UTS

CENTRE FOR MEDIA TRANSITION

Safe and responsible AI in Australia
Discussion Paper, June 2023

Submission from the Centre for Media Transition

Date: 8 August 2023

About the Centre for Media Transition

The Centre (CMT) was established in 2017 as an applied research unit based at the University of Technology Sydney (UTS). It is an interdisciplinary initiative of the Faculty of Arts and Social Sciences and the Faculty of Law, sitting at the intersection of media, journalism, technology, ethics, regulation and business.

Working with industry, academia, government and others, the CMT aims to understand media transition and digital disruption, with a view to recommending legal reform and other measures that promote the public interest. In addition, the CMT aims to assist news media to adapt for a digital environment, including by identifying potentially sustainable business models, develop suitable ethical and regulatory frameworks for a fast-changing digital ecosystem, foster quality journalism, and develop a diverse media environment that embraces local/regional, international and transnational issues and debate.

This submission was prepared by:

- Dr Michael Davis, Research Fellow, Centre for Media Transition, Faculty of Law, UTS

- Professor Monica Attard, Co-Director, Centre for Media Transition, Faculty of Arts and Social Sciences, UTS

- Professor Derek Wilding, Co-Director, Centre for Media Transition, Faculty of Law, UTS

CONTACT

Centre for Media Transition
Faculty of Law, University of Technology Sydney
Building 2, Level 15
UTS City Campus, Broadway
PO Box 123, Broadway NSW 2007

cmt@uts.edu.au
+61 2 9514 9669

cmt.uts.edu.au

Executive Summary

- The Centre for Media Transition is conducting research on the impacts of AI in the journalism industry. This submission focuses on risks arising from the use of AI in journalism and risks from AI to the information environment more broadly.

- AI has been used in newsrooms for more than a decade; however, generative AI is a game-changer, opening a much broader range of use cases and deeper risks.

- Quality journalism is critical to the strength and integrity of the public sphere. As news businesses integrate AI tools into their production processes, it is crucial that they appropriately manage any risks.

- Benefits of AI for journalism include increased efficiencies through automation of mundane tasks, potential benefits to sustainability, augmented capabilities for audience and data analysis, and as a research, verification and content-generation tool.

- AI poses risks to editorial processes, the production and sustainability of public-interest journalism, and to the integrity of the news and information environment.

- Shared accountability is key to managing and mitigating AI risk. Accountability should be assigned appropriately across the system lifecycle from development, through to deployment and use. Effective mechanisms for achieving accountability should be evaluated in the next stage of the consultation.

- Professional journalism is already held accountable via self- and co-regulatory frameworks. Although the potential applications and risks of AI in journalism are still emerging, the risks are not so great or different in kind that they cannot be managed largely via these existing frameworks.

- Maintaining a self- and co-regulatory approach is critical to the preservation of press freedom. Strong evidence of a high risk of severe harm arising from AI use would be needed were additional press regulation to be considered.

- The journalism industry should be encouraged to re-assess its codes and to develop AI-specific guidelines to ensure editorial processes are sufficiently robust to deal with AI risk. It should do this regularly as AI technology continues to develop.

- Risks posed by AI to the broader information environment go beyond journalism to implicate other actors including digital platforms, non-journalism users and AI developers. While news businesses should ultimately be responsible for their own content, responsibility for managing the risk of AI-generated or -propagated misinformation should encompass these other actors.

- AI developers and vendors could be required to certify their tools before they are released to market against a set of independent standards designed to minimise risk, including the risk of propagating misinformation, using or promulgating biased data or otherwise being used to pollute the information environment.

- Digital platform companies should be responsible for implementing safeguards against AI-assisted manipulation and for holding their users to account. They should also be held accountable more broadly for promoting a high-quality information ecosystem on their platforms.

- The government should consider the potential for AI systems to negatively affect the information environment alongside its current policy focus on mis- and disinformation and on digital platform regulation more generally. The potential impact of AI across the information landscape is of such a scale that a narrow or piecemeal approach is unlikely to adequately address the risks.

Introduction

Thank you for the opportunity to contribute to this consultation.

