Ms Meghan Quinn PSM
Secretary
Department of Industry, Science and Resources
Industry House, 10 Binara Street
Canberra ACT 2601

AIConsultation@industry.gov.au

ACMA file reference: ACMA2019/1496-39

Dear Secretary

ACMA submission: Proposals Paper for introducing mandatory guardrails for AI in high-risk settings

The Australian Communications and Media Authority (ACMA) is the independent statutory authority responsible for the regulation of broadcasting, radiocommunications and telecommunications in Australia, including spam and telemarketing. The ACMA also regulates aspects of online content advertising[1], online gambling and also oversees digital platform industry compliance with the voluntary Australian Code of Practice on Disinformation and Misinformation.

The ACMA welcomes the opportunity to contribute to the Australian Government's Safe and Responsible Artificial Intelligence (AI) work program, and comment on its proposals paper for introducing mandatory guardrails for AI in high-risk settings. The ACMA supports efforts to improve the safety of AI systems and build public confidence in AI through effective management of the risks posed by these technologies. Introducing mandatory guardrails to address high-risk uses of AI – when coupled effectively with existing regulatory tools and responses – will support Australians to engage more confidently with the technology.

A number of the sectors we regulate are early adopters of AI with many telecommunications companies, broadcasters and digital platforms deploying AI. The ACMA's role is to apply our existing regulatory frameworks to address any harms that may result from deployment of AI in those sectors. This submission explains the use of AI in some of the sectors we regulate and considers whether these uses constitute high-risk. The submission also flags the need for regulatory coherence between existing frameworks and enforcement of mandatory guardrails.

This submission should be read alongside the submission made by the Digital Platform Regulators Forum (DP-REG). DP-REG comprises the Australian Competition and Consumer Commission, the ACMA, the eSafety Commissioner and the Office of the Australian Information Commissioner. Through DP-REG, the ACMA works to better understand, assess and respond to the benefits, risks and harms of technology, including AI.

The ACMA reiterates its support for the positions put forward in DP-REG's submission.

---

[1] Online content advertising includes restrictions on gambling advertising during live-streamed sport and illegal internet gambling services.

The ACMA's contribution to the Australian Government's response to generative AI

The ACMA, as a member of DP-REG, conducts joint work to better understand digital platform technologies and their implications for consumer protection, competition, the media and information environment, privacy and online safety within the digital platform context. On 19 September 2024, DP-REG published its latest examination of technology [working paper](#), assessing the impact of multimodal foundation models (MFMs). The paper considered how the use of MFMs to generate multiple types of content, such as image, audio and video, raised concerns around enabling scams and deceptive practices, the spread of misinformation and disinformation, the generation of harmful content and the loss of control over personal information. In 2024-25, DP-REG's Digital Technology Working Group will continue to jointly explore relevant digital platform technologies and their regulatory implications.

Separately, the ACMA's June 2024 submission to the [Joint Select Committee on Social Media and Australian Society](#) highlighted the high utilisation of algorithms and recommender systems by our regulated entities in the delivery of content and advertising to Australians. This poses a range of benefits, risks and challenges – which also apply to the use of AI. These challenges, and the responses needed to address potential harms, will continue to evolve as deployment of the technology shifts in different ways in each sector we regulate.

AI and our regulatory responsibilities

The ACMA has responsibilities under the *Broadcasting Services Act 1992*, the *Radiocommunications Act 1992*, the *Telecommunications Act 1997* (Tel Act) and related Acts, like the *Spam Act 2003*, the *Do Not Call Register Act 2006* and the *Interactive Gambling Act 2001*. Our regulatory functions are set out in the *Australian Communications and Media Authority Act 2005*. Considering the use and impact of AI on the sectors we regulate supports us to deliver public interest outcomes, manage risks and protect the interests of the community in an effective and efficient way.

Many of the guardrails, outlined in the proposals paper, set out the types of actions that the ACMA already expects our regulated entities to be taking. That is supporting public accountability, providing transparency about systems and processes and addressing risks and harms.

Unsolicited communications and scams

The regulation of telecommunications scams currently occurs under the telecommunications regulatory framework, primarily the Tel Act and the Reducing Scam Calls and Scam SMS Industry Code (Scams Code), which is registered under Part 6 of the Tel Act. Current obligations in the Scams Code are outcomes-focused requiring telecommunications providers to:

- monitor the networks for scam traffic,
- notify this traffic to other providers and government agencies and
- take disruption actions.

Within the Scams Code, telecommunications providers have some flexibility to determine how they meet these obligations, including using AI tools.

