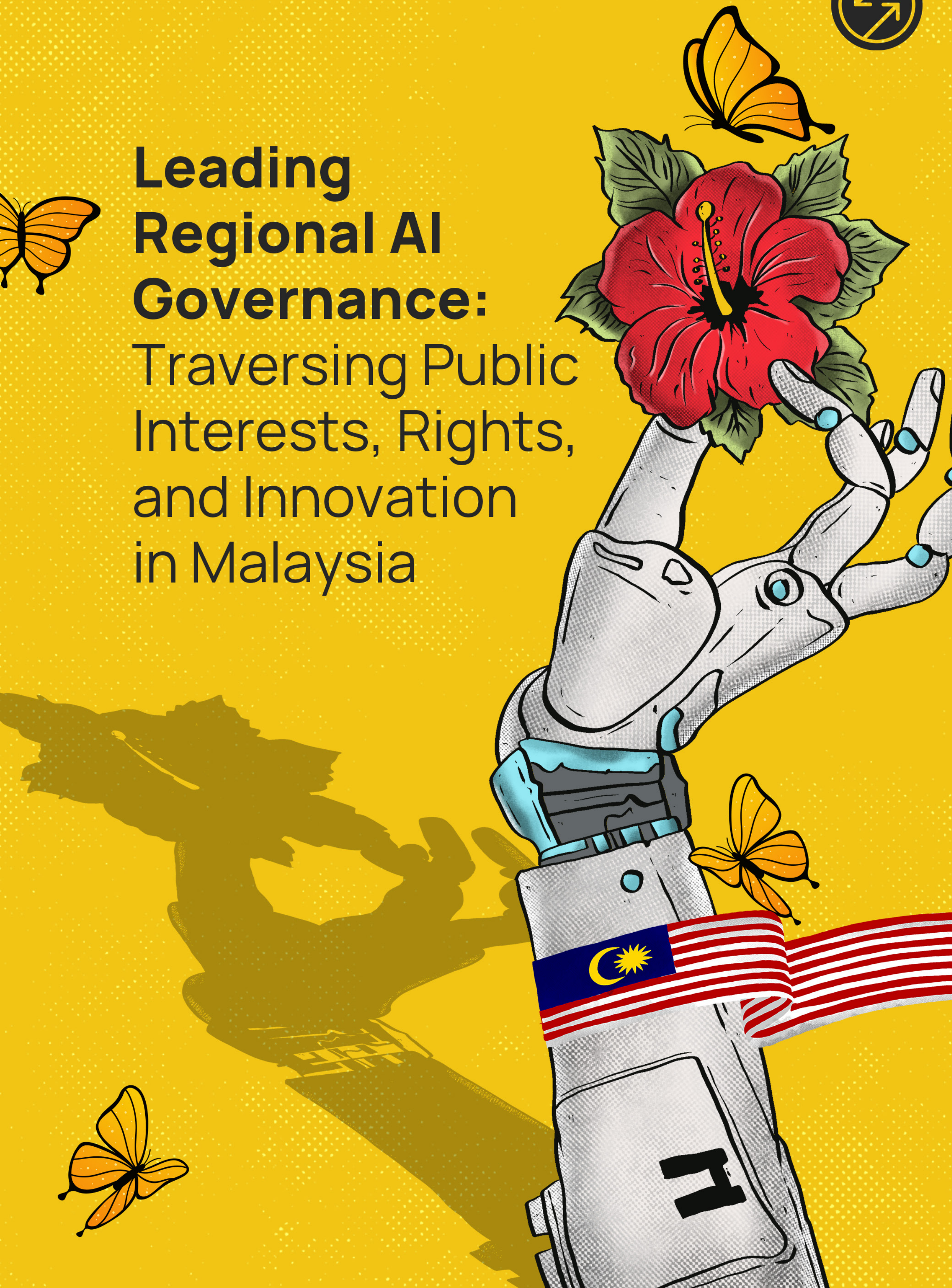
Leading Regional AI Governance: Traversing Public Interests, Rights, and Innovation in Malaysia



Authors:
Dineshwara Naidu
debby kristin

Contributor:
Azura Nasron, SUARAM
C Hari S Shankar, INITIATE.MY
Siti Nurliza S, Sinar Project
Tengku Nur Qistina, Architects of Diversity Malaysia
Center for Independent Journalism (CIJ)

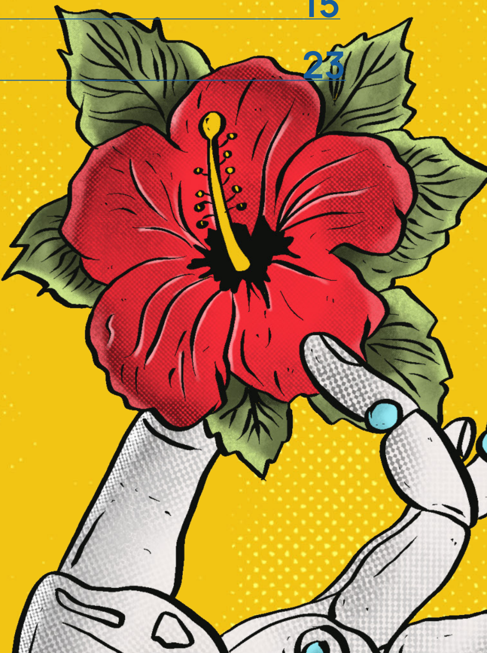
Expert Reviewer
Dr. Jun-E Tan, Khazanah Research Institute
Shabnam Mojtahedi, International Center for Not-for-Profit Law

