a binding piece of hard legislation, adopting a rules-based prescriptive approach that is enforceable through fines for noncompliance. It offers a tiered risk-based approach and imposes significant regulations on high-risk AI and AI in high-risk sectors like health care and finance and also mandates specific technical and process requirements. "Asean Eyes Hands-off AI Rules, Defying EU Ambitions," The Straits Times, October 11, 2023, https://www.straitstimes.com/business/asean-eyes-hands-off-ai-rules-defying-eu-ambitions.

190 "Discussion Paper on the Responsible Development and Use of Generative AI in ASEAN."

191 "Expanded ASEAN Guide on AI Governance and Ethics-Generative AI."

192 "ASEAN Responsible AI Roadmap (2025–2030)."

encourage regulatory harmonization.[193]

- Guide for Sustainable Data Centre Development (Ongoing). The guide for Sustainable Data Centre development is slated to act as a policy playbook for building a sustainable data center industry, addressing issues like resource demands of data centers.[194]

Key institutions and mechanisms

- ASEAN Digital Senior Officials' Meeting (ADGSOM). ADGSOM plays a coordinating role across digital areas including AI governance, cybersecurity, and privacy.

- ASEAN Working Group on Artificial Intelligence (WG-AI). Established in 2024 under ADGSOM and led by Singapore, the group is responsible for coordinating ASEAN's AI governance efforts and implementing policy recommendations.

- ASEAN Committee on Science and Technology (COSTI). Overseen by the ASEAN Ministerial Meeting on Science, Technology, and Innovation (AMMSTI), COSTI has taken the lead on initiatives like the generative AI paper and partnerships with Dialogue Partners. It generally leads initiatives related to science, innovation, and development.

3. International engagement with Dialogue Partners[195, 196]

- Australia. Generative AI (along with its adoption, and the development of infrastructure supporting it) and cybersecurity have been identified as priorities for ASEAN-Australia cooperation at recent meetings.

- China. ASEAN and China are developing a joint guide on cross-border data flows and an action plan focused on inclusive digital ecosystems for 2026–2030.[197] At the 2024 Chi-

193 "ASEAN Responsible AI Roadmap (2025–2030)."
194 "Accelerating Regional Economic Growth, Building Safer Digital Environment and Enhancing Regional Connectivity Key Takeaways for Malaysia at ADGMIN 2025."
195 ASEAN Dialogue Partners are states that formally cooperate with ASEAN and share its mission and vision. There are eleven Dialogue Partners: Australia, Canada, China, the EU, India, Japan, New Zealand, Russia, South Korea, the U.K., and the US. Australia, China, India, Japan, South Korea, and the U.S. have comprehensive strategic partnerships with ASEAN, with such partnerships entailing an elevated level of partnership above other Dialogue Partners, broad collaboration, and shared interests and goals. Only Dialogue Partners with comprehensive strategic partnerships are discussed in this section.
196 "External Relations," ASEAN Main Portal, accessed July 3, 2025, https://asean.org/our-communities/asean-political-security-community/outward-looking-community/external-relations/.
197 "Opening Remarks by H.E. Dr Kao Kim Hourn, Secretary-General of ASEAN China-ASEAN Digital Ecosystem Cooperation Forum" (Association of Southeast Asian Nations, January

na-ASEAN Symposium, both parties reaffirmed shared priorities revolving around privacy, bias, and ethics. China also pledged support for joint infrastructure, capacity building, and regional governance principles.

- India. Through the ASEAN-India Research Training Fellowship and Collaborative R&D Scheme, India supports scientific mobility and joint AI research. These initiatives fund fellowships and short-term projects in AI, data science, and high-performance computing to build capacity and foster regional innovation.

- Japan. Japan is working with ASEAN to co-develop non-Western language models and train 100,000 professionals in AI and advanced technologies.[198] The ASEAN-Japan Digital Work Plan 2025 targets policy alignment, capacity building, and interoperable digital ecosystems.

- South Korea. Korea supports ASEAN through digital talent programs, such as the ROK-ASEAN Digital Academy, and innovation platforms like the AI Development and Startup Competition, addressing challenges in health care and the environment.

- United States. The U.S. has helped develop the ASEAN Responsible AI Roadmap via USAID and cooperates on AI, cybersecurity, and digital governance. However, recent USAID cuts cast uncertainty over future ASEAN-U.S. digital projects.

- Insights based on international engagements.

    - AI governance cooperation is often embedded in broader digital economy partnerships. While this fosters synergies, it risks deprioritizing AI-specific issues in favor of other digital priorities such as cybersecurity or trade.

    - Geopolitical rivalry is a recurring undercurrent in ASEAN's AI partnerships. Japan and South Korea's initiatives are typically seen as counterweights to China's influence, while China's engagement is viewed through the lens of great-power competition. ASEAN must manage these dynamics carefully to benefit from all sides.

4. Key insights and challenges

- Policy momentum has picked up but implementation remains nascent. ASEAN has issued multiple guidance documents and plans, but there is still no harmonized regulatory regime or binding framework to enforce responsible AI governance across the region.

15, 2025), https://asean.org/wp-content/uploads/2025/01/SG-Dr-Kao-Remarks-at-China-ASEAN-Digital-Ecosystem-Cooperation-Forum-15-January-2025_As-Delivered.pdf.

198 "Japan to Work with ASEAN to Train 100,000 Digital Professionals," May 23, 2024, https://english.kyodonews.net/articles/-/48082.

- Capacity and readiness are uneven across the region. The ASEAN Responsible AI Roadmap includes a self-assessment tool to benchmark progress, but disparities in institutional maturity and technical capacity remain significant across member states.

- AI safety governance is frequently diluted within broader agendas. By embedding AI safety policy within wider digital cooperation frameworks, there is a risk that AI-specific governance concerns receive insufficient attention.

- Geopolitical dynamics complicate regional cooperation. ASEAN must balance strategic engagement with all major powers, navigating U.S.-China rivalry without becoming overly reliant on any single actor.

- Sustainability concerns are becoming increasingly salient. As AI-driven digital transformation accelerates, ASEAN is beginning to address the resource intensity of supporting infrastructure, particularly through the new Guide for Sustainable Data Centre Development.

- Traditional AI systems remain the main focus of governance efforts. While generative AI has recently been incorporated into policy discussions, other frontier AI systems like agentic AI and AGI have not yet been considered. ASEAN should broaden the policy discourse to anticipate emerging governance challenges posed by these advanced systems.

2.14. SECTION CONCLUSION

The preceding section highlighted how each Southeast Asian country is actively engaging with AI safety, albeit with different capacities, priorities, and strategic orientations. While there is broad recognition of the importance of AI risk management, national approaches diverge in response to distinct domestic contexts and development goals. These differences sometimes align with, and other times diverge from, emerging global norms led by the U.S., EU, and China. Nevertheless, a shared interest in harnessing AI responsibly provides common ground for co-operation. Building on these insights, the following section outlines policy recommendations to support more cohesive, context-aware, and future-ready AI safety governance across the region.

3.0. RECOMMENDATIONS FOR SOUTHEAST ASIA: 10 POLICY DIRECTIVES FOR 2025–2030[199]

Section 2 outlined the state-of-play of the region on AI safety initiatives and provided country-specific information. While some country profiles differ, key challenges confronting the region were apparent, which include:

- **Uneven AI safety governance capacity across the region.** Disparities in institutional maturity and technical capacity remain significant across Southeast Asia, with countries facing challenges to different degrees such as a lack of quality data, poor cybersecurity, constrained budgetary and governance capacity, limited talent, and regulatory fragmentation.[200]

- **Lack of future-readiness.** Traditional AI systems remain the focus of governance efforts while frontier AI (such as generative AI, agentic AI, and AGI) risk management,[201] catastrophic risk management and AI safety research and development remain policy loopholes.

- **Challenging international outlook.** The region faces an increasingly uncertain international environment. Most countries remain underrepresented on the global stage. At the same time, intensifying geopolitical rivalry between major powers compels the region to navigate external alignments with greater caution.

199 In addition to the authors of this report, Dr. Supheakmungkol Sarin and Zar Motik Adisuryo contributed to this section.

200 See our four characterizing themes in Section 2.1 for a more in-depth analysis on this.

201 Only Singapore and Thailand have adopted generative AI regulatory frameworks in the region. On the regional level, the Expanded ASEAN Guide on AI Governance and Ethics—Generative AI (2025) and the Discussion Paper on the Responsible Development and Use of Generative AI (2025) are recent policy forays into the frontier AI space.

Based on these challenges, analysis points to potential governance and policy paths the region could take to mitigate AI risk. The authors follow an Eisenhower matrix to propose general, high-level priority focus areas for regional stakeholders, helping to concentrate efforts within specified timeframes in Section 3.1. The following sections then detail specific actions that ASEAN and national governments can take toward these focus areas. Section 3.2 specifies four directives for ASEAN at the intraregional level from 2025 to 2028 while Section 3.3 identifies six directives at the national level from 2025 to 2030. To make these policy contributions compatible with existing regional-level proposals, each Eisenhower matrix focus area and recommendation is associated with an ASEAN Responsible AI Roadmap (RAIR) pillar.[202] Appendix C also provides an overview of each recommendation in a tabular format.

202 The pillars of the recommendation are split into "cross-cutting" (CC) pillars for regional actions/priorities, and "targeted" (T) pillars for national actions/priorities. Verbatim, the four CC pillars are Skills & Capacity Building (C1); Fairness and Inclusion (C2); Governance and Participation (C3); and Integration and Cooperation (C4). The four T pillars are Internal Governance Structures and Measures (T1); Skills & Knowledge for Responsible AI-Augmented Decision Making (T2); Risk Mitigation, Monitoring Mechanisms, and Operations Management (T3); and Stakeholder Coordination & Regional Cooperation on AI (T4).

TABLE 3

Mapping of key challenges identified, general Eisenhower matrix focus areas, and specific recommendations outlined in this section[203]

<table>
<tr><th>Key challenges (Section 2)</th><th colspan="3">Uneven AI safety governance capacity</th><th colspan="3">Lack of future-readiness</th><th>Challenging international landscape</th></tr>
<tr><td>General focus area (Section 3.1)</td><td>Develop a regional talent pool</td><td>Level all countries up to a baseline standard</td><td>Harmonize and deconflict national policy</td><td>Research and development on AI safety</td><td>Develop catastrophic management strategies</td><td>Bolster frontier AI preparedness</td><td>Increase international representation</td></tr>
<tr><td rowspan="5">Specific recommendation (Section 3.2/3)</td><td colspan="5">Establish formalized collaborative mechanisms on AI safety</td><td rowspan="5">Develop regional common positions on catastrophic risk management</td><td rowspan="5">Integrate frontier AI risk management into governance</td><td>Approach external relations pragmatically and strategically</td></tr>
<tr><td colspan="5">Structure policy priorities around readiness, niches, and complementarities</td><td rowspan="4">Increase international engagement</td></tr>
<tr><td colspan="2">Pick low hanging fruits in policy development and capacity building</td><td rowspan="3">Prioritize harmonization and interoperability over a single unified framework</td><td rowspan="3"></td></tr>
<tr><td></td><td>Clarify AI safety decision making process</td></tr>
<tr><td></td><td>Focus on policy implementation and data collection</td></tr>
<tr><td>ASEAN RAIR pillar</td><td>C 1
T 2</td><td>C 1, 4
T 1, 2, 3</td><td>C 4</td><td>C 1, 4
T 2, 4</td><td>Absent</td><td>C 1
T 1, 2, 3, 4</td><td>C 4</td></tr>
</table>

Less urgent but important (to plan on now, act later) | Urgent but less important (to delegate) | Urgent and important (for immediate action)

203 Some recommendations are multifaceted and thus span across multiple challenges and focus areas.

3.1. PROPOSED REGIONAL PRIORITIES

This section methodically adopts an Eisenhower matrix that captures the varying levels of urgency—in other words, time-sensitivity—and importance—the criticality toward achieving the goal of developing a robust AI safety governance framework in the region—of focus areas for Southeast Asia stakeholders.[204] The presented focus areas are meant to be high level and general, with specific policy recommendations to follow in the succeeding sections.

