4 October 2024

Department of Industry, Science and Resources
GPO Box 2013
Canberra ACT 2601

By email: aiconsultation@industry.gov.au

Australian Payments Network (AusPayNet) welcomes the opportunity to respond to the proposals paper for introducing mandatory guardrails for artificial intelligence (AI) in high-risk settings.

AusPayNet is the industry association and self-regulatory body for the Australian payments industry. We manage and develop standards and guidelines governing payments in Australia. Our purpose is to create confidence in payments by setting enforceable industry standards for a safe, reliable and effective payments system; leading transformation in payments to drive efficiency, innovation and choice; and being the home for ecosystem collaboration and strategic insight. AusPayNet currently has more than 150 members including financial institutions, payment system operators, major retailers and financial technology companies.

This submission builds on the feedback AusPayNet provided in response to the Department's earlier consultation on safe and responsible AI in Australia in August 2023 (Appendix 1). The views expressed in this submission are those of AusPayNet Management, and may not necessarily represent the views of each of our members.

Introduction

AI has the potential to deliver significant benefits across many areas of Australia's economy and society. This has long been recognised by the payments industry, which has been integrating AI into payments systems and processes for many years to help drive efficiency, innovation and safety for businesses and consumers. Many of the existing use cases of AI in payments – including fraud detection, customer authentication, and process automation – are widely trusted and adopted by payments service providers, businesses and customers across Australia.

As AI models become more sophisticated in their ability to process vast amounts of data, recognise complex patterns, and generate intelligent insights, the potential applications of this technology across society and the economy also continue to evolve. For the payments ecosystem, these opportunities include further enhancing fraud detection and risk management capabilities, improving operational efficiency and resilience, and supporting innovation and improvements in payments technologies and services, including through more personalised customer and merchant experiences.

At the same time, we recognise that the ongoing developments in AI also have the potential to generate significant risks if not designed and deployed responsibly. As noted in the consultation paper, this could deter adoption and limit Australia's ability to capture the full benefits of this technology. AusPayNet therefore supports the Government's continued work on ensuring that Australia has clear and proportionate governance frameworks in place to mitigate risks, while supporting the continued development and adoption of effective and trustworthy AI technology. As a standards-setting body,

AusPayNet recognises that appropriate governance mechanisms can support innovation and growth for the benefit of all stakeholders.

AusPayNet generally supports the proposed implementation of mandatory guardrails for the use of AI in high-risk settings. The proposed framework broadly aligns with the feedback provided in our earlier submission, which detailed our support for adopting a risk-based and principles-driven approach to AI regulation that balances the need for both consistency and proportionality in regulating the many different AI use cases across different industries. We also welcome the broad alignment with emerging international standards and best practice, which will support both the ability of Australian businesses to take advantage of AI technologies developed offshore, as well as continued investment in Australia's own AI capacity and capabilities.

The intention of our submission is to provide some insights and considerations about the proposed guardrails from the perspective of the Australian payments industry, to help inform any further refinements to the governance framework. We believe that such sector-specific insights can help ensure that the overarching guidelines strike an optimal balance between ensuring safety and fostering AI innovation in Australia.

Defining High-risk AI

AusPayNet broadly supports adopting a principles-based approach to defining high-risk AI. This flexible approach will facilitate consistency in interpretation across the many different applications of AI, and allow for ongoing developments in AI technology. While providing illustrative examples of high-risk use cases could serve as useful guidance, a list-based approach is unlikely to be effective, since the risks associated with the use of AI can vary significantly even within a given domain or technique. For example, Table 1 lists 'biometrics' as a high-risk use case; however, the risks associated with using facial recognition technology to unlock a phone or even authorise a mobile payment would be much lower than when used for public surveillance.

Additional guidance from sectoral regulators on the application of the definitional principles to key use cases within their industries would also be helpful. In particular, the assessment of the potential for adverse impacts to individuals' mental health, and the thresholds for the 'severity and extent' of harms beyond which use cases would classify as high risk could potentially lead to subjective and inconsistent interpretation across sectors, and even businesses within the same industry.