Editor & Proofreader
Marina Nasution

Layouter
Docallisme Studio

Published in November, 2025

Table of Content



Executive Summary

Malaysia has taken the necessary initial strategic steps to guide artificial intelligence (AI) use, exemplified in national frameworks such as the National Guidelines on AI Governance and Ethics (AIGE), Rancangan Malaysia Ketiga Belas (RMK13)[1] and the National AI Action Plan (AI NAP 2030). While the current national framework offers broad ethical principles, it urgently needs to provide operational assurances.

This paper frames its proposal within Malaysia's existing policy direction and focuses on supporting the achievement of RMK13 and the AI NAP 2030 through proposed recommendations. The government needs to address regulatory reforms in algorithmic accountability, mandatory risk assessments, and cross-sector enforcement, as well as meaningfully engage civil society organisations (CSOs) in the consultative process, drafting, and implementation of these AI-related policies. These steps are central to Malaysia's fostering innovation while maintaining accountability, safeguarding democratic norms, and strengthening its standing as a regional leader in AI adoption.

The consensus among the CSOs involved in the development of this policy recommendation advocates for and recommends several key issues, namely:

1. Upholding the right to privacy: address regulatory gaps around data governance, and establish mandatory self-assessments for companies.
2. Establishing an oversight body and AI accountability: improving AI transparency and accountability in the design, development, and deployment of AI; forming a multistakeholder governance working group and an independent AI safety oversight body, implementing a redressal tracker, and adopting a hybrid governance model.
3. Treating access to information as a key principle of AI governance: applying the principle of maximum and proactive disclosure and establishing an AI registry.
4. Embedding multistakeholder participation: adopting a whole-of-nation

[1] Digital Government Malaysia. "Press Release: Ministry of Digital to Intensify RMK-13 Initiatives." *Digital Government Malaysia.* https://www.digital.gov.my/api/file/file/02082025_PRESS%20RELEASE_MINISTRY%20OF%20DIGITAL%20TO%20INTENSIFY%20RMK-13%20INITIATIVES.pdf

strategy to ensure AI systems genuinely represent Malaysia's diverse society and safeguard fundamental rights.

Malaysia's increasing involvement in regional collaboration, particularly through the ASEAN AI Safety Network (ASEAN AI Safe), adds another layer to this agenda. By linking domestic reforms to these regional initiatives, Malaysia can reinforce rights at home while advancing shared norms and cooperation within ASEAN, supporting its ambition to lead responsibly in the global digital landscape.

At the same time, it would also strengthen Malaysia's commitments under international standards, particularly the Sustainable Development Goals (SDGs), which call for inclusive, participatory, representative decision-making (SDG 16.7). It also aligns with the International Covenant on Civil and Political Rights (ICCPR), which also upholds the right to freedom of expression, access to information, right to privacy, and participation in public affairs. Malaysia is also a signatory to the Convention on the Elimination of All Forms of Discrimination against Women (CEDAW), which carries obligations to integrate gender perspectives into policymaking. Malaysia shall ensure equal representation in public consultations, as incorporating intersectional perspectives into the implementation of AI systems is a crucial and mandatory step. Embedding these standards in AI governance strengthens rights protection and demonstrates Malaysia's adherence to its international human rights commitments, positioning the country as a responsible digital leader.

Context and Framing: Malaysia's Path in AI Governance

Context

Malaysia's dedication to incorporating human rights principles within its AI strategy is commendable. Policy coherence and enforceable safeguards are crucial within these strategic steps. While the current national framework offers broad ethical principles, it urgently needs to provide operational assurances.

Under the AI NAP 2030, Malaysia boldly stakes its claim to be a regional powerhouse not just by adopting AI across sectors such as agriculture, health, education, and energy, but also by championing homegrown AI innovations. The plan emphasises building public trust, cracking down on cybersecurity threats, expanding 5G access, and cultivating a digitally savvy workforce—all under the ambitious Madani Economy Framework. However, to truly lead, Malaysia must revamp its regulatory landscape—demanding tougher algorithmic accountability, mandatory risk assessments, and more vigorous cross-sector enforcement. Civil society must be brought into the conversation, not just as an afterthought. If Malaysia fails to act decisively, it risks falling behind other nations vying for dominance in AI—losing its regional edge and democratic credibility in the process.