Three caveats should be kept in mind:

- This categorization treats each next step as a discrete action, independent of other priorities. However, in reality, priorities are often deeply interconnected—progress in one area can drive progress in another or may depend on prerequisites elsewhere.

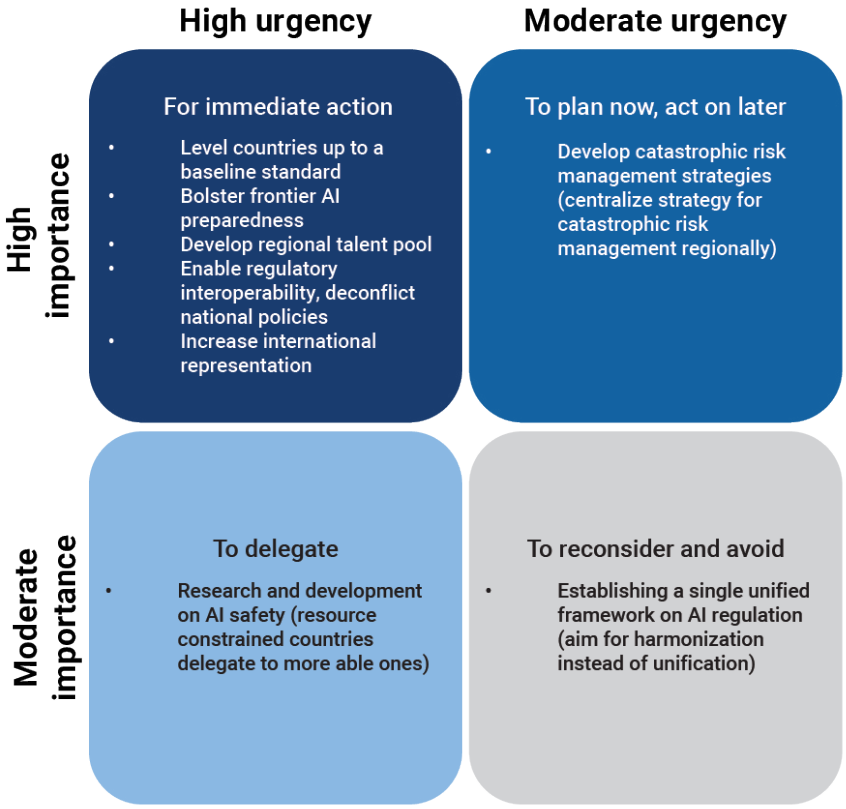
- Priority levels of action items can fluctuate across time and actors as tasks get completed and as operating environments change.

- The use of "immediate action" and other temporal markers reflects the urgency of a course of action. Specific recommendations for how long a next step should take might not be feasible since each step takes varying amounts of time.

204 Note that this Eisenhower matrix only represents action item prioritization for Southeast Asian stakeholders. While an action might be extremely important and urgent on a global level, we might not necessarily think that it is equally urgent and important for the region. For example, consideration of catastrophic risks was considered by the June 2025 UN High-Level Expert Panel on AGI to be very urgent and important on the global level. However, this report argues that for Southeast Asia, strategic sequencing is paramount; foundational capacity must be built first to effectively manage future risks (justification for this below).

FIGURE 2

Eisenhower matrix-style general focus area prioritization



High urgency, high importance

Level all countries up to a baseline standard of AI safety governance. Countries across Southeast Asia lie on a wide spectrum of AI readiness. Given the transformative and fast-evolving nature of AI, the region should galvanize action toward a baseline standard of AI safety governance. This involves two steps: first, defining what such a "baseline" entails;[205] second, developing practical measures to help countries reach it. Doing so would reduce the risk of cross-border spillovers, foster regional inclusivity and goodwill, and strengthen investor confidence. As some countries may face competing developmental priorities, the region must make basic AI safety tools readily accessible to level the playing field.

- ASEAN RAIR link. C1, T2

Bolster frontier AI preparedness. Traditional AI systems have primarily been the focus of con-

205 See ASEAN-level recommendation 1 for details of what such a "baseline" standard could include.

cern for Southeast Asian countries so far. However, consideration of frontier AI systems—such as generative AI,[206] agentic AI,[207] and AGI—remains a blind spot.[208] This is dangerous as these systems have the potential to be even more transformative and simultaneously pose greater societal and institutional risk.

- ASEAN RAIR link. C1, C4, T1, T2, T3

Develop a regional talent pool. A robust regional AI ecosystem necessarily requires experts with local expertise and solid training in relevant disciplines—not just applied AI fields, but also AI governance, safety, and other adjacent areas. Yet, developing AI talent in Southeast Asia remains difficult due to factors such as low wages and the lack of a talent development pipeline. Considering the long-term horizon needed to train AI talent, countries need to focus on this next step now so it can reap the benefits of their AI talent pool as soon as possible in the near future.

- ASEAN RAIR link. C1, T2

Enable regulatory interoperability and deconflict national policy. Southeast Asian countries are increasingly prioritizing AI safety governance and drawing from a wide range of international policy options, many of which may not be interoperable. Given the diversity of AI policies that are quickly being adopted across the region, harmonizing standards is crucial to ensuring regulatory interoperability and enabling countries to leverage each other's complementarities.

- ASEAN RAIR link. C4

Increase international representation. As the maxim goes, "if you're not at the table, you're on the menu." Similarly, if Southeast Asian countries do not get a seat at the table in the drafting of international AI governance standards, "at best, [the region] can only hope to become an AI consumer" (as was intimated by an ex-minister at our roundtable discussions). In the worst case, Southeast Asia might get exploited by AI powerhouses and Big Tech firms. Rapidly increasing representation at international forums such as AI summits and standard setting communities like the ISO and IEEE is crucial, as AI regulations are being drafted now. Delaying action risks missing the region's window to influence global affairs in this domain.

- ASEAN RAIR link. C4

206 Generative AI, as defined in the ASEAN AIGE Expanded Guide on Gen AI, is a "family of technologies that is distinct [from traditional AI] because it uses some of the same fundamental machine learning approaches, in combination with a large dataset, to produce content." "Expanded ASEAN Guide on AI Governance and Ethics."

207 AI systems that can act autonomously without constant human intervention.

208 Case in point, Singapore and Thailand are so far the only Southeast Asian countries to have come up with generative AI risk management procedures.

High urgency, moderate importance

R&D on AI safety. AI safety research and development is time-sensitive and urgently needs to be conducted for the international community to manage risks of novel AI systems. Research can also position countries as global thought leaders, enhance their credibility, and ensure their voices are taken seriously on the international stage. Yet, a high amount of resources needs to be consistently allocated to research initiatives,[209] making this difficult for all Southeast Asian countries to undertake. To maximize regional efficiency, this report recommends a "lead-and-leverage" model. Countries with greater capacity or with existing AI safety research infrastructure, such as Singapore, Thailand, and the Philippines, could take the lead in AI safety research and development and share their findings and other resources. Others could then leverage a second-mover advantage—adopting proven approaches without bearing the full costs of early-stage research—freeing up resources to focus on more immediate domestic priorities.

- ASEAN RAIR link. C1, C4, T2, T4

Moderate urgency, high importance

Develop catastrophic risk management strategies. While addressing catastrophic risk is a critical global issue, it is a less immediate priority for Southeast Asia given the developmental stages of many countries in the region, and the unsettled status of catastrophic risk management strategies even on a global level. A more realistic and pragmatic approach could be to initiate coordinated regional discussions aimed at developing a common ASEAN position on catastrophic risks. By delegating the consideration and preparedness of catastrophic risks to the regional level, national governments would then free up capacity to focus on other domestic priorities without disengaging from global catastrophic risk developments. A unified ASEAN stance would ensure continued regional awareness and preparedness, as well as amplify

209 For comparison, in February 2024, the U.K. announced plans to spend £100 million on AI safety/security research until 2030. "US, Britain Announce Partnership on AI Safety, Testing," CNBC, April 2, 2024, sec. Technology, https://www.cnbc.com/2024/04/02/us-britain-announce-partnership-on-ai-safety-testing.html; "The AI Safety Institute Network: Who, What and How?," Centre for Future Generations (blog), September 10, 2024, https://cfg.eu/the-ai-safety-institute-network-who-what-and-how/. Meanwhile, the EU AI Office and Singapore's Digital Trust Centre (the EU and Singapore's AI Safety Institute) have been allocated €46.5 million and $37 million (USD) respectively. "The AI Safety Institute Network: Who, What and How?" Centre for Future Generations, October 9, 2021. https://cfg.eu/the-ai-safety-institute-network-who-what-and-how/.

the region's international influence.

- ASEAN RAIR link. Absent in RAIR.

"We need to **capitalize on the huge market** that ASEAN is and form a bloc so that we can **negotiate and demand for accountability [and change]** that we would like to see."

Dr. Jun-E Tan, Senior Research Associate, Khazanah Research Institute, Kuala Lumpur Roundtable

While some priority areas relate to more traditional concerns the region has already been focusing on (such as institutional capacity-building measures and nurturing a regional talent pool), other areas—particularly bolstering frontier AI preparedness and catastrophic risk management strategies—have been relatively unexplored.

To address the critical gap in these fields, ASEAN could establish a high-level special taskforce to prioritize foresight, early-warning mechanisms, and strategic capacity-building on frontier AI and catastrophic risk management. Such a taskforce, potentially called the Southeast Asia Frontier AI and Emergency Risk (SAFER) Special Taskforce,[210] could oversee the adoption of a strategic plan from 2025 to 2028 to promote anticipatory governance rather than reactive regulation to strengthen scalable regional resilience and strategic agency in the context of rapidly advancing frontier AI. It could embark on projects such as:

- Identifying tangible frontier AI and catastrophic risks that directly impact governments and other stakeholders
- Facilitating scenario-based preparedness planning exercises based on the identified risks, and then creating actionable steps toward implementation
- An interdisciplinary observatory to monitor developments in agentic AI and trajectories toward AGI that issues timely regional risk assessments
- A dedicated diplomatic track that would enable ASEAN member states to coordinate positions on global AI safety norms in areas such as compute and data governance
- The development of an ASEAN technical corps to pool experts on emerging frontier and catastrophic AI risks for increased representation of the region in international standard-setting forums.

