Department of Industry, Science and Resources
Artificial Intelligence Ethics

Artificial Intelligence Ethics Unleashed: Navigating Australia's Regulatory Landscape

EXECUTIVE SUMMARY

The policy brief addresses the critical need for regulation in the field of Artificial Intelligence (AI) in Australia. The lack of knowledge and understanding of AI has led to bias and discrimination, which can compromise data privacy, public safety, and ethical standards. This brief presents a critical view of "Australia's AI Ethics Principles", the current Australian framework, while using the multiple streams framework, which is policy decisions made through the convergence of three separate streams of activity: problem, policy, and politics, to assess the issue (Cairney, 2019). The research reveals that keeping the current policy poses significant risks in the future due to the potential dangers associated with AI. This brief also shows the economic, political, and budgetary costs of implementing new AI regulations. The recommendations explore increasing AI systems' fairness, privacy protection, transparency, and explainability at various levels.

BACKGROUND AND IMPORTANCE OF THE PROBLEM

Artificial intelligence systems have become prevalent in Australia, but without proper regulation, they pose risks such as potential discrimination, bias, and malicious use, causing continued scepticism of AI systems in Australians(Mallis, 2023). This is illustrated in recent studies (Gillespie et al., 2021, 2023) indicating that Australians' trust in AI systems has only slightly improved from 32% in 2021 to 34% in 2023. According to the multiple streams framework, this falls under the problem stream, where "attention lurches to a policy problem" (Cairney, 2019, p. 196). In other words, the problem stream is the emergence of a significant and urgent concern that policymakers, the media, interest groups, or the public perceive. The issue of AI's impact on society is a cause for anxiety for both Australians and the global community and is likely to raise concern among the government (Gillespie et al., 2021; Gillespie et al., 2023). To address this challenge, the Department of Science, Industry, and Technology aspires to instil confidence in Australians regarding AI systems' safety, security, and reliability, allowing them to "realise its benefits" (Dawson et al., 2019, p. 3).

In 1955, when John McCarthy, a computer scientist, coined the term Artificial Intelligence (Dawson et al., 2019). However, recent incidents, such as using an automated system implemented in 2012 in a Houston, Texas school district to fire teachers based on students' test scores, have raised ethical concerns (Langford, 2017). This emphasises the critical need to address the ethics of AI systems in Australia and has resulted in the field's rapid expansion globally in the past few years. In 2019, Australia released a voluntary framework for AI ethics, "AI Ethics Framework,"(Dawson et al., 2019).

The key stakeholders regarding the issue are the Minister for Industry and Science and their department, the Australian Human Rights Commission, the Prime Minister and the Department of the Prime Minister and Cabinet, and the Commonwealth Scientific and Industrial Research Organisation's Data61. The Minister for Industry and Science and their department handle the policy for AI ethics. The Australian Human Rights Commission (2021) is focused on protecting human rights and with a potential interest in addressing ethical considerations and potential biases in AI systems. The Prime Minister and the Department of the Prime Minister and Cabinet could be interested in shaping AI regulatory policies, whereas Data61, Australia's leading digital research network, is responsible for driving digital innovation and transformation and may be inclined towards ensuring the safe, secure, and ethical development and deployment of AI (Range, 2019).

The politics stream of the multiple streams framework is when "policymakers have the motive to turn that solution into policy" (Cairney, 2019, p. 198). Federal politicians like Senator Ross Cadell and MP Julian Hill exemplify the politics stream. These politicians are considered critical stakeholders in shaping AI ethics policies and regulations in Australia. Senator Cadell emphasised, "To mitigate these dangers, it is important that we develop AI in a responsible and ethical manner….ensuring that AI systems are transparent and accountable…." (Commonwealth, 2023, p. 37). Lastly, MP Hill has stated that AI "presented 'existential' and 'catastrophic' risks if untamed … risks of AI increasingly surpassed 'asteroids, runaway climate change….'" and the need for "...regulation and oversight to ensure that AGI is developed and used in a responsible and ethical manner" (Hill, 2023). Policymakers must consider the potential risks associated with AI systems and ensure that these technologies are developed and used ethically and responsibly.

CRITICAL ANALYSIS

While the study of artificial intelligence has rapidly grown in the past few years, yet Australia lacks policies or regulations governing its use. However, in 2019, the Department of Science, Industry, and Technology released a voluntary framework, the "Artificial Intelligence Ethics Framework ", consisting of eight principles, including "Fairness", "Privacy protection and security", "Reliability and safety", and "Transparency and explainability", that serves as the gold standard for ethical AI in Australia(Dawson et, al., 2019). As technology and its ethical implications continue to evolve rapidly, it is necessary to re-evaluate and update existing policies on AI ethics in Australia in consultation with the stakeholders.