Whilst its future impact is uncertain, generative AI has wide-ranging potential to influence how journalism is made. Current generations of AI are already influencing how journalism is consumed. In Australia, there is already evidence that generative AI is being used to produce news articles. The Executive Chair of News Corp, Michael Miller, told the World News Media Congress in Taipei in July 2023 that a team of four staff use the technology to generate some three thousand stories each week on weather, fuel prices and traffic conditions. The ABC has put on the public record that it is evaluating how to use AI to enhance public interest journalism and make the national broadcasters more broadly accessible. Abroad, Germany's Bild newspaper, the largest tabloid in Europe, has announced a major cost-cutting program which will lead to the redundancy of 200 editorial staff due to the "opportunities of artificial intelligence." The BBC is using AI tools to streamline workflows. Reuters deploys a tool known as News Tracer to track breaking news on social media and publications in locations where Reuters has no presence. The New York Times uses its bespoke Editor tool to source and check research. The Washington Post uses a bespoke Knowledge Map to correlate information gleaned in research. It has also recently announced AI as a 'priority opportunity' and has established two teams to drive innovation in AI. No consideration of risk was apparent in the announcement.[1] Meanwhile, Wired has moved to establish internal guidelines on AI after receiving criticism for publishing a series of AI-generated articles that were riddled with errors and later retracted.[2]

Nonetheless, the application of generative AI in Australian newsrooms seems to be proceeding cautiously. An optimistic view is that newsroom managers are conscious of the risks of misuse, ungoverned use or use without human oversight, as well as job redundancies. However, even the perceived benefits of automating the more laborious tasks of journalism entailed in weather reporting, fuel pricing reports and traffic conditions should be treated with caution and assessed against the role these menial tasks have played in the training of journalists and in their capacity to progress from such beats to more fundamental public interest journalism. The optimists assert that the use of generative AI and more complex ADM tools will not result in job redundancies, rather that these new technologies will have to capacity to transform newsrooms from their current linear production lines into "networked information and engagement hubs that give journalists the ability to move more efficiently into the data driven age."[3] Whether this transformation is achievable and how it takes place is in the realm of the unknown.

One of the principal concerns arising from the latest generation of AI tools, as noted in the consultation documents, is the risk of facilitating the propagation of mis- and disinformation through inaccurate and biased outputs from generative AI systems, as well as the hugely increased capacity to produce content at scale. One side of this problem is the potential proliferation of low-quality information on the internet. Another is the potential impact on the production of high-quality information, principally via the use of AI tools or AI-generated information in the journalism industry. In our view this part of the problem has been neglected in discussions to date, and that is our focus in this submission. Australian news businesses are the principal source of high-quality, reliable news in Australia and contribute heavily to the strength and integrity of the public sphere. As these businesses inevitably seize the opportunity to integrate AI tools into their production processes, it is crucial that

---

[1] F Ryan, 'Announcement from Publisher and CEO Fred Ryan on The Post's next steps in AI innovation', in *Washington Post*, , 24 May 2023, <https://www.washingtonpost.com/pr/2023/05/24/announcement-publisher-ceo-fred-ryan-posts-next-steps-ai-innovation/> [accessed 4 August 2023].

[2] WIRED, 'How WIRED Will Use Generative AI Tools', in *WIRED*, <https://www.wired.com/about/generative-ai-policy/> [accessed 3 July 2023].

[3] C Beckett, *New powers, new responsibilities. A global survey of journalism and artificial intelligence*, London, London School of Economics and Political Science, 18 November 2019, <https://blogs.lse.ac.uk/polis/2019/11/18/new-powers-new-responsibilities/> [accessed 22 June 2023].

they appropriately manage any risks that this poses to editorial processes, the production and sustainability of public-interest journalism, and to the integrity of the news and information environment.