210 We use "emergency" risk here to refer broadly to catastrophic risks as a matter of expediency to fit the SAFER acronym.

Care should be taken to prevent bureaucratic inefficiency surrounding the SAFER Special Taskforce. After all, its purpose is to institutionalize regional consideration of frontier and catastrophic risks in a streamlined way that avoids overburdening member states, enabling them to focus on AI safety fundamentals domestically. Toward this aim, the taskforce could be integrated with existing ASEAN AI institutions like the WG-AI or could report directly to ADGSOM/ADGMIN. A clearly defined and limited mandate specifically on frontier and catastrophic AI risk management, and review periods coinciding with planning horizons adopted by other ASEAN AI-related entities should also be considered to minimize inefficiencies. For instance, a clause could be baked into the SAFER Special Taskforce mandate to require it to sunset or merge into the WG-AI after a review period unless member states renew it by consensus. Additional funding for this body could come from existing ASEAN Dialogue Partner windows like Japan's ASEAN Integration Fund and the EU's Global Gateway ASEAN envelope so as not to strain ASEAN funding.

3.2. ORIENTING ASEAN-LEVEL INTRAREGIONAL DIRECTIVES ON AI GOVERNANCE: 2025–2028

In line with the action areas identified in the Eisenhower matrix above, ASEAN can consider the following four recommendatory directives:

1. **Prioritize harmonization and interoperability over a single unified framework.** Given the region's diversity, a unified legal structure may be unrealistic. Effective governance instead requires harmonization and interoperability. In practice, this means the adoption of shared principles, reduced gaps and conflicts, and use of regulatory mappings or equivalency standards. Importantly, ASEAN member state frameworks must cohere with each other and with global AI governance arrangement. Such a strategy prevents policy duplication and discrepancies and would be fit-for-purpose in the current global geopolitical landscape, where regulatory frameworks of competing powers seem to be diverging. As first steps, ASEAN can:

    - **Develop a regional playbook for regulatory interoperability.** This playbook should build upon existing ASEAN documents (notably, the AIGE 2024, Extended Guide 2025, and the forthcoming RAIR) and could be developed by the WG-AI or a specialized task force. It should consider questions like (i) establishing clear equivalence rulings on AI safety standards,[211] and (ii) outlining common protocols for enforcement, incentive structures, and compliance monitoring. A key prerequisite for this is establishing a

211 Establishing standards that map regulations across jurisdictions is particularly important as the regional adoption of AI policy is increasing in pace and diversity. For instance, Thailand and the Philippines have considered hard legislation on AI regulation, while Singapore remains more inclined to stick to voluntary guidelines for the moment. Tools like Singapore's Project Moonshot could also be recommended as a common-sense and easy-to-use evaluation tool for generative AI systems.

shared, real-time understanding of the current policy landscape. Foundational mechanisms like the SEA AI Policy Observatory (https://seaobservatory.com/) can serve as this common reference point, enabling member states to benchmark policies and identify priority areas for harmonization.[212]

- **Concretize the AIGE and its Extended Guide on Generative AI's principles in a policy document outlining general overarching principles for AI development.** The ASEAN Guide on AI Governance and Ethics (2024) and the Extended Guide (2025) principles are merely voluntary standards as of now. An "opt-in" pilot to allow member states the flexibility of adopting the principles at their own pace might be a good, gradualist approach toward concretizing these standards across the region

- **Identify areas of complementarity within ASEAN-level AI governance documents and between global frameworks.** For example, outline (i) how principles in the AIGE and its extension can be viewed in tandem with those of other global players like the OECD AI Principles; and (ii) how self-assessment methodologies enclosed in the ASEAN AI Roadmap 2025–2030 can co-exist with UNESCO's AI Readiness Assessment Methodology, which seven ASEAN countries have participated in. Documents akin to the U.S. NIST AI RMF Crosswalks could be models for ASEAN to follow.[213]

- ASEAN RAIR link. C4

2. **Approach relations with external partners pragmatically and strategically.** ASEAN's AI safety partnerships with external Dialogue Partners are often enmeshed within broader collaborative frameworks and tinged with geopolitical competition. In an era of great-power competition and deglobalization, ASEAN should cultivate a diversity of partnerships to navigate disruptions while keeping them focused and strategic.

    - An example of how the region can hedge is by formulating cooperative agreements with the International Network of AI Safety Institutes (which includes the U.S. and other Western countries),[214] IndiaAI Safety Institute,[215] and the China AI Safety and Development Association (CNAISDA).[216] These AI safety groupings currently have limited

212 Supheakmungkol Sarin, Shi Hao Lee, Lyantoniette Chua, Philip Tham, Edward Tsoi, "The SEA AI Policy Observatory: Foundational Intelligence for Ethical AI Governance", UNESCO Global Forum on the Ethics of AI, 2025

213 "Crosswalk Documents," NIST AI Resource Center, accessed July 3, 2025, https://airc.nist.gov/airmf-resources/crosswalks/.

214

215 This institute was announced in January 2025 as part of India's broader mission to address safety challenges within its diverse socioeconomic and linguistic landscape.

216 CNAISDA is a consortium of multiple institutions rather than a singular government office. Tsinghua University and the China Academy of Information and Communications Technology (CAICT) are some key players in the consortium. Scott Singer, Karson Elmgren, and Oliver Guest, "How Some of China's Top AI Thinkers Built Their Own AI Safety Institute,"

interaction with each other, and ASEAN could prove to be a vital interstitial connection between them.

- ASEAN RAIR link. Absent in ASEAN RAIR, which focuses on intra-regional recommendations.

3. **Establish formalized collaborative mechanisms on AI safety.** Such mechanisms would foster a synergistic regional AI ecosystem, increase regional goodwill and help countries leverage each other's strengths in AI safety regulation. Regional alignment also strengthens collective bargaining power in global AI governance. Collaboration should span across multiple fields like AI safety research and development, and capacity building.

    - AI safety R&D. Proposals to enhance collaboration on AI safety research and development, such as Malaysia's AI Safety Network, are good first steps in this direction but require greater thought.

    - Capacity building. As mentioned in Section 3.1, ASEAN should focus on leveling all countries up to a baseline standard of AI safety governance.

        - First, the ASEAN WG-AI should come up with what such minimum viable standards should be. Such a baseline could include having a national AI strategy, comprehensive soft regulation on AI (such as AI governance and ethics principles), and policy consideration of frontier AI systems such as AGI, agentic AI, and generative AI. On the latter point, a possible baseline standard could be the adoption of technically grounded "red lines" for AGI development and deployment.[217] These minimum standards could also leverage the existing ASEAN AI Roadmap self-assessment methodologies to serve as a regulatory readiness checklist.

        - Second, ASEAN should then focus on leveling everyone up to this baseline. The ASEAN AIGE and its Extended Guide on Generative AI are good examples of regional boilerplate policy templates that are useful for national adoption (see Brunei for example).

        - Third, similar guides could be formulated for frontier AI systems like agentic AI and AGI. ASEAN could also promote training modules for national civil servants to boost frontier AI and digital literacy among decision makers and develop an AI

Carnegie Endowment for International Peace (blog), accessed July 3, 2025, https://carnegieendowment.org/research/2025/06/how-some-of-chinas-top-ai-thinkers-built-their-own-ai-safety-institute?lang=en.

217 This suggestion was inspired by the UN High-Level Expert Panel on AGI's recommendations published in June 2025 which included mention of red lines to bound AGI development with the intention of preventing uncontrollable AI systems. "Governance of the Transition to Artificial General Intelligence (AGI) Urgent Considerations for the UN General Assembly."

expert talent pool.[218] This kit, composed of guidelines and training modules, could serve as a primer that comes from the SAFER Plan and Special Taskforce.

- **Regional data and compute sharing.** Given regional asymmetries in data collection and compute resources, ASEAN can establish collaborative mechanisms to enable secure, equitable access to data and compute. For instance, safety-critical datasets (like accident logs from self-driving cars, health care AI misdiagnosis cases, LLM prompt-injection exploits) and public interest data (like environmental data) could be shared among ASEAN members for safety research.

- ASEAN RAIR link. C1, C3, C4.

4. **Develop a regional common position on frontier and catastrophic risk management. National governments in Southeast Asia face resource constraints and more immediate development priorities that limit attention to frontier and catastrophic risk.[219]** Frontier risks preparedness can be escalated to the ASEAN level for better coordination among the member states in the region. At the same time, "outsourcing" frontier and catastrophic risk monitoring and ideation to the regional level might also free up resource-constrained governments, allowing them to focus on immediate concerns.

   - ASEAN mechanisms for cooperation. Efforts like the ASEAN SAFE Network initiated by the Malaysian government during its chairmanship is a good step toward better frontier risks research coordination. Our proposed ASEAN SAFER Special Taskforce could also facilitate this recommendation by (i) acting as a shared platform to develop regional positions, and (ii) providing countries with regular updates on catastrophic risks globally. Common positions can then be projected onto the global stage, allowing national governments to prioritize domestic efforts on managing near-term societal impacts.

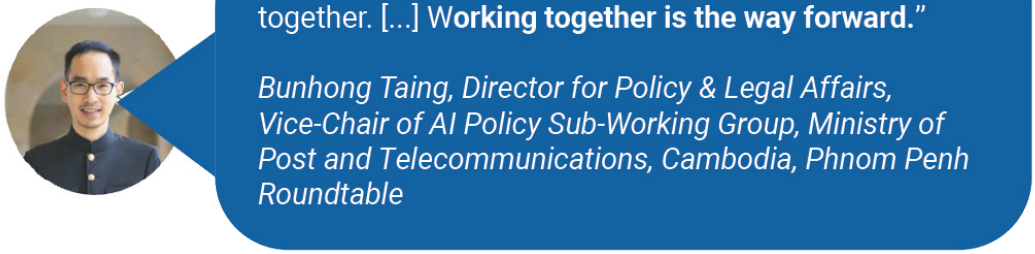
   - Step-by-step approach to managing frontier and catastrophic risks. Consideration of frontier and catastrophic risks on the regional level should take a gradualist approach. Governments could start by first identifying the socioeconomic implications of frontier AI to make the importance of frontier and catastrophic risk management clear to governments, then moving on to scenario-based planning to enhance governance capabilities to address these novel risks, and then finally developing strategies to implement previously planned proposals.

---

218 The Digital Governance, Cybersecurity, and AI Adoption for Public Service program by the Singapore Cooperation Program is one example of this. "Digital Governance, Cybersecurity and A.I. Adoption for Public Service," Singapore Cooperation Programme, accessed July 3, 2025, https://scp.gov.sg/startpublic/courses/Courses/CLS_38/9540/Digital%20Governance,%20Cybersecurity%20and%20A.I.%20Adoption%20for%20Public%20Service/0/CLS_38.

219 This view is echoed in a survey conducted in tandem with the AISA roundtable series, where close to 70 percent of participants favored government focus on societal risks over existential (catastrophic) ones (see Appendix B for details).

