AI Governance in Action:

Malaysia's National Guidelines on AI Governance and Ethics (AIGE)

Background

Malaysia's journey towards AI governance has gained significant momentum with the launch of the National Digital Economy Blueprint ("MyDigital") in 2021, which recognises AI as a core enabler of the digital economy. As AI adoption expands rapidly across various industries, growing concerns over its effects on employment, privacy, and human rights underscore the need for a robust governance framework.

In August 2022, the Ministry of Science, Technology, and Innovation ("MOSTI") unveiled the National Artificial Intelligence Roadmap 2021-2025 ("AI-RMAP"), which defines AI as a suite of technologies aimed at enhancing human ingenuity and addressing complex challenges. The roadmap also lays the foundation for the development of the National Guidelines for AI Governance and Ethics ("AIGE"), designed to ensure responsible, safe, and ethical integration of AI, while promoting transparency and trustworthiness in its application.

Overview of the AIGE

The first instalment of the AIGE, introduced in September 2024, outlines seven (7) key principles adapted from standards issued by the United Nations Educational, Scientific and Cultural Organisation ("UNESCO") and the Organisation for Economic Co-operation and Development ("OECD"), advocating for the responsible use of AI to cultivate a dynamic national innovation ecosystem. While not legally binding, the AIGE provides organisations with the flexibility to tailor and adopt its principles based on their specific sectors and use cases. Embracing these principles is a responsible course of action for organisations to align with ethical standards, in preparation for future regulatory frameworks.

Pursuit of human benefit and happiness

AI systems should improve quality of life and well-being, avoiding misuse for harm, with human oversight to protect autonomy and user experience.

Fairness

AI systems should be free from bias and discrimination, ensuring inclusivity and equal access for all.

Accountability

Developers, owners, and users must remain accountable for AI's outcomes, ensuring systems comply with ethical standards and governing policies, while considering the system's purpose, technology capability, and user sensitivity.

Reliability, safety, and control

AI should be reliable, secure from unauthorised manipulation, and enable human oversight, particularly in high-risk areas such as healthcare, autonomous vehicles, and military applications.

Transparency

AI processes and algorithms should be transparent and comprehensible, enabling stakeholders to assess risks and decisions. Clear transparency in data sources, training methods, and decision-making is crucial.

Privacy and security

AI systems should integrate security-by-design and privacy-by-design to protect personal data. Informed consent is emphasised, ensuring user data is not mishandled or unlawfully shared.

Principle

Inclusiveness

AI technology should cater to diverse societal needs, including marginalised communities, enabling them to benefit from AI advancements.
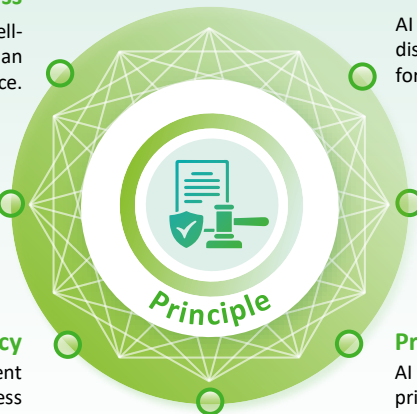
Impacted stakeholders of AIGE

The AIGE is tailored for the following three main groups of stakeholders:

End-users (Individuals and organisations)

Individuals and organisations using AI-powered applications, including those unaware that they are engaging with AI. Guidelines for end users stress the importance of knowing their rights and responsibilities, such as understanding how AI interacts with their data, safeguarding privacy, and empowering informed decisions.

Policymakers in government, agencies, organisations, and institutions

Officials, regulators, and agencies responsible for AI governance. The guidelines provide a framework to establish and monitor ethical AI practices, ensure consumer protection, and promote fair competition. Additionally, policy makers are encouraged to form advisory bodies to oversee AI deployments and align them with Sustainable Development Goals ("SDG").

Developers, designers, technology providers and suppliers

AI developers, designers, and technology suppliers are urged to adhere to ethical standards and conduct rigorous risk and impact assessments. The guidelines encourage compliance with established technical standards and include recommendations for ethical design, transparent data management, and user protection.

The AIGE brings significant impact to organisations across four key dimensions: People, Process, Policy, and Technology. These changes necessitate a comprehensive re-evaluation and adaptation of current practices

Policy

- **Development of AI governance framework:** Organisations should create governance frameworks based on the seven principles, ensuring ethical AI use with accountability mechanisms.

- **Update data privacy policies:** Privacy policies must be updated to secure explicit consent for the use of personal data for AI training purposes and protect users' rights.