Already, some telecommunications providers are implementing scam monitoring arrangements via machine learning and narrow AI systems to detect scam traffic in real time. There may also be scenarios where personal information – collected with the person's consent – could support AI deployers to disrupt deepfakes that are being used by bad actors.

The potential use of AI to benefit consumers could be better reflected in the proposed principles and mandatory guardrails in the proposals paper. This could be achieved through incorporating positive obligations to deploy and support individual users to engage with AI models where there are benefits.

There is also potential for bad actors to use AI to facilitate scam attacks on telecommunications networks. This could involve voice recordings within calls (including calls impersonating known persons) and testing of network blocks and weaknesses. While the extent that bad actors use AI technology to conduct scam activities is not known, the deployment of AI by bad actors may create risks for consumers that need to be carefully considered. It is possible that the guardrails flagged in the Department of Industry, Science and Resources' proposals paper could capture the deployment of AI models by these bad actors.

On 13 September 2024, the Australian Government released [exposure draft legislation](#) for a new Scams Prevention Framework. The new framework contemplates that scam prevention for the telecommunications sector would sit within overarching and sector specific rules under the new framework. The proposed frameworks do not explicitly refer to the use of AI.

Misinformation and disinformation

AI also has the potential to contribute to the spread of misinformation and disinformation.

This could include the use of multimodal foundation AI models to generate convincing and realistic images, videos and audio of individuals or events that never occurred (e.g., deepfake videos of authority figures spreading false information). Additionally, large language models can generate 'hallucinations' that can be difficult for both experts and non-expects to detect. These models may be used by bad actors to disseminate misinformation on a large scale, including by mimicking the style of authoritative sources of news.

In Australia, minimising the risk of harm from misinformation and disinformation on digital platforms has been the subject of self-regulation since 2021 through the [Australian Code of Practice on Disinformation and Misinformation](#). The ACMA reports on the effectiveness of the code and signatory actions to identify, assess and address misinformation and disinformation on their services.

On 18 September 2024, the ACMA released our [third report to government](#) on digital platforms' efforts under voluntary arrangements to combat misinformation and disinformation. In that report, we highlighted that some digital platforms are increasingly developing and revising their systems and processes to address the increasing manipulation of their services. For example, Adobe leads work on the Content Authenticity Initiative (CAI), which is focused on supporting transparency around the provenance of images, documents, media and content. However, we also noted that AI continues to play an influential role in the propagation of inauthentic behaviour. We consider that the use of AI to undermine information integrity in Australia constitutes a high-risk use case.

On 12 September 2024 the Australian Government introduced the *[Communications Legislation Amendment (Combatting Misinformation and Disinformation) Bill 2024](#)* to Parliament. The Bill would enable the ACMA to monitor digital platform compliance with obligations to prepare and publish current policies, media literacy plans and risk assessments and use information-gathering and transparency powers to strengthen voluntary arrangements and improve accountability of major digital platforms (not just code signatories). It would also empower the ACMA to request and approve codes, or make standards, when necessary to prevent and respond to misinformation and disinformation on digital platforms as well as systemic issues.

There is likely to be significant cross-over between the enforcement of these powers and mandatory guardrails for high-risk AI that are focused on supporting transparency and accountability from deployers of AI. For example:

- Digital platforms would be required to report publicly on the outcomes of risk assessments. This obligation may cross-over with obligations to implement risk management processes under the second guardrail in the proposals paper.

- The ACMA may make digital platform rules to support platforms' complaints handling processes for misinformation and disinformation on their services. Rules could specify

minimum standards around the role of AI in complaints handling. Any future minimum standards for complaints handling may intersect with proposed guardrails that could require platforms to have human oversight of decision-making, communicate with end-users about AI enabled decisions and establish processes to challenge the outcomes of decisions.

- The ACMA could use record keeping rules under the Bill to require that sections of the digital platform industry make, retain and report data on how AI is being used to support the enforcement of platform policies to manage misinformation and disinformation. This may duplicate or crossover with records that platforms could need to maintain to allow compliance with guardrails to be assessed.

It may be challenging for these crossovers to be effectively managed by an AI-specific regulator (in contrast to an industry specific regulator), that does not have visibility of the activities being taken under the proposed regulatory framework.

News and Media

Amid concerns regarding the sustainability of public interest journalism in both Australia and abroad, generative AI promises many potential benefits for news production[2]. These tools can help journalists to scope, research and prepare stories, analyse and draw insights from vast quantities of data, and automate time-consuming processes, freeing up editorial resources. AI tools can also help news organisations better reach, target and personalise their content to their audience.

