



AI Ethics Policy Paper

***Advancing AI Ethical Policies &
Navigating the Ethical Landscape***

April 2024

CONTENTS

A. EXECUTIVE SUMMARY	3
B. PREFACE.....	5
C. READERSHIP.....	8
D. THIS PAPER.....	9
E. APPENDIX.....	15
1. PIKOM 2024 Survey results	15
2. Existing Laws and Policies	19
3. References.....	19

While every effort has been made to ensure the accuracy of the information, all information furnished in this publication is provided strictly on an ‘as is’ and ‘as available’ basis and is so provided for your information and reference only. As such, PIKOM including their partners and associates, whether named or unnamed, do not warrant the accuracy or adequacy of the findings. Moreover, all parties concerned explicitly disclaim any liability for errors or omissions or inaccuracies pertaining to the contents of this publication. Therefore, the use of the findings presented in this publication is solely at the user’s risk. PIKOM shall in no event be liable for damages, loss or expense including without limitation, direct, incidental, special, or consequential damage or economic loss arising from or in connection with the findings in this publication.

A. EXECUTIVE SUMMARY

Introduction

The paper presents a guardrail framework for ethical AI development and usage, emphasizing risk and governance.

Aimed at stakeholders in the local IT technology industry, it promotes responsible AI deployment by incorporating ethical principles and requirements such as informed consent, continuous monitoring, collaboration, and environmental considerations.

Key Policies

Each policy is supported by a clear rationale, detailing the importance of addressing these specific risks and promoting ethical behavior in AI development and usage. The seven key guiding policies proposed are as follows.

- Fairness
- Reliability, Safety, and Control
- Privacy & Security
- Inclusiveness
- Accountability
- Transparency
- Pursuit of Human Benefits

Related Risks

Additional related risks including reputational risk, and legal risk, are highlighted, underlining the impact and consequences of not adhering to the proposed policies.

Characteristics Embedded in Risk and Governance

Seven characteristics are also advocated for inclusion in organizational enterprise risk and governance frameworks. These are:

- Informed Consent
- Continuous Monitoring
- Collaboration and Openness
- Environmental Impact

- Human Oversight
- Education and Awareness
- Checks and Balances

Collectively they are also known as **PIKOM “7 by 7” Policy and Characteristic (P&C)** framework.

Conclusion

The paper concludes by emphasizing the necessity of periodic review of policies and guidelines to ensure relevance and practicality, ultimately aiming to foster a culture of ethical AI development and usage for a more inclusive, transparent, and trustworthy AI ecosystem.

KEY MESSAGES

1. PIKOM is propagating industry self-regulation with practical and realistically-defined guard rails and policies.
2. PIKOM is proposing independent third-party review to ensure these policies are adhered to.
3. These third parties can be standards or certified bodies/organisations (similar to financial auditors)
4. These policies are to be reviewed and revised periodically to align with the industry and AI technology development.

B. PREFACE

The rapid advancement of AI technology recently has been revolutionary, offering unprecedented opportunities for innovation and efficiency to business, industry, society, and government. However, this rapid progress has also raised significant ethical concerns regarding the development and deployment of AI systems. As AI becomes increasingly integrated into all aspects of life, the need for comprehensive ethical policies has become paramount to ensure its responsible and ethical use.

Background

This report highlights the key perspectives and recommendations regarding ethical guidelines for AI development and deployment. It identifies some of the foundational principles and values for guidance, explores the global landscape of AI ethics, examines the primary ethical dilemmas and associated risks, and provides policy recommendations and implementation strategies to address these challenges.

This paper is not meant to respond to all the issues and concerns of this emerging technology but aims to propose a set of guiding principles for the industry to follow and adopt, considering Malaysia is at the early stages of adoption and development of AI including Generative AI. These policies must also be reviewed periodically to ensure relevance and practicality.

Understanding AI Ethics

Ethical development and deployment of AI systems should be guided by fundamental principles such as transparency, fairness, accountability, inclusivity, and privacy. These principles ensure that AI technologies are designed and used in a manner that respects human rights, promotes societal well-being, and mitigates potential harms.

In all cases of technology deployment over the decade, ethics have been just as paramount as their benefits. However, AI technology poses additional considerations. The pace of its evolution is so fast and the amplification of productivity gains, or ROI is

so phenomenal that ethics can sometimes be sidetracked intentionally or unintentionally.