4 Priority Areas

Presently, the data protection framework under the Personal Data Protection Act (PDPA) of 2010 and its anticipated amendment in 2025 does not sufficiently ensure accountability across both the government and the private sectors. This implies that the law does not fully mandate either sector to adhere to specific data protection and responsible management standards. Furthermore, the Official Secrets Act (OSA) of 1972 does not adequately address the use of general government data or information for AI analysis and training, thereby creating a significant oversight and accountability gap. There is heavy reliance on soft law instruments such as the AI Governance Framework for the Enterprise (AIGE) and various guidance notes issued to the public sector. These guidelines are advisory and lack enforceability, failing to impose rigid requirements on organisations to disclose risk assessments or conduct independent audits before deploying AI systems. Consequently, practices that could compromise transparency and accountability may occur.

Currently, there exists no legal provision that grants individuals or the public an automatic right to access proactive disclosures. While 'open data' initiatives are in place, they tend to be selective and discretionary, with the decision to release information often resting with the data holders rather than being guaranteed, thereby creating a transparency gap that impairs public oversight and trust. Overall, these deficiencies highlight an urgent need for clearer, more comprehensive regulations and enforceable standards to promote responsible AI development and data management. Addressing these issues is vital for enhancing public trust, safeguarding privacy rights, and fostering accountable governance in the digital age.

Hence, this brief analysis identifies four priorities: improving transparency and accountability, aligning laws and institutions for consistency, recognising access to information as a core principle of AI governance, and fostering multistakeholder participation.

1. Improving Transparency and Accountability

Malaysia's push for AI adoption expands data collection across sectors, including welfare, health, education, policing, and platforms[2]. According to Karen Hao, a journalist and author of the book "Empire of AI," AI broadly encompasses machines capable of learning, reasoning, and autonomous action by identifying patterns within **trained datasets**. Consequently, safeguarding such data is essential to uphold individuals' right to privacy in accordance with data protection regulations. Without robust and intersectional safeguards in place, AI systems can entrench surveillance, reproduce bias and weaken democratic accountability[3]. Inadequate flows of information and opaque data processing

[2] RMK-13. "Rancangan Malaysia Ke-13." *Rancangan Malaysia Ke-13.* https://rmk13.ekonomi.gov.my/

[3] Feldstein, S. *The Global Expansion of AI Surveillance.* Vol. 17, No. 9. Washington, DC: Carnegie

carry the risk of data breaches, which the law must address directly.

The current National AIGE outlines transparency and human happiness as part of its seven principles, emphasising full disclosure and human involvement in AI systems used for specific high-risk processes. We argue that ministries and agencies should mandate proactive and mandatory disclosure of whether automated decisions were involved and the key factors driving them[4].

a) High Risk[5] Application of AI in Surveillance and Security

AI systems introduced under the banner of "national security", can be deployed in areas and apply to data shared beyond borders[6] through biometric systems, predictive tools, cyber incident powers, online safety regulation, and big-data policing[7] [8].

Still, they could lead to mass surveillance and pose risks of bias, discrimination, and compromised judicial systems due to biased datasets and inaccuracies[9] [10].

Endowment for International Peace, 2019.

[4] Wachter, S., Mittelstadt, B., and Russell, C. Counterfactual Explanations Without Opening the Black Box: Automated Decisions and the GDPR. *Harvard Journal of Law and Technology 31* (2017): 841.

[5] Taken from the EU Act definition: High-risk AI systems include systems used for biometric identification, biometric categorization,emotion recognition, assessing creditworthiness, managing critical infrastructure, assessing students,risk assessment and pricing in relation to life and health insurance, making employment-relateddecisions, certain activities related to law enforcement, migration controls, administration of justiceand elections and other systems listed in the law. In addition, an AI system that is a safety componentof a regulated product (e.g., a product subject to EU health and safety legislation), or that is itself aregulated product, will also qualify as "high-risk" (e.g., cars, toys, aviation). Act, E. A. I. (2024). The eu artificial intelligence act. *European Union*.

[6] The Edge Malaysia. "*National Integrated Immigration System Known as NIISe to be Operational in October, Says Immigration DG. The Edge Malaysia*. July 15, 2025. https://theedgemalaysia.com/node/762646

[7] Astro Awani. "KDN teroka teknologi AI beri perkhidmatan terbaik pada rakyat—Saifuddin Nasution." *Astro Awani,* January 22, 2025. https://www.astroawani.com/berita-malaysia/kdn-teroka-teknologi-ai-beri-perkhidmatan-terbaik-pada-rakyat-saifuddin-nasution-505632

[8] BERNAMA. KDN guna AI kawalan keselamatan di Sabah. *Berita Harian*, October 28, 2024. https://www.bharian.com.my/berita/nasional/2024/10/1317008/kdn-guna-ai-kawalan-keselamatan-di-sabah

[9] Domingo Jaramillo, C. Risks of Remote Biometric Identification Systems in Public Places for Law Enforcement Purposes under the AI Act Regulation. 2025.

[10] Davis, J., Purves, D., Gilbert, J., & Sturm, S. "Five Ethical Challenges Facing Data-Driven Policing." *AI and Ethics* 2, no. 1: 185-198.

The problem arises when these initiatives are brought forward under the guise of ‘innovation’ and ‘national security’ without ensuring adequate safeguards, such as amending data protection laws to hold authorities accountable in the event of any mishaps. This creates a real danger that AI will reinforce existing inequalities and erode fundamental constitutional rights, such as privacy, liberty, freedom of expression, and the right to a fair trial.