Although Australia has no policies or regulations governing AI use, they have several policies related to AI ethics, which could be considered existing obligations that can impact policy changes. First off, "Australia is a signatory to seven core international human rights agreements", meaning any changes to the existing policy on AI would have to abide by the principles found in the agreements, therefore protecting the human rights of their citizens (Dawson et al., 2019, p. 18). Additionally, Australia has legislation regarding data protection, including the Privacy Act 1988 (Dawson et al., 2019), which regulates the handling how a

person's information, such as their name, address, and bank account details (Dawson et al., 2019, pp. 19-20). Any policy changes related to AI ethics must protect individuals' privacy rights.

External stakeholders, such as the private sector and academia, will likely differ in their reactions to proposed policy changes. The private sector may be concerned about the potential impact of stricter AI regulations on its development. However, some companies, such as Google and SAP, have released guidelines for responsible AI to build consumer trust and address the potential dangers of AI systems (Jobin et al., 2019). Academia would welcome policy changes that increase transparency and accountability, covering algorithmic bias and explainability. Since no Australian think tanks focus on how the government should regulate AI, they will not have a significant role (Greenleaf et al., 2019).

Implementing AI ethics policy could have positive and negative effects on the economy. Stricter regulations could increase compliance costs for businesses that use AI, leading to slower innovation and impeding economic growth (Tartaro, 2023). Conversely, insufficient regulations of AI could impede economic growth by “irreversibly destroy[ing] society’s trust in new technologies” (Erdélyi & Goldsmith, 2022, p. 2). To balance these effects, the government should create policies encouraging innovation while ensuring responsible AI development.

The government will face costs in implementing AI regulations. Although there is no research into what it would take in a government budget to implement policies, the government could estimate costs. The initial costs include research and consultation with interest groups to decide what goes into the policy. After its creation, the policy would require the government to spend money so that government agencies can ensure AI systems are regulated effectively and adhere to ethical guidelines. Additionally, there may be costs associated with informing the private sector, academia, and other applicable parties about the new policy and how it could impact them. However, the effective regulations of AI ethics are crucial in protecting the public interest, as the government is responsible for ensuring ethical and responsible AI.

The public's reaction to new AI ethics policy is an important consideration. While some Australians may be confident in the government's ability to regulate and govern AI in the public's best interests, others may be sceptical. In 2021, Gillespie et al. (2021, p. 22; 2023, p. 32) found that 68% had moderate confidence in their government's AI regulation, but their 2023 study reported an average score of 2.9 out of 5, indicating decreasing trust in AI systems. The government can clarify any misconceptions about the policy to gain public approval.

As the government considers the issue of AI ethics, it is essential to provide specific options for the minister to consider. The option chosen by the minister will exemplify the policy stream of the multiple streams framework, "a feasible solution to that problem is available” (Cairney, 2019, p. 197), which would be considered a policy instrument outlining the regulations for ethical AI in Australia. The policy development process would involve a multi-stakeholder consultation process, including industry professionals and academics, aligning with Kingdon's idea that policy development occurs within networks of individuals, such as researchers, congressional staffers, and academics (Kingdon, 1984, p. 18). Therefore, the window of

opportunity for regulating AI in Australia appears to be open, as all three streams are in the process of coming together, and there is a chance for policymakers to address this issue by creating policies and regulations that ensure the safe, secure, and ethical development and deployment of AI systems.

POLICY RECOMMENDATIONS AND CONCLUSION

The recommendations below offer three distinct options, maintaining the status quo, mild reform, and major reform.

Option 1: Maintaining the status quo

Sticking with the current voluntary “Artificial Intelligence Ethics Framework" poses significant risks for Australia and threatens the country’s status as a global leader in developing ethical AI. Without further action, biased algorithms and the lack of transparency could have significant consequences.

Option 2: Mild reform

This option is based on the European Union’s “Artificial Intelligence Act”. The approach involves implementing regulations with the following recommendations:

Recommendation 1: The ban of AI “Social Scoring”.

- This will improve “Fairness” and “Privacy protection and security, " aligning with the current Australian framework. This will ensure that AI cannot discriminate against a person for their past and use it against them (Feingold, 2023).

Recommendation 2: A fine of up to $50,000,000 or 6% of global profits for breaching the legislation (Coulter & Mukherjee, 2023).

- The proposed significant fines aim to discourage non-compliance and ensure adherence to regulations.

This approach demonstrates a solid commitment to addressing AI ethics concerns by aligning with international standards. Given its alignment with established regulations, it would be easier to pass through Parliament. Shortcomings include exempting non-Australian citizens from the rules and inadequately addressing human rights dangers with AI (Lamberti, 2023).

Option 3: Major reform

This option entails adding bold reform to the mild reform recommendations. The legislation would include the recommendations below.

Recommendation 1: The establishment of a national AI ethics database.

- This will collect data on the ethical implications of AI, including the potential biases, risks, and impact on human rights, helping the public and private sectors create AI systems in alignment with regulations.

Recommendation 2: Establish a new government agency dedicated to AI ethics.