The management of AI risk in journalism presents several challenges that do not arise in all industries. These stem from the role that journalism plays in democratic society. The many public benefits that flow from public-interest journalism were documented in the ACCC's Digital Platforms Inquiry (DPI) final report, which states that journalism 'is an important contributor to the "public sphere", democracy and the economy, and has a place within much broader constructs of societal communication and debate.' Journalism shares the function of promoting transparency and public accountability with other institutions such as law enforcement and administrative review mechanisms, but importantly, it is independent from them. Due to its independence and 'the professional expertise exercised in its production, journalism can fulfil these public interest functions in a unique and significant way.' Ultimately, the report states, journalism can promote the public interest by 'providing a volume, range and depth of information and analysis that would not otherwise be readily available. Therefore, the public also has an interest in upholding professional journalistic standards of accountability, accuracy and ethical conduct, both internal and external to media organisations.'[4]

Because the public has an interest in upholding professional journalistic standards, it has an interest in protecting public-interest journalism from any risks arising from AI. On the other hand, the public interest may also be served by any benefits that may derive from efficiency gains in the use of AI, because the sustainability of public-interest journalism is itself a matter of public interest. Given the potential impact on the public interest, the relative weight that should be placed on the risks and benefits of AI applications in journalism, and the question of how the risks should be managed, goes beyond the risk-benefit calculations of individual news businesses and instead must be treated as a public-policy concern.

This area requires extensive consideration and research. To this end, the Centre for Media Transition is currently undertaking a study to assess the current use of generative AI tools in Australian newsrooms, as well as emerging use cases and the management of risks.

In what follows, we first outline how AI is being employed in newsrooms and consider other potential use cases. We then set out some potential risks. These include risks to editorial practice and journalistic ethics, risks to the journalism industry, and risks to the broader information environment. We then consider how these risks might best be managed both by industry and through regulatory intervention.

Ultimately, we consider that existing self-regulatory mechanisms, with some adjustment, are likely to be the best means of managing AI risks in journalism. However, given the public interest that is at stake, it is important that these are supported by general accountability and transparency requirements that apply to businesses along the full AI 'value chain' including developers, vendors, and end-users.

Current approaches to AI in journalism

News journalists have been engaged with AI in many different aspects of the news workflow for more than a decade, even if these technologies are not always explicitly recognised as AI. For example, AI is used to create consumer profiles for the subscription models which dominate the newspaper landscape. AI is used to transcribe interviews, permitting journalists more time to do the heavy lifting of analysis. AI is also used, for

---

[4] Australian Competition and Consumer Commission, *ACCC Digital Platforms Inquiry Final Report*, 2019, p. 284, <https://www.accc.gov.au/about-us/publications/digital-platforms-inquiry-final-report>.

example by News Corp in Australia, to assist journalists to tailor their stories to user profiles.[5]

The London School of Economics (LSE) surveyed 71 news organisations from 32 countries in 2019 to discern how pre-generative AI technologies were being used and to understand how these newsrooms were preparing themselves for generative AI. Whilst the survey did not sample poorer, undeveloped regions, and was skewed towards the USA and Europe, the views of the editorial leadership, innovation specialists, data experts, and audience and marketing workers gave an insight into how the technology is perceived and how strategies were being developed to govern its use.[6] Broadly, the surveyed newsrooms identified six key areas where AI might make a difference to their operation and where editorial and ethical issues arise:

- economic savings

- algorithmic bias from the use of bad data

- misinformation (unless AI is used to counter information disorder)

- mistakes or 'hallucinations' which may creep into news copy if human intervention is eliminated

- changing editorial decision-making to reduce established biases in story selection

- augmenting the influence of 'big tech'.

The LSE concluded that newsrooms generally identified the need for AI literacy and skills, faster innovation, renewed ethical guidelines and better editorial processes. It also identified that to meet the competitive pressures that will arise as AI is more generally adopted across the sector, old Chinese walls between editorial and commercial would need to fall away and greater collaboration between journalism outlets and researchers – both inside those larger newsrooms which have research capacity and with academia – would need to occur.