The Ministry of Home Affairs (MOHA) has already integrated testing for public security and airport auto gates. Firstly, this would pose a significant threat to communities at risk and human rights defenders[11]. Studies have also shown that racial bias might occur as an AI system misidentifies dark complexions and minority ethnic groups at much higher rates[12]. In addition, MOHA announced that it has harnessed AI to analyse 1.2 million criminal records, using machine learning to identify “getaway offences”, in other words, **predictive policing[13],** which would amplify historical and systemic bias and disproportionately target minority and marginalised groups**.**

b) AI sovereignty

Vendor lock-in is another long-term implication. Heavy investment in a single proprietary ecosystem makes it difficult for governments to switch providers or enforce stronger privacy standards. This dependency reduces bargaining power, increases costs over time, and risks subjecting Malaysia’s digital infrastructure to external corporate priorities rather than domestic governance objectives. As scholars of “data colonialism” argue, such arrangements extend global asymmetries, in which the Global South supplies data and labour while value and control remain concentrated elsewhere[14].

[11] BERNAMA. “JIM Mula Operasikan Sistem NIISe Guna Teknologi AI Oktober Ini.” *BERNAMA,* July 15, 2025. https://www.bernama.com/bm/news.php/jenayah_mahkamah/news.php?id=2445302

[12] Warso, Zuzanna. "Human Rights Requirements for Person-Based Predictive Policing: Lessons from Selected ECtHR Case Law and its Limits." *Technology and Regulation* (2022): 71-80.

[13] Loheswar, R. “Home Ministry Leans into AI to Predict Future Crimes, Target ‘Gateway Offences’.” *MalayMail*, July 8, 2025. https://www.malaymail.com/news/malaysia/2025/07/08/home-ministry-leans-into-ai-to-predict-future-crimes-target-gateway-offences/183231

[14] Arora, A., Barrett, M., Lee, E., Oborn, E., & Prince, K. “Risk and the Future of AI: Algorithmic Bias, Data Colonialism, and Marginalization.” *Information and Organization,* 33 no. 3 (2023): 100478.

Public-sector adoption of AI often involves procuring systems and services from large multinational technology companies[15]. While such partnerships can accelerate capacity building, they introduce structural risks to privacy, sovereignty, and accountability. The state bears responsibility for recognising and safeguarding its citizens' data as an integral component of individual privacy and autonomy rights. Consequently, the government is obligated to guarantee the enforcement of sovereignty in data governance, grounded in human rights principles. The Malaysian government must undertake initiatives to address legislative deficiencies in protecting cross-border data transfers. This necessity arises from the increasing prevalence of cross-border data flows, which involve the routine exchange and processing of data between nations. Failure to address this issue may compromise the nation's sovereignty. When governments rely on proprietary cloud and AI platforms, they risk placing sensitive citizen data under the jurisdiction of external entities, even when data centres are hosted locally. This creates uncertainty about compliance with domestic privacy laws and about the reach of foreign surveillance regimes.

The reliance on proprietary systems also raises concerns about transparency. Many vendor-provided AI tools operate as "black boxes," with opaque algorithms and agencies only able to access outputs. This limits regulators' and the public's ability to evaluate how decisions are made or to identify bias and error. In practice, the opacity of such systems undermines key principles of fairness and accountability, as individuals are unable to challenge decisions that affect their rights and entitlements. A recent example is the partnership between Google Cloud and the Ministry of Digital to roll out "AI at Work 2.0", which integrates Google Gemini (LLM) into civil servants' daily workflows. While framed as a step toward efficiency, the heavy reliance on a global 'big tech' company raises questions about transparency and data sovereignty, particularly when public administrations' decisions may rely on outputs from systems beyond independent scrutiny. Malaysia must refrain from deploying

[15] Free Malaysia Today. "Malaysia Rolls Out Generative AI Tool to 445,000 Civil Servants." *Free Malaysia Today,* February 5, 2025. https://www.freemalaysiatoday.com/category/nation/2025/02/05/malaysia-rolls-out-generative-ai-tool-to-445000-civil-servants

AI systems now and act later when integrating them, as they could affect the public directly.

c) PDPA and other data governance

Research institutes, including the Khazanah Research Institute (KRI) and the Institute of Strategic & International Studies (ISIS), have also published papers arguing that Malaysia's AI governance lacks comprehensive, enforceable legal frameworks. The papers also highlight that existing laws like the Personal Data Protection Act 2010 (PDPA, amended in 2025) and Cyber Security Act 2024 address only specific aspects, such as data breaches, mandatory notification and critical infrastructure, leaving broader AI risks such as AI bias, accountability and transparency, unregulated[16] [17].

The PDPA applies primarily to private sector actors and commercial transactions, leaving the public sector exempt. This means many high-risk deployments, such as welfare automation, predictive policing, and biometric ID, lack the same level of oversight. The act at this point does not adequately address modern risks, which are central to AI.