- This ensures the enforcement of ethical standards to the highest level.

Recommendation 3: Tax breaks for companies that use ethical AI.

- This incentivises companies to comply with regulations.

Recommendation 4: Investing greatly into the Australian ethical AI research

- This ensures that Australia is a leader in the innovation of responsible AI.

Australia has the opportunity to become a global leader in responsible AI adoption through these reforms, that the government's responsibility to address AI ethics. To achieve this, significant resources is necessary to ensure the legislation's successful governance. However, some members of the public and Parliament may oppose the big spending and the tax cuts, perceiving them as favouring "big business" and excessive government intervention.

The recommendations underscore the importance of addressing AI ethics and guide the minister in shaping AI policy in Australia. Maintaining the status quo may be inadequate, while mild reforms aim to balance concerns and comply with international standards. Major reforms require significant investment but can establish Australia as a leader in AI ethics. Implementing a comprehensive legal and regulatory framework for AI ethics is essential for Australia to realise the benefits of AI while minimising the risks and challenges.

References List

Australian Human Rights Commission. (2021). *Human Rights and Technology*. https://tech.humanrights.gov.au/sites/default/files/2021-05/AHRC_RightsTech_2021_Final_Report.pdf

Cairney, P. (2020). *Understanding public policy: Theories and issues*. Red Globe Press.

Commonwealth. *Parliamentary Debates*. Senate. 8 March 2023. https://parlinfo.aph.gov.au/parlInfo/download/chamber/hansards/26443/toc_pdf/Senate_2023_03_08.pdf;fileType=application%2Fpdf

Coulter, M., & Mukherjee, S. (2023, March 22). *Explainer: What is the European Union AI act?*. Explainer: What is the European Union AI act? | Reuters. https://www.reuters.com/technology/what-is-european-union-ai-act-2023-03-22/

Dawson, D., Schleiger, E., Horton, J., McLaughlin, J., Robinson, C., Quezada, G., Scowcroft, J., & Hajkowicz, S. (2019). *Artificial Intelligence: Australia's Ethics Framework*. https://www.csiro.au/-/media/D61/Reports/Artificial-Intelligence-ethics-framework.pdf

Erdélyi, O. J., & Goldsmith, J. (2022). Regulating Artificial Intelligence: Proposal for a global solution. *Government Information Quarterly*, 39(4), 101748. https://doi.org/10.1016/j.giq.2022.101748

Feingold, S. (2023, March 28). *The EU's Artificial Intelligence Act, explained*. World Economic Forum. https://www.weforum.org/agenda/2023/03/the-european-union-s-ai-act-explained/

Gillespie, N., Lockey, S., & Curtis, C. (2021). *Trust in Artificial Intelligence: A Five Country Study*. https://doi.org/10.14264/e34bfa3

Gillespie, N., Lockey, S., Curtis, C., Pool, J., & Akbari, A. (2023). *Trust in Artificial Intelligence: A Global Study*. https://doi.org/10.14264/00d3c94

Greenleaf, G., Clarke, R., & Lindsay, D. F. (2019). Does ai need governance? – the potential roles of a 'responsible innovation organisation' in Australia (submission to the Australian Human Rights Commissioner on the White Paper Artificial Intelligence: Governance and leadership). *SSRN Electronic Journal*. https://doi.org/10.2139/ssrn.3346149

Hill, J. (2023, February 6). *Artificial general intelligence (AGI)*. https://www.julianhillmp.com/Artificial-general-intelligence-(AGI)_06022023.html

Jobin, A., Ienca, M., & Vayena, E. (2019). The global landscape of AI ethics guidelines. *Nature Machine Intelligence*, 1(9), 389–399. https://doi.org/10.1038/s42256-019-0088-2

Kingdon, J. W. (1984). *Agendas, Alternatives, and Public Policies*. HarperCollins.

Lamberti, L. (2023, January 26). *Is the AI act missing safeguards on migration?*. The Parliament Magazine. https://www.theparliamentmagazine.eu/news/article/ai-act-migration-technology

Langford, C. (2017, May 8). *Houston schools must face teacher evaluation lawsuit*. Courthouse News Service. https://www.courthousenews.com/houston-schools-must-face-teacher-evaluation-lawsuit/

Mallis, A. (2023, April 18). *Cover story: How generative AI has impacted ethics*. Digital Nation. https://www.digitalnationaus.com.au/video/cover-story-how-generative-ai-has-impacted-ethics-593342

Range, J. (2019, April 18). *CSIRO's Data61: A global inspiration for Digital Innovation*. CSIRO's Data61: A global inspiration for digital innovation - Algorithm. https://algorithm.data61.csiro.au/csiros-data61-a-global-inspiration-for-digital-innovation/

Tartaro, A., Smith, A. L. L., & Shaw, P. (2023). *Assessing the Impact of Regulations and Standards on Innovation in the Field of AI*. https://arxiv.org/abs/2302.04110