A 2023 international survey found that half of the world's newsrooms are already using generative AI. However, its use is still infrequent, with less than 15% of journalists using it weekly or more. Content summaries are the most common use (54%), but 32% said that AI was being used for article creation and 44% for research. The survey also gauged perceptions of risk. Inaccuracy and quality of news were a major concern for 85% of respondents, with 67% concerned about copyright, 46% about data protection and privacy, and 38% about job security. However, 82% thought that their roles were likely to change with the adoption of AI. Despite the rapid adoption of generative AI, the survey found that only 20% of newsrooms have internal guidelines in place to govern its use.[7]

Reporters in the US are being permitted to use generative AI as an 'editorial assistant'.[8] This is to say they can use tools such as ChatGPT to locate facts, rewrite sentences where the reporters encounter blockages in the creative process, trawl documents for supporting data and synthesise information to create new news stories in an iterative process. The risks associated with this usage are many: information used to train AI systems may not be broad enough in scope or even correct. AI might throw out sentences written by humans, which places the journalist at risk of plagiarising. Synthesised information requires all sources to be scoured whilst AI systems may be only narrowly trained or subject to training cutoff dates. Large language models, on the other hand, may be trained on extremely large datasets comprising information of varying accuracy. LLMs' probability-derived outputs may also contain inaccuracies or biased information. Newsroom workflows are notoriously complex and story-dependent; as a result, any streamlining processes need to account for

---

[5] A Meade, 'News Corp using AI to produce 3,000 Australian local news stories a week', in *The Guardian*, 31 July 2023, <https://www.theguardian.com/media/2023/aug/01/news-corp-ai-chat-gpt-stories> [accessed 3 August 2023].
[6] Beckett.
[7] D Roper et al., *Gauging Generative AI's impact on newsrooms*, World Association of News Publishers (WAN-IFRA), 2023.
[8] J Herrman, 'How AI Will Change the News Business: 3 Theories', in *Intelligencer*, 1 August 2023, <https://nymag.com/intelligencer/2023/08/how-ai-will-change-the-news-business.html> [accessed 4 August 2023].

the multitude of variations of fact, story angle and verification to ensure newsrooms are robust purveyors of truth or the closest approximation of the truth.

Similarly, Australian newsrooms are beginning to experiment with more sophisticated models of AI, but the extent to which this is occurring, and in which areas, requires research. Importantly, adjustments to editorial guardrails and ethical guidelines are not yet apparent, though a cautionary note has been published by the Media Entertainment and Arts Alliance.[9]

Many challenges stem from algorithmic bias. Australian news organisations are known to be investigating ways in which such bias can be eliminated. However, this will require developers to share algorithms and source material. Given the reluctance of social media platforms to share such information, it seems apt to assume there will be similar reluctance on the part of AI developers, particularly once newsrooms become reliant on the technology for the production of news.

AI developers and vendors are both actively targeting their products at newsrooms and entering agreements with news businesses for access to their news story archives. In July, Associated Press agreed to license part of their text archive to OpenAI, the manufacturer of ChatGPT, for training purposes.[10] OpenAI is also providing substantial research funds in software-use credits to the American Journalism Project, an NGO which supports small newsrooms in the USA.[11]

In part, these are welcome moves. Developers secure access to high-quality information for AI training, and news businesses receive remuneration or other financial support. Yet these benefits must be weighed against potential risks. For example, the licensing of particular news sources over others creates a new avenue for consolidating US-centric views in training data that may entrench bias against perspectives from other locations, especially the global south. These biases can find their way into Australian news when local newsrooms use foreign news outlets as sources for their own output.

Similarly, agreements with small and underfunded newsrooms, whether in the US, Australia or developing nations, create more alarm bells than solace. There is a substantial risk of creating a reliance in small newsrooms on proprietary AI technology, funded by the developer itself. Google's moves in the US to roll out its Genesis assistant,[12] which promises to take in information about current events and generate news items in response, should be seen in a similar light: generously self-serving and inherently dangerous to the methodologies usually employed by public interest journalism. Whilst there is no evidence that OpenAI or Google are currently offering similar 'assistance' to Australian newsrooms, we can assume this will occur in some form.