The Public Sector AI Adoption Guidelines (2025) and supporting instruments such as KRISA, the Public Sector Data Sharing Policy (2021), the Big Data Analytics Circular (2017), the Open Data Circular (2017), The Data Sharing Act 2025 and the Data Dictionary Standard (2022) prove useful for harmonisation and data management, but it is treated as a means for efficiency and not a tool for privacy safeguards implementations.

2. Fragmented Oversight and AI Accountability

a) Principles and Objectives

As we move towards an AI Act in the near future[18], it is crucial to address the issue of fragmentation in Malaysia's regulatory landscape. Whilst we note the establishment of NAIO and AIGE as steps forward, these remain

[16] Khazanah Research Institute, *AI Governance in Malaysia.* (n.d.).

[17] Said, F., & Nabilah, F. *Future of Malaysia's AI Governance.* 2024.

[18] The Malaysian Reserve. "Malaysia Targets AI Law by 2026 as Govt Weighs Risks and Rules." *The Malaysian Reserve,* July 3, 2025. https://themalaysianreserve.com/2025/07/03/malaysia-targets-ai-law-by-2026-as-govt-weighs-risks-and-rules/

non-legally binding and lack statutory backing and enforceable risk assessment protocols for AI deployment. Current reliance on voluntary ethical guidelines (similar to the EU's "Trustworthy AI" Approach risks creating an accountability vacuum: broad principles are stated, but operational mechanisms for oversight, auditing and rights-based safeguards are absent[19].

While we understand the need for 'soft laws' as a quick solution to fill policy gaps, they come with significant drawbacks. Firstly, the unenforceability of these soft laws and confusion and overlap in AI governance[20]. Without binding requirements and an oversight body to monitor and evaluate, high-risk AI policing, welfare, health, military, and surveillance can be deployed without scrutiny of impacts on rights, equality, or due process[21] [22].

Voluntary ethics underperform, as mentioned. A comprehensive package that includes impact assessments, audits, explainability, and disclosure should be implemented. It enhances accountability, while standards (ISO/NIST) make it operational across agencies and sectors. Using horizontal and vertical regulation, joint supervision, and the Right to Information could be a way for Malaysia to plug gaps quickly, without a complete overhaul.

b) Mechanisms and Processes

As mentioned above, core instruments such as the PDPA and CSA focus on data and infrastructure security but do not address **algorithmic accountability, bias mitigation, or the integrated human rights impact of autonomous systems.** We believe that the best way to regulate is through mandatory AI audits, impact assessments, or transparency reports that ensure alignment with NAIGE's principles and other countries' aspirations.

Without mandatory AI audits and standardised risk assessment tools, Malaysia

[19] Mittelstadt, B. "Principles Alone Cannot Guarantee Ethical AI." *Nature Machine Intelligence* 1, no. 11 (2019): 501-507.

[20] Marchant, G. "'Soft Law' Governance Of Artificial Intelligence." 2019.

[21] Binns, R. "Algorithmic Accountability and Public Reason." *Philosophy & Technology*, 31, no. 4 (2018): 543-556.

[22] Eubanks, V. *Automating Inequality: How High-Tech Tools Profile, Police, and Punish the Poor.* St. Martin's Press, 2018.

risks falling into **regulatory non-interoperability** with global AI frameworks such as the EU AI Act, OECD AI principles, and ISO/IEC 420001, which are the standards for AI governance at this point. Misalignment could hamper Malaysia's participation in cross-border data initiatives or collaborative research with international partners, including major investors such as Google and Microsoft, who already operate under these frameworks. As the OECD (2023) notes, interoperability among AI risk management frameworks enhances efficiency, reduces enforcement and compliance costs and prevents administrative burdens that hinder cross-border cooperation[23].

c) Implementation Pathways

Malaysia's AI governance remains hampered by fragmented oversight and unclear pathways. Oversight is currently scattered across multiple ministries and agencies, which are working on ways of implementing AI systems for public sector use. For instance, the Malaysia Digital Economy Corporation (MDEC) and the Malaysian Communications and Multimedia Commission (MCMC) are advancing their own mandates and initiatives, often without coordination. This siloed approach weakens overall policy coherence, leads to regulatory overlap, and diminishes the effectiveness of law enforcement efforts.

At the point of writing, a crucial aspect that we must point out is the fragmented oversight and the absence of clear pathways, particularly in areas such as policy formulation, regulatory enforcement, and risk management in AI systems. While there has been a move toward centralization with the establishment of the National Artificial Intelligence Office (NAIO), fragmentation remains a significant issue, particularly in high-risk applications like surveillance and biometric systems, where accountability is paramount.

3. Treating access to information as a key principle of AI governance

Access to information should be treated as a cornerstone of AI governance

[23] OECD. "Common Guideposts to Promote Interoprability in AI Risk Management." *OECD,* November 2023. https://www.oecd.org/content/dam/oecd/en/publications/reports/2023/11/common-guideposts-to-promote-interoperability-in-ai-risk-management_9629ed36/ba60time of writing, a crucial aspect we must point out is the fragmented oversight and the absence of clear pathways, particularly in policy formulation, regulatory enforcement, and risk management forcentralisation2d18-en.pdf#page=15.10

because it underpins democratic oversight and public trust. When governments and companies deploy AI systems that directly impact public life, be it in policing, welfare distribution, or content moderation, the public must have the right and means to know how these systems function, how their data is processed, and the common risks they pose.