- ASEAN RAIR link. Absent in ASEAN RAIR, which neglects catastrophic risk management.



Deliverables in the four recommendations above can be realized within the immediate to medium horizon from 2025 to 2028. This window coincides with a pivotal regional leadership moment, as Malaysia (2025), the Philippines (2026) and Singapore (2027) assume the ASEAN chairmanship. These three countries are actively engaged in AI diplomacy and digital governance, making this an opportune period to shape regional norms and cooperation frameworks.[220] Furthermore, regional governance requires significant diplomatic lead time—consensus-building, alignment across regulatory systems, and relationship cultivation with global partners. Laying down three-year regional orientations in this pivotal period ensures Southeast Asia has a unified and strategic voice in the rapidly evolving global AI governance landscape. A fallback "ASEAN Minus X" implementation pathway[221] could be considered for the specific policy area of AI safety governance so early adopters can move without full–block sign off given the urgency and criticality of some AI risks.

220 See Appendix D for an excerpt from an opinion editorial on what Malaysia, the Philippines, and Singapore could prioritize for their respective chairmanships of ASEAN from 2025–2028.

221 ASEAN operates on a consensus-based decisionmaking approach. For decisions related to the ASEAN Economic Community however, a more flexible "ASEAN Minus X" approach could be adopted, where two or more member states can move forward with an agreement or initiative, even if not all members are ready or willing to participate initially. Joel Ng, ed., "Extending the 'ASEAN Minus X' Formula," in Contesting Sovereignty: Power and Practice in Africa and Southeast Asia (Cambridge: Cambridge University Press, 2021), 224–52, https://doi.org/10.1017/9781108854320.012.

3.3. ORIENTING NATIONAL DIRECTIVES ON AI GOVERNANCE: 2025–2030[222]

Each national government faces different conditions and priorities. The six general policy directives below provide a flexible, adaptable approach toward overcoming the challenges identified earlier. As much as possible, specific recommendations are classified according to the ASEAN RAIR "readiness tiers" of Advanced, Promising, and Emerging.[223]

5. **Pick low-hanging fruit in policy development and capacity-building.** Most countries in Southeast Asia face constrained budgetary and governance capacity. Yet, AI safety governance is an expansive policy field. Resource-constrained governments (i.e., Emerging or Promising AI ecosystems) should thus prioritize low cost and high reward initiatives to quickly level up, especially if they are starting from a low baseline.

    - Regulatory low-hanging fruit. The ASEAN AIGE and the Discussion Paper on the Responsible Development and Use of Generative AI distills AI governance principles and policy toolkits from ASEAN member countries and the international community. These stand as low-hanging fruit of AI policy that countries without comprehensive soft regulation on AI can consider adapting (see Brunei's policy contextualization efforts). Evaluation tools like UNESCO's RAM, and ASEAN's self-assessment frameworks in its AI Roadmap stand as simple, practical mechanisms for assessing readiness, identifying gaps, and guiding responsible AI adoption. Regional micro-grants could be offered with funding from Dialogue Partner programs such as Japan's ASEAN Integration Fund and EU's Global Gateway ASEAN envelope to offset costs incurred from RAM and self-assessment audits.

    - Capacity building low-hanging fruit. To bolster digital literacy and awareness of the risks of AI systems among the civil service, governments could consider sending civil

222 In addition to the authors of this paper, Nitya Kuthiala also contributed to this section.

223 This three-tiered AI readiness national AI ecosystem categorization system is based on a self-evaluation checklist outlined in Annex Table 4 of the RAIR. The RAIR proposes recommendations targeted toward each. "ASEAN Responsible AI Roadmap (2025–2030)."

servants for training programs.[224] Establishing a centralized office dealing with AI safety governance affairs might also be a quick and efficient way to bolster capacity (see point nine below).

- ASEAN RAIR link. T1, T2 (specifically E2.7 on building foundational digital skills and knowledge), T3 (specifically E3.7 on building risk mitigation and monitoring mechanisms).

6. **Integrate frontier AI risk management into AI safety governance.** Policy in the region has largely focused on traditional AI, leaving frontier AI risk management as a blind spot. Yet, this latter category, stands to pose even more transformative and potentially dangerous impacts on societies.[225] Our proposed SAFER Special Taskforce focusing on frontier and catastrophic risks could help coordinate regional assistance to national governments on this front as well. To move forward, all national governments across the RAIR tiered spectrum could consider:

   - Building foundational capacity on AI frontier preparedness and risk management. To ensure civil servants keep abreast of frontier developments, consider integrating frontier AI risk topics into existing digital policy training programs.

   - Institutionalizing risk assessment processes. Consider calling for periodic risk reviews of frontier AI developments, particularly in critical sectors such as national security, education, and health care, while leveraging ASEAN-level AI safety research coordination and exchange. This would strengthen internal capacity and position governments to respond effectively to the rapidly evolving frontier AI challenges.

   - Leveraging regional blueprints for national regulation. Use ASEAN's Extended Guide on Generative AI, recommendations from the Discussion Paper on generative AI, and future ASEAN-level documents on frontier AI as a low-hanging fruit for national regulation adoption.

   - ASEAN RAIR link. T1, T2, T3.

7. **Structure policy priorities around readiness, strategic niches, and complementarities.** While gaps in current AI safety policy exist across the board, roundtable panelists stressed that governments should not rush to address everything all at once. Instead, national policy development should be prioritized based on several factors.

---

224 See footnote 218 on Singapore's training program.

225 The aforementioned UN High-Level Expert Panel on AGI has called for "urgent, coordinated international action under UN leadership" to consider AGI risk management. Mara Di Berardo, "High-Level Report on AGI Governance Shared with UN Community," The Millennium Project (blog), May 28, 2025, https://millennium-project.org/high-level-report-on-agi-governance-shared-with-un/.

- Foundational policy first. For emerging AI ecosystems specifically, in line with recommendations from roundtable panelists, low-resource developing countries could take a multi-step approach to bolstering their AI safety regime.[226] Basic digital policy that is not AI specific but nevertheless an important prerequisite for AI safety such as data privacy and cybersecurity laws are also immediate gaps that need to be addressed before more advanced initiatives can be adopted.

- Focus on sectors where governance is most critical. Policy efforts should concentrate on sectors where AI governance is most critical (such as high-risk applications) and where national comparative advantages can be leveraged. For example, Thailand may prioritize medical AI, Singapore can build on its leadership in AI safety R&D, and Indonesia's large market offers a valuable data ecosystem.

- Synergies and complementarities with neighbors. Strategic complementarities within the region should actively be sought out by individual countries. Collaborating with neighbors can enhance resilience, share regulatory insights, and pool resources for AI development. Countries should also establish clear legal carve-outs to prevent cross-border AI testbeds from running into data sovereignty and export-control barriers. The region should resist the impulse to compete on identical grounds with leading AI market hubs.

- Case study of the Johor-Singapore Special Economic Zone. The Johor-Singapore Special Economic Zone agreement, signed in January 2025, represents a strategic economic partnership between Malaysia and Singapore.[227] By establishing this collaborative agreement, both nations aim to leverage their complementary economic strengths. Malaysia gains access to Singapore's robust financial and investment ecosystems, while Singapore benefits from Malaysia's land resources and competitive labor market. The agreement is strategically designed to enhance regional economic resilience, enabling both countries to navigate increasingly challenging global trade dynamics more effectively.

- ASEAN RAIR link. T1, T3 (specifically E3.6 on strengthening data protection and privacy enforcement), T4 Stakeholder Coordination & Regional Cooperation on AI.

8. **Focus on policy implementation and data collection.** Even if principles and policies have been published, implementation is key. Governments should (i) ensure that policies are not overly general or high-level, which can hinder practical uptake (see Thailand), and (ii) follow through with or streamline proposed initiatives (see Indonesia's Satu Data or Viet-

226 As suggested by a UNDP representative (see Laos in Section 2.5), countries could invest more resources on immediate priorities like building minimum basic digital literacy and infrastructure before enhancing compute capacity, data, and research capacity.

227 Danial Azhar, "Malaysia, Singapore Announce Deal on Johor Economic Zone," Reuters, January 7, 2025, sec. Asia Pacific, https://www.reuters.com/world/asia-pacific/malaysia-singapore-announce-deal-johor-economic-zone-2025-01-07/.

nam's national innovation centers).

- Talent pipeline programs are a crucial example of policies that are often talked about but rarely formalized and implemented due to practical difficulties (see the Philippines). Secondment, internship, and researcher-in-residence matching programs focusing on AI safety and security could be simple options for governments to consider developing talent.

- Robust, disaggregated data is also essential to inform evidence-based policymaking. Countries with limited data infrastructure but already possessing foundational digital frameworks (i.e., Promising AI ecosystems) should begin by targeting key sectors (e.g., health care, education, or electricity) where data is most critical for policymaking and establish national protocols to guide consistent data collection with appropriate protections in these areas. Data can also be used to develop contextualized AI systems (such as multilingual chatbots like Singapore's Project SEA-LION or Cambodia's Khmer translation tools) to equitably distribute the benefits of AI.

- ASEAN RAIR link. T1, T2 (specifically P2.6 on strengthening data and digital literacy with a focus on usage, processing, and analysis of datasets), T3.

9. **Formalize AI safety decisionmaking processes.**

   - On one hand, there is a need to formalize balanced multi-stakeholder approaches involving the active participation of government, industry, academia, civil society, and the public. Participation should be managed carefully—unstructured consultations (see the Philippines) risk becoming counterproductive to AI safety governance if not complemented by clear frameworks. Governments could formalize modalities for participation in AI governance policy formulation to strike a balance between efficiency, objectivity, and multi-stakeholder deliberation. For instance, formulating clear conflict-of-interest policies,[228] and transparent avenues for receiving industry/civil society feedback could be good first steps.

   - AI safety governance institutional structure could also be streamlined.

     - For Emerging AI ecosystems, while not a necessity, a centralized agency or department that draws synergies from different sectors and coordinates policy might help build robust regulatory frameworks. Several countries (see Indonesia and Thailand) have multiple government entities looking at different aspects of AI

228 This recommendation on setting conflict of interest policies was inspired by a policy paper on the Development of the UN Scientific Panel on AI by Mila-Quebec AI Institute published in March 2025. Their recommendation for the establishment of conflict-of-interest policies with clear guidelines on disclosure for industry-affiliated individuals was meant to insulate the policymaking process from misaligned industry influence. "The Development of the UN Scientific Panel on AI" (Mila-Quebec Artificial Intelligence Institute, n.d.).

safety governance, and centralization could bolster efficient resource allocation and confidence in government.

- For Advanced and Promising AI ecosystems with pre-existing AI safety institutions, integrating AI safety governance departments into existing institutions that develop AI policy from economic, strategic, or security angles might also be helpful. As mentioned in Section 2.1, countries are increasingly aware that safety and innovation are not mutually exclusive, but rather that the former is a crucial prerequisite for the second. Housing government units looking at AI safety, economic strategy, and more under one roof might tighten policy coordination while streamlining costs.

- ASEAN RAIR link. T1 (specifically E1.5 on naming a national AI body/agency to coordinate the implementation of national AI initiatives), T3, T4.