Current Landscape of AI Ethics

Globally, various frameworks and initiatives have been established to address ethical considerations in AI. These include guidelines from organizations such as the IEEE, OECD, and the EU, as well as initiatives like the Partnership on AI and the Montreal Declaration for Responsible AI. Malaysia has also published an AI Roadmap (AI-RMap) 2021-2025 that will be referenced here. These frameworks aim to promote ethical AI development and deployment by providing principles, guidelines, and best practices for stakeholders.

Challenges

Despite the potential benefits of AI, there are significant ethical dilemmas and risks associated with its development and deployment. These include issues related to bias and discrimination in AI algorithms, lack of transparency and accountability in AI decision-making processes, threats to privacy and data protection, and the potential for AI systems to exacerbate existing societal inequalities. In most cases, AI-driven automation enables humans to do more and faster, resulting in the amplification and perpetuation of the degree of harm. Further, AI can lead to the emergence of new sources of harm.

Policy Recommendations

An ethical approach to AI and risk management practices can help to prevent or mitigate these risks. To ensure the ethical development and deployment of AI, we should consider implementing effective but practical risk-based policy frameworks. They include establishing clear guidelines and standards for AI development, ensuring transparency and accountability in AI systems, promoting diversity and inclusivity in AI research and development, and fostering collaboration to address ethical challenges. Our guiding policies must also be aligned with the current state of development and deployment of

AI in Malaysia. In our adoption of best practices, we must also ensure they are aligned with our industry.

With such frameworks, PIKOM is of the opinion that the industry should practice self-regulation on AI ethics with guiding principles based on risk and governance. Further, we propose that AI development and deployment be subject to third-party review either by the authorities or a standards body. Over-regulation through mandates and bans should be avoided as this may have the unintended effect of curbing AI based innovation.

Implementation Strategies

Effective implementation of ethical AI policies requires collaboration and coordination among various stakeholders, including governments, industry, academia, civil society, and international organizations. Stakeholders should work together to develop a risk-based governance model that includes enforcement of ethical AI standards as a baseline, promote responsible AI research and innovation, raise awareness about AI ethics, and provide education and training on ethical AI principles and practices. The policies will also require effective implementation strategies including a consistent set of measurement metrics to be in place to monitor the outcomes that drive continuous improvement.

In conclusion, addressing the ethical challenges posed by AI requires a concerted effort from all stakeholders to develop and implement holistic ethical policies that prioritize human well-being and societal values. By working together, we can harness the transformative potential of AI while mitigating its risks and ensuring a more ethical and responsible AI future. Lastly, the industry should develop the necessary mechanisms for self-regulation based on risk and governance, and supported by an external review process.

C. READERSHIP

These guidelines are intended for a broad audience involved in various aspects of AI development and deployment within the industry. The target readership includes:

- AI Developers and Engineers
- Organizational Leaders and Decision-Makers
- Data Scientists and Analysts
- Regulatory and Compliance Officers
- Policy Makers and Government Officials
- End-Users and Consumers
- Ethics and Compliance Professionals
- Educators and Researchers

By targeting this diverse readership, the guidelines aim to foster a collective commitment to responsible AI development and usage, ensuring that ethical principles are ingrained across different roles and levels within the industry.

Organizations are encouraged to define their set ethical principles, and then apply a risk management approach (for example ISO 42001 - AI Management System) to drive the governance.

D. THIS PAPER

Introduction

As prefaced in earlier sections, AI continues to permeate various aspects of society, requiring its development and usage to be aligned with ethical principles. This paper outlines a guardrail framework for ethical AI development and usage, with a focus on risk and governance. Bearing in mind that our local adoption and development of AI are at a formative stage, our guiding principles must also reflect this although regulating this factor is certainly not the intention of these recommendations.

Henceforth, the intent and purpose of this proposed framework is to provide overall guidance to stakeholders that develop or use AI including Generative AI, to navigate the complex landscape of AI ethics. By pivoting on complying with pre-determined ethical principles and additional requirements such as informed consent, continuous monitoring, collaboration, and environmental considerations, this framework aims to promote responsible AI deployment through mitigating potential risks and ethical lapses.

As iterated earlier, it is only through collaboration, education, and accountability that this framework seeks to foster a culture of ethical AI development and usage, ultimately contributing to a more inclusive, transparent, and trustworthy AI ecosystem.

