Article Information

Authors: Michael Bacina, Steven Pettigrove, Luke Higgins
Service: Blockchain, FinTech
Sector: Financial Services, IT & Telecommunications

Australia issues interim response on Safe and Responsible AI

Michael Bacina, Steven Pettigrove and Luke Higgins of the Piper Alderman Blockchain Group bring you the latest legal, regulatory and project updates in Blockchain and Digital Law.

The Australian Government has released its interim response to the "Safe and responsible AI in Australia" consultation paper. The 25-page response states the adoption of AI and automation could inject an additional $170 – $600 billion annually into Australia's GDP by 2030, heralding a transformative era for the technology and Australia alike.

The response reiterates that the extensive consultation process made clear the benefits of AI systems and its applications to improving the wellbeing and qualify of life for Australians. However, the Australian Government noted that there was currently a significant lack of public trust in the safety and responsible use of AI systems.

A recent survey reveals that only one-third of Australians believe that adequate guardrails are in place to ensure the safe design, development, and deployment of AI, with this skepticism acting as a roadblock to widespread business adoption and public acceptance.

Addressing these concerns, the response paper calls for a balanced approach, recognising the various ways AI technologies are employed. While many applications, such as spam filters and business optimisation tools, pose low risks, other applications like predicting recidivism, job suitability or enabling self-driving vehicles, are deemed higher risk due to their potential negative impacts.

The proposed risk-based approach in the discussion paper emphasises setting up additional guardrails to mitigate harms in high-risk settings, a crucial step in fostering responsible AI development and deployment. Several submissions received by the Australian Government identified specific risks in relation to advanced AI models, referred to as "frontier" models. These cutting-edge systems can generate new content quickly, raising concerns about their potential misuse.

The global focus on creating sufficient guardrails for AI development and deployment is intensifying, with governments demanding proactive steps from companies in high-risk contexts. Jurisdictions worldwide are exploring a variety of measures, from voluntary commitments to mandatory obligations, to ensure the safety of AI products. Notably, the EU has etched a deal for the world's first comprehensive laws for artificial intelligence (AI), a groundbreaking move with repercussions echoing far beyond the continent's borders. The recent AI Safety Summit hosted by the United Kingdom in November 2023 further emphasised the urgency of addressing AI safety concerns on a global scale. Australia, along with the EU and 27 other countries, signed the Bletchley Declaration, committing to collaborative efforts in AI safety testing and the development of risk-based frameworks.

Looking domestically, the Government's response noted that Australia is currently undertaking work to strengthen existing laws in areas that will help to address known harms with AI, including amendments to privacy laws and the Online Safety Act. The Government's response notes a preliminary analysis of submissions found at least 10 legislative frameworks that may require amendments to respond to applications of AI.

While there was a consensus in submissions that voluntary guardrails are insufficient, the Governments appears to have accepted the need for a nuanced regulatory approach. Mandatory guardrails, it is proposed, should be tailored to high-risk applications, allowing low-risk innovation to flourish unimpeded.

As the Australian Government considers what these mandatory guidelines may look like for AI development and use, the

response notes that it will be taking immediate action in the following three ways:

1. working with industry members to develop a voluntary "AI Safety Standard", implementing risk-based guardrails for industry;
2. working with industry members to develop options for voluntary labelling and watermarking of AI-generated materials (e.g. AI generated images and videos); and
3. establishing an expert advisory body to support the development of options for further AI guardrails.

Based on the Government's interim response, a two track strategy to regulating AI in Australia appears likely, combining voluntary measures and legislative changes in a number of areas. As AI technologies continue to develop rapidly, the Government will need to strike a balance between an approach which mitigates known harms caused by AI, while allowing Australians to harness the benefits of the technology and develop local expertise and enterprise in this emerging field.