However, use of this technology raises legal, financial and ethical risks for media outlets, including, the impacts on journalistic workforce due to greater automation, and the potential loss of editorial quality and audience trust due to AI 'hallucinations', inaccuracies, or biases.[3]

More than half of Australians are uncomfortable with the idea of news produced mainly by AI (59%) – higher than the global average (45%),[4] and the Australian news media industry has been cautious around widespread implementation of new generative AI tools.[5] Nevertheless, most Australian news outlets are reportedly experimenting with the technology,[6] and many have, or are currently developing, policies and guidelines around its responsible use.[7] These policies and guidelines may include requirements requiring media outlets be transparent about their use of AI.

In addition to these internal policies, many Australian newsrooms are also required to adhere to existing regulatory safeguards that could help mitigate some of the potential harms from AI adoption. Broadcasting codes of practice include requirements for accuracy and impartiality in TV and radio news content, which are enforced by the ACMA and apply regardless of whether news content was created through AI tools. Additionally, the Australian Press Council's Statement of General Principles also requires online and print publications within its jurisdiction to be accurate.[8]

[2] Beckett, Charlie and Mira Yaseen (2023), Generating Change: A global survey of what news organisations are doing with AI, London School of Economics, Generating+Change+_+The+Journalism+AI+report+_+English.pdf (squarespace.com)

[3] Working Paper 2: Examination of technology – Large Language Models | The Digital Platform Regulators Forum (DP-REG)

[4] Park, Sora, Caroline Fisher, Kieran McGuinness, Jee Young Lee, Lerry McCallum, Xiaolan Cai (2024), Digital News Report: Australia 2024, News and Media Research Centre, University of Canberra, pg 24 apo-nid326816_4.pdf

[5] Attard, Monica, Michael Davis, Lisa Main (2023), Gen AI and Journalism, Centre for Media Transition, University of Technology Sydney, Australia, GEN AI and Journalism_web version 9 April 2024.pdf (uts.edu.au)

[6] See, for example, Buckingham-Jones, Sam (2024). 'ABC builds its own AI model', Australian Financial Review, 24 Jan

[7] Attard, Monica, Michael Davis, Lisa Main (2023), Gen AI and Journalism, Centre for Media Transition, University of Technology Sydney, Australia, pg. 49 GEN AI and Journalism_web version 9 April 2024.pdf (uts.edu.au)

[8] Statements of Principles - Australian Press Council

Many online Australian news services are not covered by the accuracy and impartiality rules outlined above.

Of interest to media companies is the reliance on their news material to train AI models. This potentially raises questions of attribution and compensation.

Given these issues, the adoption and impact of AI on media companies, newsrooms and news output is of ongoing interest to the ACMA and is one of several issues that could be examined in future iterations of the Media Diversity Measurement Framework.

Implementing mandatory guardrails in high-risk settings

While the ACMA supports guardrails for AI, we consider the introduction of mandatory guardrails should not add regulatory burden for industry participants where the objectives of the guardrails are already being achieved through compliance with other regulatory frameworks. Industry-specific regulators remain best placed to understand how the obligations under mandatory guardrails may apply in the sectors they regulate. We consider industry regulators remain best placed to implement guardrails in a way that complements, rather than duplicates, existing regulatory activities. We do not think this cohesive approach would be achieved under a single AI regulator or legislation that is developed for a specific technology.

A single AI specific Act, enforced by an 'AI regulator' will not have the depth of knowledge of high-risk industries or the harms they may present to effectively apply guardrails in a way that achieves the best possible outcomes for specific sectors. An AI specific Act would also likely struggle to keep up with the pace of change that AI developers and deployers are working in. This makes it even more critical that technologically-agnostic approaches – which are focused on harm mitigation – are central to how guardrails are implemented.

The ACMA supports a framework approach as the most suitable of the options presented for implementing mandatory guardrails. This approach would establish clear definitions to support individual regulators to implement the guardrails where the sectors they regulate constitute 'high-risk use cases'. It would also enable the ACMA, as a specialist regulator, to apply guardrails in a way that complements existing regulatory frameworks and activities.

While it is our preferred approach, modest steps are still likely to be needed to prepare existing regulators to support a framework approach to applying mandatory guardrails. This may include government leveraging the AI Advisory Body to support technical uplift for domain-specific regulators. It could also include government supporting efforts to improve regulatory cohesion by addressing barriers to information-sharing between the regulators that enforce guardrails in high-risk settings.

We welcome the opportunity to share our expertise across telecommunications, broadcasting as well as online content regulation, and our experience supporting regulatory coherence (via DP-REG) to help improve coordination and any proposed legislative design.

We also welcome further opportunities to engage with the Department as this important work continues.

Yours sincerely

Nerida O'Loughlin PSM
2 October 2024