10. **Increase international engagement.** Failure to have a seat at the table where the standards of emerging technologies would mean that the region's interests are neglected. Countries need to step up engagement with international partners, not just within ASEAN, but also in other forums. Our proposed SAFER Special Taskforce's ASEAN technical corps could help smaller member states develop government capacity to also shape international standards as well.

    - Active participation in India's 2026 Global AI Summit will be an important entry point—particularly as governance in the Global South is expected to take prominence. But to meaningfully shape outcomes, ASEAN should not stop at attendance. It should pursue sustained representation—whether through permanent membership in relevant multilateral forums or the formation of a dedicated ASEAN subgroup within such bodies—to articulate shared priorities and influence agenda-setting.

    - Moreover, engaging in the ISO/IEC JTC 1 on AI as active contributors, rather than peripheral participants, will allow ASEAN to co-develop technical standards that reflect its unique social, cultural, and economic contexts.

    - Common positions on standard setting and a unified stance in international negotiations on AI governance could be considered through ASEAN mechanisms like the WG-AI or our proposed SAFER Special Taskforce for frontier and catastrophic risk management. Even as "AI consumers" or "price takers" with limited ability to control the course of frontier AI development, Southeast Asia might still have much to contribute in terms of AI system deployment and regulatory practices.

    - ASEAN RAIR link. Absent in RAIR, which focuses on building regional and domestic capacity and neglects international representation.

"Singapore and ASEAN benefit from global norms—**norms for countries and enterprises to follow**. We want to **contribute to the development of these norms and to offer our points of view.**"

Lee Wan Sie, Cluster Director, AI Governance & Safety, IMDA, Singapore Roundatble

In contrast to the three-year intraregional horizon, the national-level directives are set in a five-year cycle starting with 2025–2030. This aligns with the common planning cycles of national roadmaps across Southeast Asian governments.[229] Once ASEAN-wide orientations are established, individual countries can adapt and implement them in line with their distinct political, economic, and technological contexts. This longer runway allows governments to prioritize relevant sectors, build regulatory capacity, implement policies, and develop institutional structures suited to their needs. It also provides space to integrate lessons learned from early AI governance pilots and from participation in global initiatives—such as India's 2026 Global AI Summit or standard-setting efforts under ISO/IEC JTC 1.

A summary of the key takeaways from this section on recommendations for Southeast Asian AI safety follow.

229 For instance, Malaysia's MyDIGITAL plan goals are set to 2025 and 2030, while Thailand's National AI Strategy and Action Plan spans 2022 to 2027. The ASEAN Responsible AI Roadmap is intended for 2025–2030.

KEY TAKEAWAYS: RECOMMENDATIONS FOR SOUTHEAST ASIAN AI SAFETY

- Uneven AI safety governance capacity, lack of future-readiness, and a challenging international landscape complicate AI safety policy in the region. Leveling all countries to a baseline AI safety standard, bolstering frontier AI preparedness, developing regional talent, harmonizing policy, and increasing international representation are priority areas of high urgency and importance.

- The authors propose the establishment of a Southeast Asia Frontier AI and Emergency Risk-management (SAFER) Special Taskforce to work in tandem with the WG-AI focusing specifically on frontier and catastrophic risk management. This regional-level taskforce would allow ASEAN and national governments to expand consideration of AI risks beyond those emanating from traditional systems, while not overburdening national governments.

- On the regional level, the emphasis should be on harmonizing approaches rather than enforcing regulatory uniformity, while managing great-power dynamics with pragmatism. Strengthening cooperation in areas such as capacity building, research, and catastrophic risk management is also essential.

- National governments should pursue practical, high-impact measures—such as improving policy implementation, enhancing data collection, addressing frontier AI risks, and streamlining institutional processes—while actively engaging with international partners.

4.0. LESSONS FOR THE WIDER WORLD: AI SAFETY GOVERNANCE, THE SOUTHEAST ASIAN WAY[230]

Southeast Asia's approach to AI safety governance offers a compelling model for other regions grappling with the rapid evolution of AI. That said, "AI safety" is not a neutral or universally defined term, but a concept that carries various interpretations and implications depending on technical, cultural, and geopolitical contexts. This section adopts a deliberately grounded, policy-relevant interpretation of the concept focusing on institutional readiness, risk management, and developmental goals, first defining what AI safety governance the Southeast Asian way means, and subsequently exploring the applicability of the region's approach to other contexts.

230 In addition to the authors of this report, Dr. Supheakmungkol Sarin and Zar Motik Adisuryo contributed to this section.

TABLE 4

Summary table of AI safety governance, the Southeast Asian way

Characteristic	Distinctive SEA feature	Examples	Global takeaway
1. Localized government	Tailored to local culture, language, and capacity; pluralistic rather than uniform.	- ASEAN AIGE risk template - Vietnam & Thailand adapting EU/NIST frameworks	Local context matters—customize global models to fit national realities.
2. Regional cooperation	Collaboration as strategy; ASEAN centrality; complementarity across countries.	- ASEAN WG-AI, AI SAFE - Singapore: governance, Malaysia: compute, Philippines: auditing - ASEAN Power Grid precedent	Regionalism can scale governance without enforcing uniformity.
3. Inclusive governance	Co-creation with industry, academia, and global partners.	- Google (Malaysia), UP (Philippines) - UNESCO RAM support in Indonesia (completed), Cambodia, the Philippines, Thailand, Vietnam (in process), Laos, Malaysia, and Timor-Leste (in preparation)	Inclusive processes boost legitimacy and policy relevance.
4. Open-source AI safety tools	Promotes access and transparency while embedding safety standards early.	- AI Verify, SEA-LION (open-source + standards) - Regional reuse encouraged	Open-source tools can make AI safety accessible to all.

4.1. DEFINING AI SAFETY GOVERNANCE, THE SOUTHEAST ASIAN WAY

Far from following global AI governance trends, Southeast Asia is shaping a distinctive approach to AI safety grounded in its political realities, developmental goals, and cultural diversity.

While many principles may echo global norms, the ways they are interpreted, implemented, and institutionalized in Southeast Asia in response to the unique constellation of local challenges are particular to the region. This section distills four defining features of the Southeast Asian approach to AI safety governance, each offering practical insights for policymakers worldwide. These features are discussed below.

Localized governance grounded in pluralism and pragmatism

Unlike more centralized governance models seen in the EU or China, Southeast Asia favors localization, sovereignty, and adaptability. Each country has an awareness of the cultural and developmental diversity of the region, and their profiles reflect a deliberate intent to ensure national legitimacy and practical effectiveness, including:

- Context-specific risk assessment tools. The ASEAN Guide on AI Governance and Ethics,[231] along with national initiatives like Singapore's AI Verify,[232] embody a shift away from generic risk checklists. Instead, they integrate locally meaningful markers of harm, such as ethnic or religious stereotyping, that are often overlooked by global benchmarks.

- Pluralism over regulatory uniformity. With vastly different political systems and

231 "ASEAN Guide on AI Governance and Ethics."

232 "AI Verify Foundation-Building Trustworthy AI," AI Verify Foundation, accessed July 3, 2025, https://aiverifyfoundation.sg/.

institutional capacities across the region, Southeast Asia leans into pluralism, using the ASEAN consensus model to forge coherence without enforcing conformity. This stands in contrast to the EU's harmonized legal instruments.

- Tiered approach embedded in ASEAN-led initiatives. Recognizing that member states vary widely in technological maturity, ASEAN frameworks often (but admittedly not always) go beyond a one-size-fits-all model and include graduated recommendations to enable meaningful participation across the development spectrum. An example of this is the ASEAN RAIR (2025) that elaborates on key target areas in detail for national governments categorized based on their level of AI readiness.[233]

Regional cooperation as a strategic asset

AI governance is not treated solely as a domestic issue but as a shared regional challenge requiring coordination, burden-sharing, and institutional alignment. Southeast Asia's consideration of AI safety governance on a regional level is unique among Global Majority regions—case in point, the Africa AI Council, a rough cognate to the WG-AI, was only formed in 2025 and with little information being available about its agenda and tasks.[234]

- Institutionalized platforms for cooperation. ASEAN bodies like the Working Group on AI Governance (WG-AI), the ADGSOM, and Malaysia's proposed ASEAN AI SAFE initiative are already functioning as nodes of regional policy convergence, despite differing national levels of AI readiness.

- Proven capacity for cross-border infrastructure coordination. Past successes, such as the ASEAN Power Grid and the Johor-Singapore SEZ, provide proof-of-concept for regional AI safety infrastructure, whether in compute sharing, regulatory sandboxing, or interoperable compliance.

- Complementary regional strengths. Rather than competing, countries are exploring specialization: Singapore as a policy and governance bridge; Malaysia as a compute and semiconductor hub; the Philippines in AI auditing and annotation; Thailand in sectoral AI regulation (e.g., health and tourism); and Brunei in religiously informed AI governance. This emerging division of labor is rare globally and uniquely Southeast Asian.

233 "ASEAN Responsible AI Roadmap (2025–2030)."

234 Joanna Wiaterek, Cecil Abungu, and Chinasa T. Okolo, "Building Regional Capacity for AI Safety and Security in Africa," Brookings (blog), accessed July 3, 2025, https://www.brookings.edu/articles/building-regional-capacity-for-ai-safety-and-security-in-africa/.

Inclusive, multi-stakeholder governance

Southeast Asian governments are intentionally designing AI governance processes that bring together diverse actors including academia, civil society, and the private sector. This inclusive model is not only about consultation but about co-creation of norms, risk frameworks, and governance structures. Admittedly, the structure of these multi-stakeholder processes is still a work in progress (see the Philippines), although bright spots do exist.

- Inclusion of Big Tech firms. In contrast to a general anxiety about working with Big Tech firms in the EU, Southeast Asian countries have embraced such collaboration. For instance, Malaysia has enlisted Google's help to create an AI Policy and Skilling Lab (2024) to develop policy recommendations for secure AI. Singapore's Project Moonshot evaluation toolkit for generative AI and other Model Governance Frameworks have been developed in collaboration with Big Tech firms. When managed strategically and with appropriate caution, these Big Tech and government tie ups could accelerate innovation while safeguarding public interest.

- UNESCO and international organizations as capacity builders. In countries like Cambodia, Laos, and Timor-Leste, international partnerships are playing a foundational role in building AI policy capacity, in ways not always seen in other Global Majority regions.

- Building legitimacy through transparency. Big-tent governance builds trust and legitimacy—especially critical in emerging economies where skepticism toward government-led tech initiatives can be high. Southeast Asia's inclusive approach is helping overcome this credibility gap.

Open-source AI safety initiatives

Southeast Asia is embracing open-source AI as a developmental necessity to lower barriers, increase transparency, and share benefits more broadly. In particular, it is making tools such as multilingual LLMs that spread AI benefits equitably and AI safety toolkits more accessible to users across the region via open sourcing.

- Shared technical baselines. Singapore's AI Verify and Project SEA-LION exemplify how Southeast Asia is developing shared technical baselines for safe and transparent AI.