As a result, Australian newsrooms need to be alert, if not alarmed, to the risks brought by AI and the need to protect both their own interests and those of the public. In deploying AI in the pursuit of efficiency and other benefits, they need to ensure they do not become reliant on the help being offered by the manufacturers or licensees of the AI tools they intend to deploy. While licensing agreements are welcome for newsrooms, they should also be viewed with some scepticism, as perhaps in part a means for forestalling copyright reform or other regulation such as a bargaining code, or to license public-interest content for exclusive use. Larger news businesses are already alert to the risk that generative AI tools may siphon news content – and therefore advertising dollars – away from their own webpages to search engines and other digital platforms. Newsrooms must also be vigilant

[9] Media, Entertainment & Arts Alliance, 'Caution and consultation needed as AI is rolled out in media and entertainment', 22 June 2023, <https://www.meaa.org/mediaroom/caution-and-consultation-needed-as-ai-is-rolled-out-in-media-and-entertainment/> [accessed 28 June 2023].
[10] M O'Brien, 'ChatGPT-maker OpenAI signs deal with AP to license news stories', in *AP News*, 14 July 2023, <https://apnews.com/article/openai-chatgpt-associated-press-ap-f86f84c5bcc2f3b98074b38521f5f75a> [accessed 4 August 2023].
[11] H Tameez, 'OpenAI will give local news millions to experiment with AI', in *Nieman Journalism Lab*, 2023, <https://www.niemanlab.org/2023/07/openai-will-give-local-news-millions-to-experiment-with-ai/> [accessed 20 July 2023].
[12] Herrman.

in protecting the integrity of their product through the robust review of editorial guardrails and ethical guidelines.

From the above discussion it is apparent that there are many parallels between the circumstances which triggered the DPI and those currently emerging with the advent of generative AI. There is certainly value for news businesses in using AI tools. At the same time, there are many risks to journalism as an industry and to its product of public-interest news. The importance of these for the public sphere means it is crucial to manage and mitigate these risks.

Opportunities for AI in journalism

Our preliminary research has identified a range of opportunities and benefits for AI in journalism:

- Automating the production of news which is iterative, repetitive or mundane, thus freeing journalists to pursue investigative research projects of high public interest value.

- Augmenting audience analysis and identifying stories of relevance to audiences to maximise engagement and promote loyalty to a particular news brand.

- When correctly trained, a reduction in biased interpretations of data. This bias reduction process when performed manually can be extremely time consuming and is open to mistakes.

- AI can also codify data more efficiently than humans such that journalists can use AI to spot trends more easily than can be done manually.

- In the sphere of information disorder, properly trained AI can examine large databases to identify false information and misinformation.

- Continued use in the transcribing of interviews as well as audio and visual content allowing journalists to focus more fully on providing analysis and insight.

- Overall increases in efficiency.

- Other advantages and opportunities may emerge in the future. In the vision for the future of journalism outlined by the LSE, in which linear identification and production of news content is replaced by networked approaches, the hope is that editorial managers and proprietors will invest in research to successfully identify and manage risk. It is important that this research occurs within a framework of expectations for the protection of information in the public sphere as set out by regulation.

Risks of AI in journalism

Our preliminary research indicates that there are a number of risks which require consideration and further research. We have categorised these risks into the following areas:

1. Risks to editorial practice and ethics

- Risks to the accuracy of news arising from AI output errors and difficulties in verification and authentication due to unidentified provenance.

- Risks to impartiality and discrimination protections from biased data.

- Copyright issues arising from both the input and output of generative AI systems.

- Risks of defamation or violation of privacy due to the use of personal information in training data or inputs.

- Risk of unintended plagiarism due to uncertain provenance of AI-generated information.

- Threats to editorial independence and control arising from the use of proprietary models and uncertain provenance.

2. Risks to the journalism industry

- Impacts on / reduced demand for editorial, design and other staff.

- Implications for news business models including a potential increase in 'content farms' and news recyclers.

- Loss of trust if AI generated news contains inaccuracies or if insufficient transparency is attached, such that consumers are unaware that the news item they are reading is AI generated.

- 'Walled garden' problem with a loss of audience to digital platforms due to incorporation of AI in search.

- Loss of expertise or of editorial attention to particular areas with certain low demand beats left to AI.

- Drop in job satisfaction and a rise in low-paid verification work.

- Adoption of proprietary systems that lack transparency and impose terms of use on small newsrooms, which may include using risk-mitigation frameworks developed by the AI companies themselves.

3. Risks to the information environment

- Risk of the proliferation of misinformation and fake or low-quality news produced or distributed by AI tools.