We also query whether the risk assessment framework (or at least the subsequent application of some of the guardrails) should consider the counterfactual of *not* using AI for that particular application. In the context of payments, for example, not using AI for fraud detection would very likely lead to a significant increase in financial crime. There may be a case for weighing such factors against any potential risks posed by the AI systems, to help ensure that beneficial AI applications are not unduly restricted.

Developers and Deployers

We support the Government's proposal that both developers and deployers of high-risk AI systems should share responsibility for ensuring the safe development and use of AI technology, with the guardrails distributed according to which actors are best equipped to address the risks associated with a particular stage of development.

However, we do have some concerns about the breadth and clarity of the two definitions, and particularly the point at which a deployer might classify as a developer. Many entities in the payments ecosystem – particularly fintech companies – may not have the capacity to develop their own AI models, and would therefore chose to adapt or fine-tune existing models for their specific use cases. A prominent example would be fraud detection systems. These are often developed by larger businesses, such as payments system operators, but allow entities leveraging those tools to make minor adjustments to certain parameters of the model. The currently proposed definitions are unclear about the extent of such fine-tuning that would classify those smaller entities as developers, and potentially subject them to a disproportionate regulatory burden. We therefore suggest providing more detailed guidance on the degree of model adaptation that would classify an entity as a developer, perhaps with industry-specific examples to help clarify this distinction. It may also be useful to consider a potential 'middle' category for entities that significantly adapt but do not fundamentally develop AI systems.

AusPayNet also notes that many Australian entities, particularly smaller ones, are likely to rely on AI systems developed by large global technology companies. These local entities may not have the bargaining power to enforce all of the proposed guardrails on these global developers. While the alignment of Australia's mandatory guardrails with international standards should help mitigate this risk, we encourage the Government to consider how this potential dynamic might affect the practical implementation of the guardrails for some entities.

Relatedly, we also encourage the Government and/or sectoral regulators to implement regulatory sandboxes for AI applications. This would allow smaller entities in particular to carry out controlled development and testing of innovative AI systems under regulatory supervision, without incurring an upfront regulatory burden that may inhibit innovation.

Mandatory Guardrails

AusPayNet supports the Government's commitment to establishing robust guardrails for high-risk AI applications. The proportionate and tailored implementation of these guardrails should significantly enhance the safety and trustworthiness of AI systems. As noted earlier, we also welcome the broad alignment of the guardrails with emerging international standards.

A key concern for us is ensuring that organisations have appropriate flexibility in how they implement these guardrails, based on the specific context and application of their AI systems. In particular, some use cases in the payments industry highlight that there may be instances where strict adherence to one guardrail could potentially compromise another. For example, in AI-based fraud detection systems, there might be tension between the requirements for explainability and those for system security and effectiveness. A highly sophisticated AI model would likely be more effective at detecting complex fraud patterns, but could be less explainable than a simpler model. In such cases, the ability to balance these requirements based on the specific use case and risk profile is crucial. We suggest that the final guidelines explicitly acknowledge such potential trade-offs, and provide guidance on how organisations should approach such situations.

We also strongly support the emphasis on rigorous testing and ongoing monitoring of AI systems. This is particularly important in the payments industry, where AI systems often operate in real-time and deal with sensitive financial data and decisions. However, for AI applications such as fraud detection systems, some of the most effective development and testing often occurs in real-world deployment,

where the system can continuously learn from and adapt to actual transaction patterns and emerging fraud techniques in real time. We suggest that the guidelines consider allowing for phased or limited real-world testing for certain types of AI systems, under close monitoring and with appropriate safeguards. This could be particularly valuable for adaptive AI systems that improve their performance over time based on real-world, real-time data.

Similarly, while we generally support the principles of explainability and transparency, we suggest that these should be applied proportionately based on specific use cases and risk profiles. In the case of AI-based fraud systems, for example, full explainability of each decision might not always be necessary or beneficial if other safeguards are in place – such as rigorous testing to ensure appropriate accuracy and bias mitigation. Moreover, too much transparency around the decisions made by fraud detection systems could potentially introduce new risks and compromise the effectiveness of the system, by assisting criminal actors in learning how to evade detection. We therefore suggest a nuanced approach that balances the need for explainability and transparency with other considerations such as system effectiveness and security.