Importantly these policies and guidelines must be reviewed periodically to ensure relevance and practicality.

Based on our research and collective input from the industry, PIKOM has proposed 7 Policies for responsible AI adoption in the use and development of AI technology including Generative AI.

These core policies become the key priorities for any risk-based governance. In addition, organizations may embed additional ethical principles listed here that resonate well with the enterprise Risks and Governance areas of organizations.

Key Policies

PIKOM terms this collectively, a “7 by 7”, ie **Policies and Characteristics (P&C)** framework on the developmental and usage ethics.

1. **Fairness**
2. **Reliability, Safety, and Control**
3. **Privacy & Security**
4. **Inclusiveness**
5. **Accountability**
6. **Transparency**
7. **Pursuit of Human Benefits**

PIKOM proposes to prioritize the following policies you may adopt:

1. **Fairness:** We commit to conducting risk assessment on AI bias, and addressing this risk by utilizing various approaches such as diverse datasets, employing fairness-aware algorithms, and others to mitigate unfair treatment, discrimination, and negative societal implications.

Rationale: AI systems may inadvertently perpetuate biases and discriminate against certain groups. By conducting AI and employing approaches to mitigate AI bias such as fairness-aware algorithms and others, we mitigate the risk of unfair treatment and discrimination, promoting inclusivity and equity in AI development and usage.

2. **Reliability, Safety, and Control:** We will rigorously test AI systems for reliability and safety, including inherent internal controls, to ensure user trust and prevent potential harm and unreliable AI.

Rationale: Rigorous testing and internal controls ensure AI systems operate reliably and safely, maintaining user trust and preventing potential harm or unreliable performance.

- 3. Privacy & Security:** We will adopt privacy-by-design principles, encrypt data, and implement robust security protocols to safeguard against violations that could lead to legal consequences, reputational damage, or loss of user trust.

Rationale: Adherence to privacy-by-design principles and robust security measures protect user data, preventing violations that could damage trust and reputation or lead to legal consequences. The increase in cybersecurity breaches and fraud has propagated major concerns in all digital platforms, which is amplified in AI deployment. If the data falls into the wrong hands of an AI developer with ill intentions, the consequences can be extremely disastrous.

- 4. Inclusiveness:** We will ensure AI benefits all stakeholders, avoiding societal divides to prevent exclusion that may lead to societal disparities and hinder the positive impact of AI.

Rationale: Ensuring AI benefits all stakeholders, promotes societal cohesion, and prevents exclusion leading to the exacerbation of societal divides, fostering a more equitable and inclusive society.

- 5. Transparency:** We will document decision-making processes and use explainable algorithms as far as possible to foster trust, mitigate scepticism, and prevent unintended bias.

Rationale: Documenting decision-making processes and using interpretable algorithms as far as possible fosters trust and accountability, reducing scepticism and unintended biases.

- 6. Accountability:** Entities deploying AI will be held accountable for success or failure to ensure responsibility is taken for AI implementation.

Rationale: Holding entities and/or individuals accountable ensures responsible AI development and usage, mitigating the risk of potential failures or negative outcomes and thereby, fostering trust.

- 7. Pursuit of Human Benefit:** We will prioritize user input and feedback, focusing on benefiting humanity over efficiency to prioritize welfare and prevent dissatisfaction.

Rationale: Prioritizing user input and focusing on benefiting humanity over efficiency ensures AI serves human welfare, preventing dissatisfaction and promoting societal well-being.

Related risks

1. Without independent oversight or review, self-regulation may not be effective and may be subject to interpretations. Further, we are not in favour of regulations from the authorities.
2. Reputational risk is perhaps a key risk to all organizations and individuals, especially if legal issues are also involved. Hence, floundering or side-tracking any of the policies above can be catastrophic to organizations and businesses.
3. Legal risk is yet another key risk as it will not only entail loss of reputation but also monetary loss in fines and legal costs.

7 Characteristics to be Embedded in Risk and Governance Areas

In addition to the 7 key guiding policies, PIKOM also advocates incorporating the following characteristics into the enterprise risk and governance areas. In essence, these tones or characteristics must be inherent in your respective risk and governance framework.