Access to information empowers all levels of society to participate meaningfully in the implementation of AI systems, ensuring accountability beyond institutional processes. This approach reflects Malaysia’s current aspirations to establish a right-to-information regime and SDG 16.10, which calls for public access to information as a condition of effective governance.

4. Embedding Multistakeholder Participation

The RMK13 outlined an inclusive and sustainable AI nation, and realising this requires a collective ‘Whole-of-Nation’ strategy. Similarly, the National AI Roadmap (2021–2025) calls for AI to serve as a catalyst for both economic competitiveness and improved governance outcomes. These policy frameworks highlight that technological progress cannot be achieved in isolation by the state or industry alone; instead, it requires the collective action of government, academia, industry, civil society, and potentially representatives of impacted groups (e.g., indigenous peoples, gender minority groups, and child protection groups).

Adopting a whole-of-nation strategy ensures that AI systems reflect Malaysia’s diverse societal needs and protect fundamental rights. CSOs are uniquely positioned to contribute by monitoring risks, facilitating community engagement, and promoting accountability in both the public and private sectors for AI deployment. Their involvement can help bridge gaps between technical expertise, policymaking, and lived experiences on the ground.

Without such inclusive collaboration, AI, when integrated into government systems and processes, can pose the risk of reinforcing structural inequalities, eroding trust in public institutions, and enabling harmful applications such as unchecked surveillance or discriminatory profiling. By contrast, embedding multistakeholder participation into AI governance strengthens Malaysia’s ability to meet RMK13’s goals of inclusive growth and sustainable development while ensuring alignment with international human rights standards. The Whole-

Nation Approach is best for engaging all key stakeholders from the outset of the AI product life cycle.

Recommendations

1. Fragmented Oversight and AI Accountability

a. Improving AI Transparency and Accountability (Modality)

Actor	Role	Alignment
Multistakeholder Working Group (MSWG)	Co-design standards (Human rights, trials via sandboxes.	Recommendations would feed into the Independent Oversight Body
Independent Oversight Body	Approve, restrict, or ban high-risk AI systems, and conduct AI human rights auditing.	Carry out recommendations by the MSWG, and provide an appeals and grievance mechanism
High-risk AI Registry	Monitor and evaluate high-risk AI systems.	Provide data for oversight reviews and for the MSWG to develop further recommendations.

b. Establishing a Multistakeholder Working Group (MSWG) 'Human-in-the-loop' Governance

The creation of an MSWG, which comprises government, academia, civil society, and industry players in Malaysia's AI governance, aligned also with Malaysia's AIGE principles and anchored to RMK13's development elements of "Building Policies" and "Increasing Digital Trust". The MSWG will co-draft guidance, test it in sandboxes, and route recommendations into RMK13's monitoring and GovTech reforms. There are certain areas of involvement from CSOs, such as:

- **Rights-respecting input**: Provide input on what constitutes the high-risk tiers of AI system deployment, human-review thresholds, escalation procedures, and records, rooted in international human rights standards.
- **Co-development of Human Rights Impact Assessments**: Such a framework aims to clarify the responsibilities of all parties across the AI system life cycle, as well as the

safeguards and measures they must implement.

- **Gathering input from all levels of society**: Impacted groups and marginalised communities, such as indigenous people, women, children, refugees, and LGBTQIA groups, are exposed to intended and unintended risk from AI systems. Inputs from these groups will help formulate adequate safeguards to protect them.

c. Establishing an Independent AI Safety Oversight Body

Maintaining the spirit of multistakeholderism, the establishment of an Independent AI oversight body will safeguard rights and uphold accountability in the use of AI systems. Its role would be to review and monitor high-risk systems, set enforceable standards, and provide a fair redress mechanism for people harmed by automated decisions and products. The oversight body may operate regulatory sandboxes limited to high-risk AI systems used by government agencies[24].

This makes AI use transparent, protects fundamental rights such as privacy, equality, and access to redress, and prevents conflicts of interest when government agencies eventually deploy and regulate their AI tools.

The oversight body would also work in synergy with the Multistakeholder Working Group, which brings government, industry, CSOs and academia together to co-design policies and test them in practice (sandboxing).

a) MSWG = develop standards rooted in rights-based principles, HITL safeguards.
b) Independent oversight body = enforcement and monitoring.
c) AI Registry = data, monitoring and evaluation.

[24] These sandboxes allow controlled testing under strict safeguards, independent monitoring, and clear exit criteria, ensuring that high-risk tools meet legal and human-rights standards before deployment.