Finally, the effective implementation of Guardrail 6, on informing end-users regarding AI-enabled decisions and interactions, should also be considered. There are many cases in which such transparency might be necessary from an ethical and trust-building perspective, particularly when customers should be able to make informed decisions about whether to engage in the AI-driven services. However, we believe that such notification would not be appropriate or helpful in some circumstances. For example, in the fast-paced, high-volume payments environment – where most transactions have been approved or denied on the basis of AI-led decisioning for many years – constant notifications about AI use could reduce payments efficiency and lead to customer alert fatigue, potentially causing users to overlook important information in the future.

Noting that some consumers may have general anxiety or mistrust around the use of AI, explicit disclosure of AI use in every instance of fraud monitoring and similar decisioning could also lead to unwarranted concerns about fairness and accuracy – and a corresponding increase in complaints that businesses needs to handle – even where the AI model has been proven to reduce bias and increase accuracy relative to human-based decision making. Where businesses deploying AI systems conduct rigorous ongoing testing to ensure their accuracy and fairness, we consider that a customer's decision to challenge an outcome should not be unnecessarily biased by whether the decision was made by AI, a human, or a mix of the two. Furthermore, it should not be possible for a customer to opt out of AI-based transaction monitoring, even if they wished to do so. Recurring notification would therefore not provide the customer with any actionable insights regarding the use of AI to inform decisions relevant to them. We therefore suggest that the guidelines allow for a nuanced approach to disclosure, considering factors such as the specific use case, the potential impact on system effectiveness, and the overall user experience. For example, it might be more appropriate to provide general information about AI use in terms and conditions or privacy policies for certain systems, rather than real-time notifications for each AI-driven interaction. The goal should be to ensure meaningful transparency without compromising security or overwhelming users with excessive notifications.

Regulatory Framework

As detailed in our earlier submission, we support a regulatory approach to AI that balances consistency of regulatory outcomes across the economy, with the flexibility to proportionately tailor the

regulatory response to specific contexts and use cases. This is supported by the discussion above, which highlights the need for tailored implementation of certain guardrails to mitigate the introduction of new risks, undue regulatory burden, or other unintended consequences. Our earlier submission had also expressed support for maximising reliance on existing legislation and regulatory expertise where possible.

AusPayNet therefore broadly supports the framework approach (option 2), which we believe best balances consistency with flexibility, by ensuring that all AI systems are subject to the same baseline standards, while allowing for sector-specific implementation of those standards through well-established regulatory frameworks and expertise.

As indicated in the consultation paper, standards can also play an important role in supporting this kind of regulatory framework, to ensure the effective practical implementation of the overarching guardrails in specific settings. As the self-regulatory body for the payments industry, AusPayNet is ready to lead the development of any technical standards for the Australian payments industry (or the adaptation of relevant standards developed by international working groups), where the industry and our sectoral regulators see a need for such technical standards to effectively comply with any of the mandatory guardrails.

Finally, we note that implementing these guardrails will require significant adjustments for many organisations, particularly those with existing AI systems. Appropriate transition arrangements will be crucial to minimise disruption to existing services. These arrangements might include phased implementation, with priority given to the highest-risk applications, and could potentially vary by sector and risk level.

Conclusion

AusPayNet appreciates the opportunity to respond to the Government's consultation on mandatory guardrails for AI in high-risk settings. This is a critical step in ensuring that the potential risks of AI can be mitigated appropriately, while promoting the continued development and adoption of AI so that Australia can realise the full benefits of this technology. AusPayNet looks forward to continuing our engagement with the Government as this work progresses over the coming years. Please contact Kateryna Occhiutto, Head of Policy & Insights [redacted] if you have any further questions.

Yours sincerely,

[redacted]

Andy White
Chief Executive Officer
Australian Payments Network