- Culturally attuned model development. Efforts such as Singapore's Project SEA-LION and Cambodia's Translatekh prioritize language and cultural alignment in LLMs. By training models on low-resource languages like Thai, Vietnamese, Bahasa Indonesia, and Khmer, Southeast Asia is explicitly rejecting the dominance of Western or Chinese value sets in foundation models—and charting its own course for value alignment and representation.

In contrast to the EU, Southeast Asia's approach is more inclusive and overarching with less apprehension about involving Big Tech firms. In fact, there's a lot of collaboration with Big Tech for capacity building and AI system development.

In contrast to China, Southeast Asia' approach is more decentralized. Each ASEAN member state is a sovereign country, and ASEAN's role is to coordinate and harmonize. This is why China's instinct is to unify and centralize, while Southeast Asia's imperative is to harmonize.

Similar to the U.S.'s approach, Southeast Asian countries recognize the importance of innovation. But they also are cognizant that safety and innovation are not necessarily mutually exclusive, and that both can be sought after at the same time.

Unique to Southeast Asia is that it is a developing region with real resource constraints, diverse cultures, countries with distinct priorities and different developmental profiles.

4.2. APPLICABILITY TO THE WIDER WORLD

However, Southeast Asia's approach to AI governance is not without challenges. Gaps remain in areas such as international representation, implementation capacity, and inclusive stakeholder engagement. Nonetheless, the region offers useful lessons for others with similar profiles.

These lessons may be especially relevant for other regions in the Global Majority that navigate similar structural and contextual complexities. Like Southeast Asia, many countries in Africa and South America see AI as a tool to advance development goals in areas such as education, health care, agriculture, and employment. Both regions are also characterized by institutional and linguistic diversity, with uneven progress in national AI strategies. In Africa, countries are advancing AI policy at different speeds, shaped by linguistic and legal differences across Francophone, Anglophone, and Lusophone systems. In South America, Brazil, and Chile have taken early steps in policy development, while others are still in the initial stages.

Several aspects of Southeast Asia's approach to AI safety governance could provide useful reference points for other regions.

- Coordinated approach to context-specific LLM projects. Southeast Asian countries have made significant progress in advancing LLM development for local languages, while implementing robust evaluation mechanisms to test prominent models for harmful bias stereotypes about social groups. These efforts are crucial for advancing contextualized AI safety in ways that meet unique cultural and linguistic needs.[235] Project SEA-LION is an example of this. Trained on content produced in regional languages like Thai, Vietnamese, and Bahasa Indonesia, the project aims to ensure better representation in data and value alignment compared to Western or Chinese models, better serving the needs of non-western, industrialized, rich, educated, and democratic

235 Major academic studies have shown that existing LLMs exhibit a bias toward Western, industrialized, rich, educated, and developed (WIRED) countries. Anthropic, an AI company that produced the widely used LLM Claude, found that responses from its chatbot tracked closer to the opinions of populations in the U.S. and Europe vis-à-vis other parts of the world. Esin Durmus et al., "Towards Measuring the Representation of Subjective Global Opinions in Language Models," April 12, 2024, https://doi.org/10.48550/arXiv.2306.16388; Mohammad Atari et al., "Which Humans?," Department of Human Evolutionary Biology, Harvard University, September 2023, https://doi.org/10.31234/osf.io/5b26t. Another paper published by Harvard showed that OpenAI's GPT3 tended to reflect the personalities of WIRED individuals over that of other countries. Atari, Mohammad, Mona J. Xue, Peter S. Park, Damián E. Blasi, and Joseph Henrich. 2023. "Which Humans?." PsyArXiv. September 22. doi:10.31234/osf.io/5b26t.

(WIRED) societies (see Singapore's country snapshot in Section 2.9).

- Academia-private-government partnerships. Southeast Asian countries have made significant efforts to steward local AI innovation through academic-government and private partnerships (see the point on inclusive, multi-stakeholder governance in Section 4.1). To make similar advancements, other countries in the Global Majority must invest significantly in capacity-building efforts to build strong research institutions and establish partnerships to strengthen AI research. Strong multi-stakeholder dialogue also reduces the chances that discussions are dominated by Big Tech influence, potentially opening the door to more bespoke and innovative approaches for the public good. That being said, robust conflict-of-interest safeguards should also be put in place to prevent regulatory capture by external interests.

- Localized AI system safety evaluation metrics. Southeast Asian countries have made significant progress in developing context-specific evaluation metrics to examine AI systems for limitations. For instance, Annex A of the ASEAN AIGE provides a locally created AI Risk Impact Assessment template. Singapore, Indonesia, Malaysia, Thailand, and Vietnam (among others) have also produced AI governance principles and templates of a similar vein. These metrics are essential for enabling a more nuanced understanding of the harms AI systems can enact on culturally and linguistically diverse regions like Southeast Asia, while also moving from boilerplate metrics that accurately represent this diversity.

- ASEAN centrality in regional AI safety governance. While many efforts have been spearheaded by national governments, regional policy is primarily coordinated through ASEAN mechanisms like the WG-AI and ADGSOM and has produced important documents such as the AIGE and the Roadmap on Responsible AI. Pan-regional AI safety institutions in other Global Majority regions, like the Africa AI Council, are often less defined and less integrated with existing bodies—for example, the Africa AI council lacks a clear public agenda and formalized ties to the African Union (AU).[236]

- Open-source AI safety toolkit development. Southeast Asian countries have promoted a focus on the open-source development of AI safety toolkits, which is crucial to improved information sharing, reducing the costs of AI development, and increasing transparency regarding data sources used to train AI systems and the underlying weights of models (see the point on open-source AI safety tools in Section 4.1). Singapore's Project SEA-LION and AI Verify are open-source, encouraging regional uptake and collaboration on these models to spread the benefits of these systems (see Singapore's country snapshot in Section 2.9).

Still, important differences between Southeast Asia and other Global Majority regions must not be neglected.

236 Wiaterek, Abungu, and Okolo, "Building Regional Capacity for AI Safety and Security in Africa."

- Talent distribution. While Southeast Asia and other Global Majority regions face similar issues with a lack of concentration of AI talent, AI hubs are emerging in key cities such as Jakarta, Bangkok, Kuala Lumpur, and Singapore.[237]

- Research capacity. Compared to other developing regions, Southeast Asian countries experience similar challenges in access and quality of education.[238] However, AI research output in Southeast Asia is greater than other regions in the Global Majority, with notable standouts such as the National University of Singapore and the Nanyang Technological University.[239]

- Basic infrastructure. Compared to regions like Africa and Oceania, Southeast Asian countries have less of a need to focus on establishing basic infrastructure (e.g., electricity and telecommunications services), given the already relatively high internet penetration and access to stable, reliable electricity. Within Southeast Asia, internet penetration is estimated at 78.2 percent,[240] compared to 37 percent in Africa,[241] and the global average of 68 percent.[242] The International Energy Agency reports that 95 percent of ASEAN households have electricity,[243] compared to 43 percent in sub-Saharan Africa.[244]

- AI-specific infrastructure. Across the Global Majority, Southeast Asia benefits from a relatively higher concentration of data centers equipped with state-of-the-art GPUs. For instance, Johor Bahru, Malaysia's second-largest city, is projected to be one of the largest data center hubs in Asia if its planned data center capacity comes online.[245]

---

237 Bhaskar Chakravorti et al., "50 Global Hubs for Top AI Talent," Harvard Business Review (blog), accessed July 3, 2025, https://hbr.org/2021/12/50-global-hubs-for-top-ai-talent.

238 Eduardo Velez Bustillo and Harry A. Patrinos, "Four of the Biggest Problems Facing Education—and Four Trends That Could Make a Difference," World Bank Blogs (blog), accessed July 3, 2025, https://blogs.worldbank.org/en/education/four-biggest-problems-facing-education-and-four-trends-could-make-difference.

239 "Artificial Intelligence," Emerging Technology Observatory, accessed July 3, 2025, https://almanac.eto.tech/topics/ai/#countries.

240 "Internet Usage in Southeast Asia," Statista, July 1, 2025, https://www.statista.com/topics/9093/internet-usage-in-southeast-asia/.

241 "Facts and Figures 2023 - Internet Use," ITU, 2023, https://www.itu.int/itu-d/reports/statistics/2023/10/10/ff23-internet-use.

242 "Facts and Figures 2024 - Internet Use," ITU, accessed July 3, 2025, https://www.itu.int/itu-d/reports/statistics/2024/11/10/ff24-internet-use.

243 "Executive Summary-Southeast Asia Energy Outlook 2024," IEA, 2024, https://www.iea.org/reports/southeast-asia-energy-outlook-2024/executive-summary.

244 "SDG7: Data and Projections Access to Electricity," IEA, accessed July 3, 2025, https://www.iea.org/reports/sdg7-data-and-projections/access-to-electricity.

245 Dylan Butts, "Malaysia Is Emerging as a Data Center Powerhouse amid Booming Demand from AI," CNBC, June 16, 2024, sec. Technology, https://www.cnbc.com/2024/06/17/malaysia-emerges-as-asian-data-center-powerhouse-amid-booming-demand.html.

However, improving where these centers are distributed should be a priority.

Southeast Asia differs in key respects from other Global Majority regions, boasting more developed infrastructure, stronger research institutions, and emerging AI hubs. Yet, it ultimately shares more with them than with dominant AI powers like the U.S., China, or the EU. These similarities lie not only in shared challenges such as limited resources, fragmented talent pools, and cultural and linguistic diversity, but also in a shared imperative to shape AI governance models that reflect local needs rather than imported templates.

Compared to other major AI governance models, Southeast Asia's approach is distinct. Unlike the EU, which often adopts a precautionary stance and keeps Big Tech at arm's length, Southeast Asia takes a more inclusive, overarching approach, actively collaborating with technology firms to build capacity and co-develop AI systems. In contrast to China's centralized and state-driven model, Southeast Asia's governance is inherently decentralized, shaped by the sovereignty of its member states and ASEAN's coordination. Harmonization rather than uniformity is ASEAN's imperative. While the region shares the United States' emphasis on innovation, it recognizes that safety and innovation are not mutually exclusive and strives to pursue both in tandem. What ultimately sets Southeast Asia apart is its reality as a developing region: resource-constrained, culturally diverse, and composed of countries at different stages of digital maturity. These complexities shape a uniquely pragmatic, pluralistic, and adaptive approach to AI safety—one that offers important lessons for the rest of the Global Majority.

KEY TAKEAWAYS

- **Southeast Asia is pioneering a distinctive AI safety governance model.** Instead of passively following global norms, its approach is shaped by its cultural diversity, political pluralism, and development spectrum. Key features include localized, context-sensitive governance; regional coordination through ASEAN platforms; inclusive multi-stakeholder processes with pragmatic Big Tech engagement; and open-source, culturally aligned AI tools.
- Southeast Asia's emphasis on **local-language LLMs, open-source safety tools, multi-stakeholder collaboration, and ASEAN centrality** are points that other Global Majority regions with cultural, institutional, and developmental diversity could consider. The region's relatively stronger infrastructure, academic capacity, and context-specific evaluation tools make it a useful reference point for regions like Africa and South America, though differences in talent, research, and baseline infrastructure remain.