This balance of participation, independence, and transparency strengthens human rights protections while also supporting Malaysia's role in the region as set out at ADGMIN 2025. It would link domestic policies to ASEAN's initiatives on AI governance, security and trust.

d. Establishing a high-risk AI Registry

The high-risk AI registry is a central system for recording, tracking, and disclosing information about government-owned AI systems. The AI registry should serve as an extension to the oversight body. Each entry would include the system's purpose, developer, risk level, and the safeguards in place, especially whether people can contest or override automated decisions.

e. Establishing Referral/Redressal Tracker

The importance of formulating a remedial procedural mechanism in AI governance resides in recognising that AI-enabled risks and harms can impact human rights at both micro and macro levels. The lack of a redress mechanism shows how states and the private sector often overlook the perspectives and needs of current and potential users, as well as the need for redress for human rights violations or breaches related to their activities, in both the design and operation of these systems. States have a vital role in ensuring and safeguarding human rights. Nevertheless, technology companies' execution of their corporate responsibility to respect human rights can also significantly affect how well the remedy system functions in practice.

Simultaneously, technology companies can enhance aspects of the state-based regulatory framework through self-regulation for direct redress as a key component of the smart mix approach. AI companies, including both developers and deployers, should establish and implement effective operational-level grievance mechanisms, as discussing effective redress logically follows from recognising the diverse risks that may occur across the AI lifecycle. This notion is enshrined not only in international customary law and the UN Guiding Principles on Human Rights but also in the mandate of the Consumer Protection Act 1999.

There are numerous opportunities within the realm of consumer protection law. Firstly, implementing a presumption of liability throughout the AI lifecycle would offer alternatives and promote responsible AI development. Secondly, adopting an evidentiary model that reallocates the burden of proof to business actors would address the challenge of opacity, which frequently constitutes a substantial obstacle for victims in providing evidence. Thirdly, establishing a redressal mechanism to address communal harm would be pertinent in light of the concept of relational autonomy.

f. Adopting a Hybrid Model of Governance

a) Horizontal Regulation: A Common Floor of Protections

Horizontal regulation establishes baseline safeguards across all AI systems, regardless of sector. This includes uniform definitions of "high-risk AI," mandatory human rights and algorithmic impact assessments, and obligations for explainability and independent audits. Such cross-sector requirements respond to calls for algorithmic accountability frameworks that apply system-wide rather than piecemeal[25]. Importantly, horizontal regulation should be anchored in the right to information (Article 19, UDHR; ARTICLE 19, 2024), ensuring that citizens, civil society, and journalists have enforceable access to information about how AI is procured and deployed.

[25] Veale, M., & Edwards, L. "Clarity, Surprises, and Further Questions in the Article 29 Working Party Draft Guidance on Automated Decision-Making and Profiling." *Computer Law & Security Review* 34, no. 2 (2018): 398-404.

b) Vertical Regulation: Sector-Specific Safeguards

While a shared floor is essential, AI's risks are context-specific. Vertical regulation tailors safeguards to the unique risks of sectors such as policing, healthcare, education, welfare, finance, and online platforms. In policing, predictive algorithms may exacerbate discrimination, warranting moratoriums or strict thresholds[26]. In welfare, automation may wrongly deny benefits, requiring mandatory human review[27]. On online platforms, recommender systems amplify harmful speech, necessitating systemic risk assessments[28]. Thus, vertical rules ensure proportionality: while not all AI poses a high risk, stronger safeguards are necessary when human rights are directly implicated.

c) Horizontal and Vertical Together

A hybrid governance model works in tandem to fill gaps. A purely horizontal approach risks remaining abstract principles without enforcement. A purely vertical approach risks sectoral silos with no common accountability language. Combined, they create a coherent architecture: a shared baseline of rights and accountability (horizontal) reinforced by sectoral rules (vertical). This hybrid approach aligns with emerging global best practices, including the EU AI Act's layered risk framework and the OECD's principles on AI (OECD, 2019).

[26] Ferguson, A. G. *The Rise of Big Data Policing: Surveillance, Race, and the Future of Law Enforcement*. New York: New York University Press, 2017.

[27] Eubanks, V. *Automating Inequality: How High-Tech Tools Profile, Police, and Punish the Poor.* St. Martin's Press, 2018.

[28] Gorwa, R. "The Platform Governance Triangle: Conceptualising the Informal Regulation of Online Content." *Internet Policy Review* 8, no. 2 (2019): 1-22.

2. Privacy and Security

a. Address Regulatory Gaps around Data Governance

- Expand the scope of the PDPA 2010 (amended in 2025) to include federal and state governments, ensuring that all processing of personal data by government bodies is subject to statutory duties of accountability and limitations.
- Harmonise the PDPA, the Cyber Security Act 2024, the Data Sharing Act 2025, and sectoral laws with the National AIGE, which serves as a baseline for responsible AI.
- Implementation of Data Protection Impact Assessment (DPIA). Particularly for high-risk AI systems.
- Related governmental agencies, such as NAIO (policy/AI governance), JPDP (privacy), NACSA (cybersecurity), and ministries (data owners) must work together to ensure AI systems that are put in place go through scrutiny and wire reporting under an Independent oversight body, as mentioned above.

b. Mandatory Self-assessments for Companies

This requires an annual AI & Data Governance Self-Assessment for suppliers to the government using AI and firms operating high-risk AI (safety-critical, finance, health, and rights-impact). File with the relevant regulator and keep it on site for the audit.