- Potential to exacerbate the trust crisis by reducing transparency, objectivity and accuracy of news.

- Potential to entrench mainstream audience preferences.

- Threats to original news and creativity where AI is used to spot and piggyback on trends. This may lead to an increase in 'cookie cutter' news, stenography in place of critical journalism, and algorithm-driven content in place of news.

Managing AI risks

The importance of shared accountability

As set out in the NSTC's Rapid Response Information Report, the extent to which AI risks are realised or mitigated will depend on the actions of multiple stakeholders including government, industry, developers and consumers. Trust will be critical, and trustworthiness will be built through 'appropriate levels of reliability, transparency, accountability and legal, policy and other safeguards.'[13]

In our view, the key to managing and mitigating risk is accountability. It is accountability that will ensure that developers build reliable systems that minimise potential risks including error and bias. It is accountability that will ensure that those who use AI systems do so in fair, equitable and appropriate ways, and that privacy, security and safety are maintained. And it is accountability that will ensure that potential risks to individuals, society and the environment through the use of AI are acknowledged and addressed. Transparency is

---

[13] G Bell et al., *Rapid Response Information Report: Generative AI - language models (LLMs) and multimodal foundation models (MFMs)*, Australian Council of Learned Academies, 24 March 2023.

important for trust, but it is especially so when all parties in the AI value chain or system lifecycle are accountable for their actions. As noted by the Human Technology Institute, 'the opacity and flexibility of generative AI systems pose challenges to accountability. Regulators and corporate policies should carefully attribute legal liability for harms or errors to entities across the AI value chain and incentivise the creation of effective safeguards at the most effective and appropriate points.'[14] Similarly, accountability is one of the eight principles in Australia's AI Ethics Framework, which states, 'People responsible for the different phases of the AI system lifecycle should be identifiable and accountable for the outcomes of the AI systems, and human oversight of AI systems should be enabled.'[15]

A significant problem in this respect is that AI will be (is already) so pervasive that it will not be practical or even possible for every use of AI to be transparent. For example, it may not be possible to identify every instance of AI-generated material on the internet (or even in hardcopy). This is made more difficult by the sheer variety of tools that come under the AI umbrella. For this reason, a risk-based approach is likely to be appropriate. Further, industry users – including journalists and news businesses – should be held accountable for their use of AI, and AI developers and vendors should be accountable for ensuring that the capabilities and limitations of their products are transparent to users.

If we look at AI applications in the news media industry, accountability should be assigned appropriately across the system lifecycle from development, through to deployment and use. To the extent that AI risks are transparent and predictable, newsrooms must ensure that their internal governance processes are sufficiently robust to minimise those risks. To the extent that AI risks are not transparent or predictable, an appropriate degree of caution should be applied in the use of AI systems. Importantly, if the news industry is to successfully manage and mitigate those aspects of AI risk that are under its control, it is critical that the risks are transparent to the industry, and that AI developers and vendors are accountable for making these risks transparent to industry. This includes ensuring that users are aware of when they are using AI tools or sourcing information that has been generated by AI.

A system of accountability could be usefully built around the elements presented in attachment C to the discussion paper. These can be adapted as necessary to different points of accountability in the value chain or system lifecycle. Effective mechanisms for achieving accountability across the system lifecycle should be evaluated in the next stage of the consultation.

News media regulation

News media businesses are subject to a range of laws, including defamation, privacy and copyright laws, that intersect with several of the AI risks set out in the previous section. Though these intersections have largely not been tested, they may require increased vigilance from news media businesses to avoid inadvertently incurring liability. For example, news publishers may risk incurring liability where AI is used to generate a story that inadvertently contains defamatory statements about a person.

However, our focus here is rather on the risks to the integrity of journalism practice and to the broader information environment than on these legal risks. In this area too, the news industry is subject to regulation, albeit largely in the form of self- or co-regulatory codes that govern journalism practice via complaints-based accountability mechanisms.

Although the potential applications and risks of AI in journalism are still emerging, and further research needs to be undertaken on the robustness of current safeguards, our

---

[14] L Solomon & D Nicholas, *The State of AI Governance in Australia*, Human Technology Institute, University of Technology Sydney, 31 May 2023, , p. 24, <https://www.uts.edu.au/sites/default/files/article/downloads/HTI%20The%20State%20of%20AI%20Governance%20in%20Australia%20-%2031%20May%202023%20.pdf> [accessed 14 July 2023].