1. **Informed Consent:** Obtaining explicit consent for using AI and providing user-friendly interfaces ensures users can make informed decisions on AI usage, mitigating the risk of dissatisfaction, legal challenges, and damage to credibility while fostering trust and transparency.
2. **Continuous Monitoring:** Committing to implementing regular audits and establishing accountability frameworks to monitor AI systems will mitigate the risk of unanticipated harms and ethical lapses.
3. **Collaboration and Openness:** Fostering collaboration, sharing insights, and encouraging open dialogue to address ethical challenges collectively promotes innovation and progress.
4. **Environmental Impact:** Optimizing algorithms for energy efficiency and adopting eco-friendly practices mitigates the environmental impact of AI and demonstrates social responsibility.
5. **Human Oversight:** Integrating human oversight mechanisms ensures human judgment is applied where necessary, preventing AI from making inappropriate decisions and upholding ethical standards.
6. **Education and Awareness:** Regular training sessions, workshops, and awareness campaigns educate stakeholders on ethical considerations, mitigating the risk of unintended misconduct and uninformed decision-making.

7. **Checks and Balances:** In proposing a self-regulating framework, independent third-party review and reporting is essential, not just for compliance but to promote continuous areas of improvement.

It is intended for these characteristics to complement and support the 7 key policies (above) in safeguarding the Ethics and Moral of AI adoption and development.

Alignment of Policies and Characteristics

This direct alignment endeavours to emphasize the specific role each characteristic plays in supporting and promoting the corresponding policy within the PIKOM "7 by 7" framework.

1. **Fairness / Collaboration and Openness** - Ensure diverse collaboration and openness in AI development to mitigate biases and promote fairness.
2. **Reliability, Safety, and Control / Continuous Monitoring** - Implement continuous monitoring to ensure AI systems maintain reliability, safety, and control, with the ability for timely human intervention when necessary.
3. **Privacy & Security/Informed Consent** - Prioritize informed consent mechanisms to uphold privacy and security standards in AI development and usage.
4. **Inclusiveness/Education and Awareness** - Promote education and awareness to ensure inclusiveness in AI technologies, considering diverse user needs and perspectives.
5. **Accountability/Checks and Balances** - Establish checks and balances to ensure accountability for AI-related decisions and actions within organizations.
6. **Transparency/Human Oversight** - Incorporate human oversight mechanisms to enhance transparency in AI processes, allowing for explanations and accountability.
7. **Pursuit of Human Benefits/Environmental Impact** - Assess and mitigate the environmental impact of AI technologies, ensuring they contribute positively to human well-being and sustainability.

E. APPENDIX

1. PIKOM 2024 Survey results

In March 2024, PIKOM conducted a survey of our members to gauge their perspectives on the ethics revolving around AI development and deployment. The following captures the essence of the survey findings.

- **93%** responded positively on whether AI developers should prioritize fairness and ensure non-biased assessments in their algorithms.
- **100%** highlighted the importance of using diverse datasets in AI development to ensure inclusivity.
- **97%** indicated the need for AI systems to prioritize transparency in their decision-making processes.
- **90%** felt there must be greater emphasis on human oversight in the deployment of AI technologies.

Recommendations include assessing AI systems for potential misinformation and disclosing any dependencies or biases, implementing human oversight for a period after implementation, emphasizing human interactions, developing conducive guidelines and regulations, using tracking systems and analytics for monitoring, establishing a repository for feedback from diverse focus groups, promoting AI usage in educational institutions while prioritizing intent for societal good over profit, and fostering human-AI collaboration for mutual benefit. These suggestions highlight the importance of incorporating human oversight, transparency, accountability, and collaboration into AI processes to mitigate risks and maximize benefits.

- **97%** are of the opinion that it is crucial for AI systems to prioritize user consent and provide user-friendly interfaces.
- **97%** believe collaboration and openness among industry players and government are essential to address ethical challenges in AI.
- **What roles can universities play in promoting ethics and morals in AI?**

Universities play a crucial role in the ethical development and application of AI. They can provide early ethical education to students, integrate AI knowledge into various courses, and promote ethical usage of AI through research and development. Additionally, universities can organize open forums to facilitate collaboration between academia, AI developers, industry players, and government agencies. They can also enforce systems to ensure students use AI responsibly in their work, offer classes on ethical programming, and conduct research to understand the evolving landscape of AI technologies. By empowering stakeholders with knowledge and promoting transparent and ethical AI practices, universities contribute to creating a responsible AI ecosystem for the betterment of society.
- **90%** emphasized the significance of prioritizing the pursuit of human benefit and happiness over efficiency in AI development.
- **100%** agreed that AI systems be rigorously tested for reliability and safety, even if it may slow down innovation including greater investments.
- **97%** stressed the importance of AI developers to consider the environmental impact of their algorithms.
- **100%** noted that continuous monitoring and accountability frameworks are necessary to ensure ethical AI practices.