- **Consumer protection measures:** Policies should protect consumers' rights in AI interactions, ensuring transparency, explanation of AI decisions, and avenues for redress.

Process

- **AI lifecycle management:** Organisations should integrate responsible AI practices across the entire AI lifecycle, from design to deployment and monitoring.

- **Risk assessment and mitigation:** Compliance requires organisations to implement strong risk assessments, focusing on ethical risks and safeguards against data misuse.

- **Data trust:** Organisations should regularly review data practices, securing consent, minimising usage, and ensuring compliance with privacy regulations to uphold data privacy and security.

- **Continuous monitoring and auditing:** Organisations should establish or enhance processes for monitoring and auditing AI systems, including impact assessments, bias audits, and performance checks to ensure ongoing adherence to ethics principles.

People

- **Training and awareness:** All employees should be trained on AI ethics, privacy, and responsible use, fostering a culture that emphasises fairness, accountability, and ethical implications.

- **New roles and responsibilities:** Organisations may create AI ethics roles and a governance committee to ensure compliance with regulations and ethical AI development/ usage.

- **Diversity and inclusion:** Diverse perspectives in AI development/ usage should be included to avoid biases, especially among data scientists, AI developers, and decision-makers.
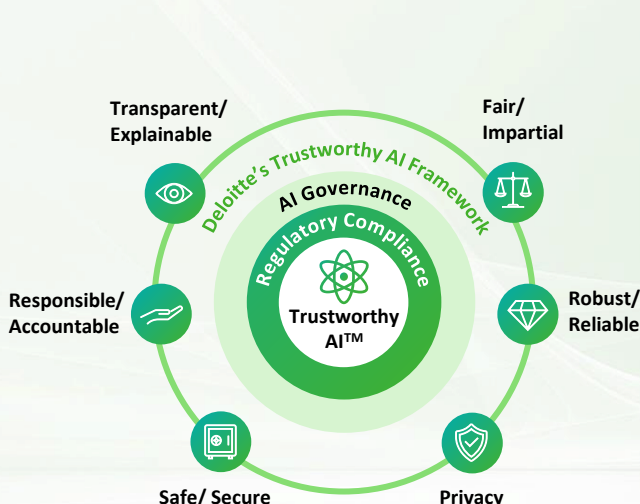
Technology

- **AI design and development:** Technical teams must incorporate ethical considerations in AI design, including fairness, transparency, and accountability.

- **Explainability and transparency tools:** Organisations may need to invest in developing dependable fairness metrics and explainable AI (XAI) tools to help users and developers and users understand AI decision-making.

- **Security, robustness and privacy by design:** AI systems should be built with security, robustness and privacy considerations from the start, including secure coding, data encryption, and strict access controls.

- **Human oversight mechanisms:** Organisations should incorporate mechanisms such as human-in-the-loop ("HITL"), human-on-the-loop ("HOTL"), or human-in-command ("HIC") models, to enable human oversight and intervention in automated processes.

Our transformation programme

We provide services to assist organisations navigate the complexities of AI governance, promoting responsible and ethical deployment at every stage. Above all, we can assist organisations to effectively manage risks, ensure regulatory compliance, and uphold principles of Trustworthy AI throughout the entire AI lifecycle.

Deloitte's Trustworthy AI Framework — AI Governance — Regulatory Compliance — Trustworthy AI™

- Transparent/ Explainable
- Fair/ Impartial
- Responsible/ Accountable
- Robust/ Reliable
- Safe/ Secure
- Privacy

Assess

- Evaluate the organisation's AI governance framework, programme, applications, and algorithms against local regulations, industry standards, and selected maturity model.
- Provide recommendations to enhance overall maturity, compliance, and alignment with Trustworthy AI principles.

Implement

- Establish a strategic direction and governance framework to ensure effective oversight of AI risks.
- Enhance the organisation's AI governance framework by developing relevant policies and implementing appropriate technology solutions.
- Support the organisation in preparing for the selected certifications (e.g. ISO/IEC 42001).

Operationalise

- Advise on operationalising the organisation's AI governance programme and framework, including training and awareness initiatives.
- Provide continuous support to ensure optimisation and ongoing compliance with local regulations and industry standards.

Ho Siew Kei
Malaysia Cyber Risk Leader
sieho@deloitte.com

Venkat Paruchuri
**SEA Data Privacy and
Protection Risk Leader**
veparuchuri@deloitte.com

Melvin Toh
**Senior Manager, Technology &
Transformation**
Deloitte Southeast Asia
mtoh@deloitte.com

Melbourne Lim
**Manager, Technology &
Transformation**
Deloitte Southeast Asia
melblim@deloitte.com

Deloitte.