[15] Department of Industry, Science and Resources, Australian Government, *Safe and responsible AI in Australia*, , June 2023, , p. 14, <https://storage.googleapis.com/converlens-au-industry/industry/p/prj2452c8e24d7a400c72429/public_assets/Safe-and-responsible-AI-in-Australia.pdf> [accessed 5 June 2023].

current assessment is that the risks are not so great or different in kind that they cannot be managed largely via these existing self- and co-regulatory frameworks. For example, all journalism codes currently operating in Australia require news to be accurate and inaccuracies to be corrected. Material produced by AI is subject to these same constraints.

This is not to say that journalism codes as they stand provide adequate protection against AI risks; our point is only that self- and co-regulatory schemes are likely to prove adequate. It may be argued that the complaints-based accountability system, supported in the case of broadcasting by ACMA's enforcement powers but completely self-regulatory in the case of the press, may not provide the strongest protection against careless or unethical journalistic practices. But this is a point that would apply generally rather than specifically to AI. Furthermore, maintaining a self- and co-regulatory approach is critical to the preservation of press freedom. Strong evidence of a high risk of severe harm arising from AI use would be needed were additional press regulation to be considered.

Though the self- and co-regulatory system may be adequate on the whole, industry bodies should be encouraged to re-assess their codes and to develop AI-specific guidelines to ensure their editorial processes are sufficiently robust to deal with AI risk. They should do this regularly as AI technology continues to develop. This point no doubt applies to many industries, but it is critical in the case of journalism, given its importance to the public sphere.

One area where existing codes may be inadequate is the lack of an obligation to be transparent about the publication of AI-generated content. This will be critical for preserving audience trust, particularly as the use of AI tools expands beyond data-heavy reporting such as weather, finance and sports results.

Other areas of risk are more difficult to assess. This includes the risk of bias in news arising from deficiencies in AI training data or probability-driven outputs. A related problem is a potentially increased risk of misrepresentation of viewpoints, persons or groups, or of discrimination and vilification, through the use of AI-generated text, images or videos. And to address the risk of AI-generated inaccuracies, newsrooms will need to ensure appropriate safeguards are in place, such as AI-specific verification or fact-checking procedures that include human oversight.

Despite these concerns, the principles included in journalism codes are likely to be appropriate for the most part – what will be critical is that newsrooms ensure these principles are front of mind when they use AI tools in the production process. Targeted AI guidelines are likely to help ensure this is the case, but if risks are to be managed, users will need to be aware of when they are using AI tools or using information generated by AI tools but sourced elsewhere. They will also need to be aware of which tools they are using, as well as the capabilities and limitations of the tools.

Addressing risks to the information environment

Current problems with the information environment, including the proliferation of misinformation and other forms of information disorder, are well documented. Professional journalism plays a critical role in countering information disorder and ensuring the public can find quality news and information on matters of public interest.

However, any regulatory approach to mitigate AI risk to the broader information environment must be at least partly located outside the journalism industry since it implicates other actors including digital platforms, non-journalism users and AI developers. While news businesses should ultimately be responsible for their own content, responsibility for managing the risk of AI-generated or -propagated misinformation should encompass these other actors.

This could be achieved in part by requiring AI developers and vendors to certify their tools before they are released to market against a set of independent standards designed to minimise risk, including the risk of propagating misinformation, using or promulgating

biased data or otherwise being used to pollute the information environment. Any proprietary tools developed or customised for news businesses should also be required to comply with these general standards.

Digital platform companies should be responsible for implementing safeguards against AI-assisted manipulation and for holding their users to account. They should also be held accountable more broadly for promoting a high-quality information ecosystem on their platforms. In our view, the government should consider the potential for AI systems to negatively affect the information environment alongside its current policy focus on mis- and disinformation and on digital platform regulation more generally. The potential impact of AI across the information landscape is of such a scale that a narrow or piecemeal approach is unlikely to adequately address the risks.