The responses emphasize the importance of transparency, oversight, and accountability in the development and deployment of AI systems. Suggestions include full documentation and disclosures that can be reviewed and audited periodically, internal governance, third-party reviews, human oversight with periodic checks, selective review by third-party audits, self-regulation, adoption of AI management systems such as ISO 42001, continuous monitoring and audit processes, viable tracking systems for forensic analytics, properly designed review teams, and regular updates to AI programs to prevent false manipulation. These measures aim to ensure that AI systems are developed and used responsibly, with mechanisms in place to detect and address any potential biases, errors, or misuse.

- **73%** called for AI development and deployment to be governed by “guardrails” and regulated by third-party review.

- **How do we ensure compliance and risks on AI ethics and morals be mitigated?**
To ensure compliance and mitigate risks associated with the suggestions provided, several measures can be taken. These include:
 - Establishing stringent compliance frameworks and policies that are regularly updated based on developments in AI technology.
 - Ensuring transparency and auditability of use cases, datasets, and algorithms by third parties.
 - Implementing mechanisms for monitoring and reporting suspected wrongdoing or breaches, potentially involving regulatory bodies like MCMC.
 - Promoting the adoption of AI management systems standards such as ISO 42001 to guide ethical AI practices.
 - Utilizing blockchain technology to enhance transparency and mitigate risks in AI content.
 - Implementing real-time self-regulatory systems to monitor and enforce compliance.
 - Enforcing regulations through agile regulatory frameworks or AI acts, ensuring accountability and adherence to ethical guidelines.

- Assigning authorized bodies or professional organizations to develop and enforce industry-wide frameworks and policies.
 - Facilitating responsible and ethical AI development through human oversight, robust safeguards within AI systems, and ongoing public discourse and collaboration.
 - Conducting research and development to continuously improve AI governance and mitigate emerging risks.
- **70%** of respondents are developing or/and using AI in their business environment or intend to.

➤ **Additional comments**

Several respondents emphasized the need for collaboration among industry stakeholders, government entities, and academia to ensure that proper ethical frameworks are in place to prevent adverse impacts on society, businesses, and the environment. Others highlighted the significance of government regulation to provide a regulatory framework while still enabling innovation in AI technologies. Additionally, there is recognition that AI's potential benefits can only be fully realized if guided by ethical and moral principles. Some respondents expressed optimism about the future of AI and its potential to propel humanity forward in terms of productivity, profitability, and sustainability. Overall, the responses underscored the importance of considering ethical implications and collaboration across sectors to harness the full potential of AI responsibly.

By implementing these measures, stakeholders can work towards fostering a responsible and ethical AI ecosystem that maximizes benefits while minimizing potential risks to society, businesses, and individuals.

2. Existing Laws and Policies

Existing laws that are applicable include:

- Personal Data Protection Act (PDPA) 2010 can be used as a foundation. The PDPA includes principles of data protection, like disclosure, security, retention, and data integrity, which can be used as a basis to guide the responsible management of data in relation to the development of AI.
- Additionally, there are other regulations that apply to the use or misuse of technology, such as the Communications and Multimedia Act 1998, Sedition Act 1948, Defamation Act 1957, Penal Code, Child Act 2001, Child (Amendment) Act 2016, and Sexual Offenses against Children Act 2017.
- There are also copyright laws to protect the intellectual property of software, and consumer protection laws like the Consumer Protection Act 1999 and the Consumer Protection (Electronic Trade Transactions) Regulations 2012 that can safeguard consumers' rights in the context of AI technology.

3. References

1. Policy paper: The Bletchley Declaration by Countries attending the AI Safety Summit published 1 November 2023
2. Malaysian AI RoadMap (AI-RMAP) 2021-2025 by MOSTI
3. WITSA AI Position Paper - BUILDING TRUST AND DELIVERING ON THE PROMISE OF ARTIFICIAL INTELLIGENCE, September 28, 2023