5.0. CONCLUSION

This report illustrates how the AI safety landscape in Southeast Asia is developing quickly. Bright spots clearly exist. Based on our roundtable sessions, decisionmakers are increasingly aware of the importance of robust AI risk management. They can bring their lessons learned to the global discourse on AI safety. In particular, Southeast Asian governments seem to be approaching policy development in a clear-eyed and pragmatic manner in collaboration with a variety of stakeholders. Local projects aiming to equitably share the benefits of AI technology are also taking root across the region.

Yet, obstacles undoubtedly remain. Uneven AI safety governance maturity levels across the region might result in differing national priorities and gridlock that could hamper regional harmonization. Consideration of looming frontier AI and catastrophic risks also stands as a major blind spot across the board. Southeast Asia has the added challenge of navigating these issues against the backdrop of an evermore complicated international environment. The suite of recommendations outlined in this report aim to tackle some of these issues while complementing existing proposals.

Southeast Asia's successes, areas for improvement, and possible future directions may resonate with other Global Majority regions. As AI developments occur at lightning speed and global approaches to AI safety governance harden, largely shaped by the U.S., EU, and China, Southeast Asia offers a timely and necessary counterpoint. Its nuanced, context-aware approaches, shaped by diversity and pragmatism, show how governance models can be adapted to different realities without compromising on safety. These distinct pathways expand the global vocabulary of AI safety and show that one-size-fits-all approaches may fall short in a multipolar world. The region's challenges in achieving this aim also primes other governments to be cautious of particular pitfalls.

Ultimately, it is our hope that this report not only offers insight into Southeast Asia's evolving approach to AI safety but also urges global and regional actors alike to recognize, engage with, and build upon the Southeast Asian way.

6.0. APPENDICES

6.1. APPENDIX A: ROUNDTABLE PARTICIPANTS

AISA greatly appreciates the contributions of the panelists and moderators who participated in our roundtable series in 2024 and 2025. Salient points raised during these roundtable sessions formed the basis of the country profiles in Sections 2.2 to 2.13. While the recommendations and lessons for the wider world in Sections 3 and 4 respectively were not directly discussed during roundtable discussions, the authors used insights gleaned from the sessions to develop their ideas. These speakers also reviewed the country profiles that they were invited to speak on.

Note that opinions shared by speakers at the roundtable sessions do not necessarily reflect the opinions of the organizations that they are affiliated with.

An exhaustive list of speakers listed by roundtable session follows.

Roundtable	Speakers
Indonesia **(Sep 2024)**	Esben Kran (Founder of Apart; Co-founder, European Network for AI Safety; Adviser to AISA), Gita Wirjawan (Former Minister of Trade of Indonesia; Chairman and Founder, Ancora Group and Ancora Foundation), Prof. Hammam Riza (President of Kolaborasi Riset dan Inovasi Industri Kecerdasan Artifisial; Delegate to U.K. AI Summit on behalf of Indonesia; Adviser to AISA), Prof. Robert Trager (Co-Director, Oxford Martin AI Governance Initiative; International Governance Lead, Centre for the Governance of AI; Senior Research Fellow, Blavatnik School of Government, University of Oxford; Adviser to AISA)
Philippines **(Oct 2024)**	Alejandro Reyes (Scholar-in-Residence, Asia Society Hong Kong Center; Adjunct Professor, University of Hong Kong; Adviser to AISA), Benjamin Prud'homme (Vice President, Policy, Safety and Global Affairs, Leadership Team, Mila – Quebec Artificial Intelligence Institute), Elmer C. Peramo (Senior Science Research Specialist, Department of Science and Technology – Advanced Science and Technology Institute, Philippines)
Singapore **(Nov 2024)**	Lee Wan Sie (Cluster Director, AI Governance & Safety), Mike Nelson (Senior Fellow, Asia Program, Carnegie Endowment for International Peace), Prof. Mohan Kankanhalli (Director, National University of Singapore (NUS), Artificial Intelligence (AI) Institute), Rina Chandran (Southeast Asia Editor, Rest of World)
Malaysia and Thailand **(Jan 2025)**	Abhishek Vats, (Sustainability Director, ASEAN Youth Organization), Alex Moltzau (Policy Officer, European AI Office of the European Commission), Jam Kraprayoon (Researcher, Institute for AI Policy and Strategy), Dr. Jun-E Tan (Senior Research Associate, Khazanah Research Institute), Raymond Sun (aka TechieRay) (Lawyer, Herbert Smith Freehills; Director, Data Science and AI Association of Australia Technology Lawyer; Developer of Global AI Regulation Tracker)
Brunei, Laos, and Vietnam **(Feb 2025)**	Alex Read (Chief Technical Specialist, UNDP Lao PDR), Danh Nguyen (Global Head of Legal at GreenFuture.tech, Legal-Tech Expert at LegalTech.org), Dr. Hoda Alkhzaimi (Co-Chair, Global Future Council for Cyber Security, World Economic Forum; Adviser to AISA), Isyrah Fahmi Osman (Head of Research and Development, Information Technology Protective Security Services),
Cambodia, Myanmar, and Timor-Leste **(Mar 2025)**	Barani Maung Maung (Tech Policy and Safety Expert; Online Safety Policy Lead, Ofcom), Bunhong Taing (Director of Policy and Legal Affairs, Digital Government Committee), Justo Fernandes (Director of Application and Information, Siriwat William Chhem (Co-Founder and President, AI Forum), Yonah Welker (Public Technologist; Former Tech Envoy, Ministry of AI).

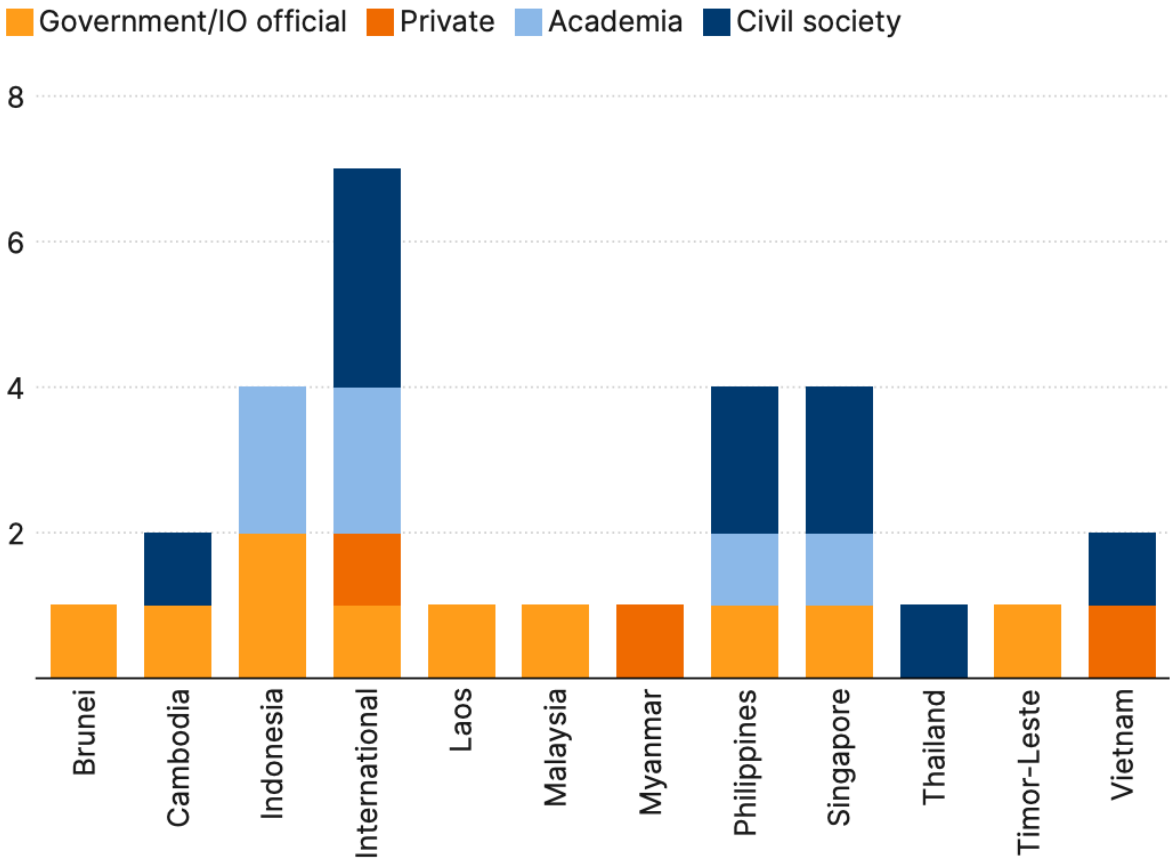
6.2. APPENDIX B: METHODOLOGICAL CLARIFICATION

The regional and country reports in Section 2 were based on salient points raised during roundtable discussions organized by AISA from September 2024 to March 2025. These discussions were structured around the key theme of "AI Safety: X Risks or Spectrum of Risks?"[246] and sought to elucidate regional perspectives on how to balance between existential AI risks, more immediate terms risks, and other policy preoccupations.

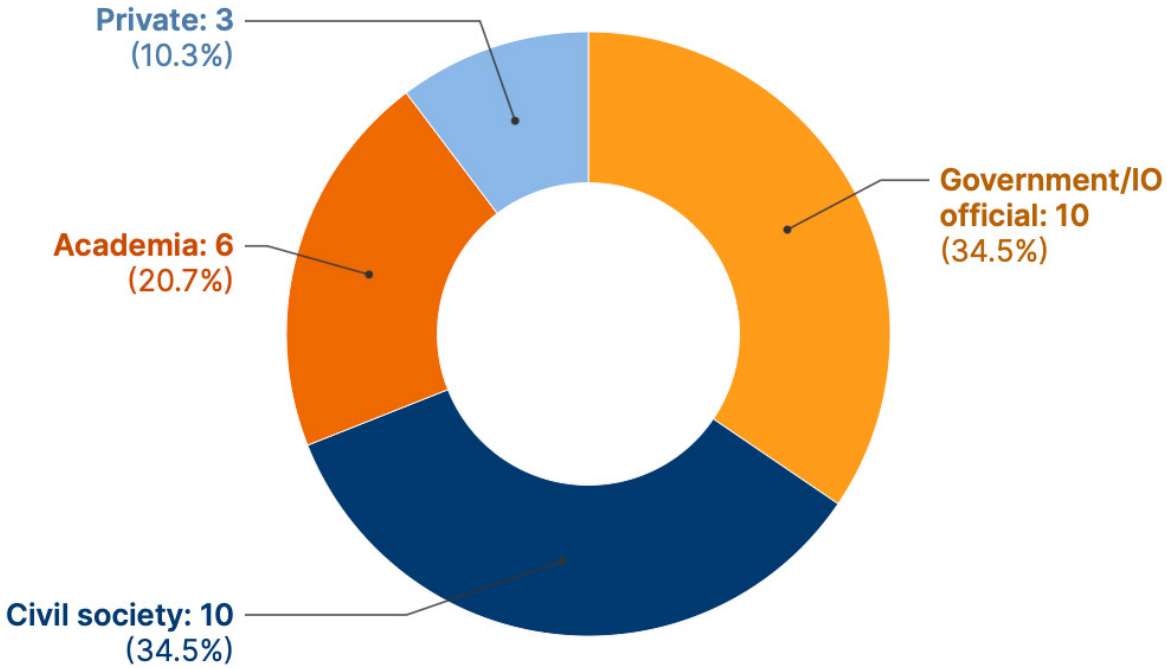
Across all discussions, the research team invited at least one AI safety governance expert to represent the perspective of each Southeast Asian country examined in this report. Panelists were selected based on their demonstrated expertise in AI policy, involvement in their country's digital ecosystem, and their ability to represent a key stakeholder perspective (government, academia, private sector, and civil society). Whenever possible, these representatives were locals of the countries they were speaking about. They were chosen from a mixture of affiliations to gather a diversity of views. Speakers representing an international perspective—for instance, an ASEAN, EU, or Global Majority perspective—were also included to enrich the discussion. A total of 29 panelists were invited across six discussions that focused on (i) Indonesia, (ii) Philippines, (iii) Singapore, (iv) Malaysia-Thailand, (v) Brunei-Laos-Vietnam, and (vi) Cambodia-Myanmar-Timor-Leste. The panelist breakdown by affiliation and country represented is shown below.