Base the template on AIGE, ISO/IEC 42001[29] (AI management systems), and the NIST AI RMF (incl. the GenAI profile)[30]. Include DPIA, documented model lineage, security controls, testing, and an incident playbook. Reporting can be wired to the AI registry.

c. Clear and Narrow Definitions for use of AI in National Security

[29] ISO. *ISO/IEC 42001:2023*. (n.d.). Retrieved August 25, 2025. https://www.iso.org/standard/42001

[30] NIST. "AI Risk Management Framework." *NIST,* 2021. https://www.nist.gov/itl/ai-risk-management-framework

Establish strict, clear and narrow definitions for National Security so that it does not overlap with law enforcement activities (prevention, detection, investigation, and prosecution of criminal offences)[31]. To avoid abuse, support accountability and protect rights.

3. Treating access to information as a key principle of AI governance
   a. **Principle of Maximum and Proactive Disclosure**
   An AI governance instrument should be grounded on the principle of maximum and proactive disclosure, meaning that all information held by public bodies is public and must be automatically disclosed.

A Right to Information Memorandum prepared by the Centre for Independent Journalism, Article 19, and the Sinar Project stresses that the right to information is recognised under Article 19 of the Universal Declaration of Human Rights (UDHR) and the International Covenant on Civil and Political Rights (ICCPR). It further notes that the spirit of RTI goes beyond limiting government secrecy but also requires proactive disclosure of information that affects people's livelihoods and democratic accountability[32]. To ensure the government's commitment to building digital trust through mandatory disclosure of AI use by ministries, agencies, and contractors, existing laws and rules that obstruct the development of a practical RTI framework must be reviewed, repealed, or amended.

The proactive disclosure obligation requires states to publish information of public interest proactively. When it comes to AI, data disclosure must be a priority for public-interest AI use. Malaysia's existing data-sharing and open data guidelines, outlined in the Public Sector AI Adaptation Guidelines, emphasise efficiency and interoperability but fall short of recognising people's right to know how AI systems are designed and deployed.

[31] Montasari, R. (2024). *National security in the Artificial Intelligence era: Challenges and implications of advanced technologies* (Doctoral dissertation, Cardiff Metropolitan University).

[32] Article 19. "Malaysia: Uphold International Standards in Tablings of Right to Information Bill." *Article 19*, (n.d.). https://www.article19.org/resources/malaysia-uphold-international-standards-in-tabling-right-to-information-bill/

4. Embedding Multistakeholder Participation

   Promoting a comprehensive national strategy ensures that AI systems align with Malaysia's diverse social needs while safeguarding fundamental rights. Civil society organisations (CSOs) possess a unique position to evaluate risks, promote community involvement, and improve accountability across government and private-sector AI initiatives. Their involvement connects technological expertise, policy development, and practical experiences. Incorporating multistakeholder participation in AI governance bolsters Malaysia's ability to achieve the RMK-13 objectives of inclusive growth and sustainable development, while aligning with international human rights standards. The Whole-Nation Approach represents the most effective method for engaging all principal stakeholders as a shared effort to shape how AI develops in Malaysia, and that decisions reflect lived realities on the ground, not only technical or commercial views.

Annex

The section contains meeting notes from the Multi-stakeholder Discussion on the Development of Public Interest AI Governance in Southeast Asia held on August 28, 2025 at the AICB Centre of Excellence, 10, Jalan Dato Onn, Kuala Lumpur, 50480 Wilayah Persekutuan, Wilayah Persekutuan Kuala Lumpur, Malaysia. This session is a side event of the Digital Rights in Asia-Pacific assembly (DRAPAC25), a two-day regional gathering that convenes diverse actors across the Asia-Pacific to shape rights-based digital governance and foster solidarity within the digital rights community.

Center for Independent Journalism (CIJ), Wikimedia Foundation, Oxfam, and EngageMedia are among the civil society organisations that have been working on rights-based governance on AI and other emerging tech in Southeast Asia. We are co-conveners in providing expert feedback and organising a dialogue that brings together domestic and international expertise. NAIO, Indonesia National AI Task Force, Khazanah Research Institute, International Center for Not-for-Profit Law, Consumer Unity & Trust Society (CUTS), United Nations Development Programme (UNDP) Malaysia, the Office of the United Nations High Commissioner for Human Rights (OHCHR) and GIZ Asia, among others, are attending it.

In this meeting, civil society groups in Malaysia emphasised that, in efforts to regulate AI governance, the government should consider six key aspects: public participation, accountability, transparency, data protection and privacy, oversight, censorship and security, and legal reform.

Several key points discussed during the meeting, which led to a shared understanding, included:

- The state holds significant power due to its role in procurement, implementation, and regulation.
- Ensuring robust data protection is crucial, as AI is closely tied to issues of quality, discrimination, use, privacy, surveillance, and extensive data collection. This includes concerns about unethical business practices, violations of consumer rights, and discrimination in automated content distribution.