246 "X-risks," short for existential risk, refers to the potential for AI to pose a catastrophic and existential threat to humanity. Such a scenario might occur as a result of an "intelligence explosion" or a "technological singularity" wherein AI becomes uncontrollable, unpredictable, and fundamentally misaligned with human values and well-being. This is often juxtaposed against a "spectrum of risks," which encompasses a wider range of risks that do not rise to the level of posing an existential threat to humanity. Such risks include misinformation and disinformation, job displacement, and an erosion of trust in institutions.

Stakeholder breakdown by country representation



Distribution of panelists by affiliation



This methodology of capturing a snapshot of AI safety governance in a country based on conversations with (mostly) local experts of various affiliations is advantageous as it offers a balanced, authentic, and contextualized look into the domestic policy landscape.

While this report draws on insights shared by speakers during the roundtable series, the recommendations and interpretations presented reflect the authors' analysis and are not necessarily the views of the speakers. The content aims to build upon and be inspired by the discussions, rather than represent direct endorsements by individual participants.

One caveat to this approach is that the country snapshots are not meant to be comprehensive. Rather, they are based on salient points raised during the discussions. This means there naturally would be some biases in terms of the perspectives and initiatives discussed in the report. For instance, despite the strong presence of civil society representatives in the panels, these representatives mostly critiqued and analyzed government policy and did not talk much about local civil society action in AI safety governance. While this could have been because local civil society action was limited, it could also have been due to the presence of government officials in the panel who could have motivated the panel discussion to skew more toward government policy action. Another limitation is that some countries received greater attention in the panel discussions vis-à-vis others—Indonesia, Philippines, and Singapore had whole panel sessions focusing on their AI safety governance landscape while the other countries shared sessions with one another. On one hand, the organization of the panel discussions was reflective of the relative maturity of the AI safety governance landscape among the different countries. However, this also might have made certain countries receive less attention.

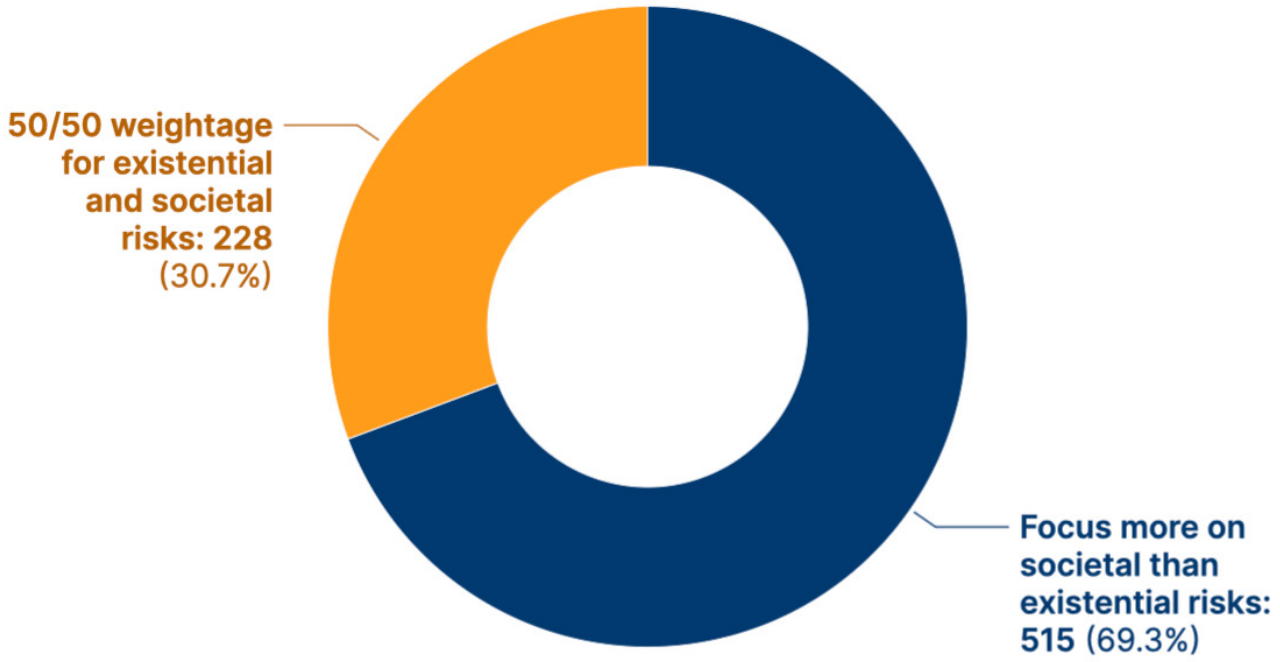
To mitigate these limitations, panelist input is complemented with independent research to plug the gaps. This included a review of official government policy documents, reports from international organizations (e.g., UNESCO, OECD), and reputable academic, private sector, and civil society reports, Each country snapshot—consisting of points raised during the roundtable and information independently sourced by this research team—was then submitted for review by the panelists. Feedback from the panelists was then incorporated into the report once more.

This report summarizes findings from the inaugural series of roundtable discussions focusing on Southeast Asia AISA organized in 2024 and 2025. Looking ahead, AISA aims to continue the momentum of this initial series by convening further roundtables across other subregions in Southeast Asia. The goal of these discussions is to surface regionally grounded insights, stimulate cross-border dialogue, and ultimately paint a picture of AI safety policy—the Southeast Asian way.

In tandem with the roundtable, AISA also asked all participants (listeners and panelists) what their ideal allocation of government resources to x-risk vs. societal risk management was. These participants represented a mix of sectors, including students, researchers, industry professionals, and government staff from across the region. The exact wording of the question and the results are shown in the donut chart below.

Question posed to AISA roundtable participants:

What is the percentage of resources you think should be allocated existential risks versus societal risks?



6.3. APPENDIX C: OVERVIEW OF RECOMMENDATIONS

Level	Specific recommendation	General priority area	Key challenge	ASEAN RAIR pillar
Regional	Prioritize harmonization and interoperability over a single unified framework	Harmonize and deconflict national policy	Uneven AI safety governance capacity	C 4
	Approach relations with external partners pragmatically and strategically	Increase internationl representation	Challenging international landscape	Absent
	Establish formalized collaborative mechanisms on AI safety	Develop a regional talent pool Level all countries up to a baseline standard Harmonize and deconflict national policy R&D on AI safety	Uneven AI safety governance capacity Lack of future-readiness	C 1,3,4
	Develop a regional common position on catastrophic risk management	Develop catastrophic risk management strategies	Lack of future-readiness	Absent

Level	Specific recommendation	General priority area	Key challenge	ASEAN RAIR pillar
National	Pick low-hanging fruit in policy development and capacity building	Develop a regional talent pool Level all countries to a baseline standard	Uneven AI safety governance capacity	T 1,2,3
	Integrate frontier AI risk management into AI safety governance	Bolster frontier AI preparedness	Lack of future-readiness	T 1,2,3
	Structure policy priorities around readiness, strategic niches, and complementarities	Develop a regional talent pool Level all countries up to a baseline standard Harmonize and deconflict national policy R&D on AI safety	Uneven AI safety governance capacity Lack of future-readiness	T 1, 4
	Focus on policy implementation and data collection	Level all countries up to a baseline standard	Uneven AI safety governance capacity	T 1, 3
	Formalize AI safety decision-making process	Level all countries up to a baseline standard	Uneven AI safety governance capacity	T 1, 3, 4
	Increase international engagement	Increase international representation	Challenging international landscape	Absent

6.4. APPENDIX D: EXCERPT FROM OPINION EDITORIAL ON AI GOVERNANCE IN THE ASIA-PACIFIC CENTURY

AI Governance in the Asia-Pacific Century: From Diversity to Cohesion—The ASEAN Way (Chua, Lee, Thum, 2025)

A recommendatory three-year plan for ASEAN chairmanships. Now more than ever, it is imperative that Southeast Asian nations coalesce together and function as an integrated global market system on AI. To solidify ASEAN's position as a proactive player in global AI governance, the authors propose that the leaders and digital ministries of ASEAN countries consider putting in motion a coordinated three-year plan to align the efforts of successive chairmanships toward a constructive trajectory for regional AI governance.

Malaysia, as chair in 2025, could lead with a focus on inclusivity through an ASEAN AI Safety Network. This initiative would unite member states, foster cross-border collaboration to address AI safety risks, and establish foundational guidelines tailored to the region's diverse technological, legal, and cultural contexts.

Building on this, the Philippines, in 2026, could emphasize intergenerationality among high-level experts and key stakeholders with a focus on capacity-building programs for AI governance. By engaging youth, policymakers, and industry leaders, this phase would ensure the development of a skilled, future-ready workforce while embedding regionally-aligned domestic AI governance capabilities into broader socioeconomic priorities encompassing long-term thinking.

Finally, Singapore's chairmanship in 2027 could culminate in the creation of an interoperable governance framework, driven by a Southeast Asia Declaration on AI Governance, accompanied by a 2027–2030 roadmap. This declaration would not only outline shared principles but also provide actionable steps for harmonizing diverse national policies.

Together, these initiatives—grounded in three I's: inclusivity, intergenerational engagement, and interoperability—would position ASEAN as a credible voice in shaping a responsible, resilient, and people-centric AI governance ecosystem in the AI global market chain.

Ultimately, while ASEAN's current capabilities on global AI governance may be limited, its in-

fluence and potential are far from insignificant. Being able to play a constructive role in global AI discourse, however, hinges on ASEAN's ability to get its own house in order. Addressing internal disparities, aligning national and regional policies, and demonstrating a commitment to shared principles will not only enhance the region's standing, but also position it as a credible and constructive partner in shaping the future of AI.


B
Center for
Technology Innovation
at BROOKINGS
AISA
AI SAFETY ASIA