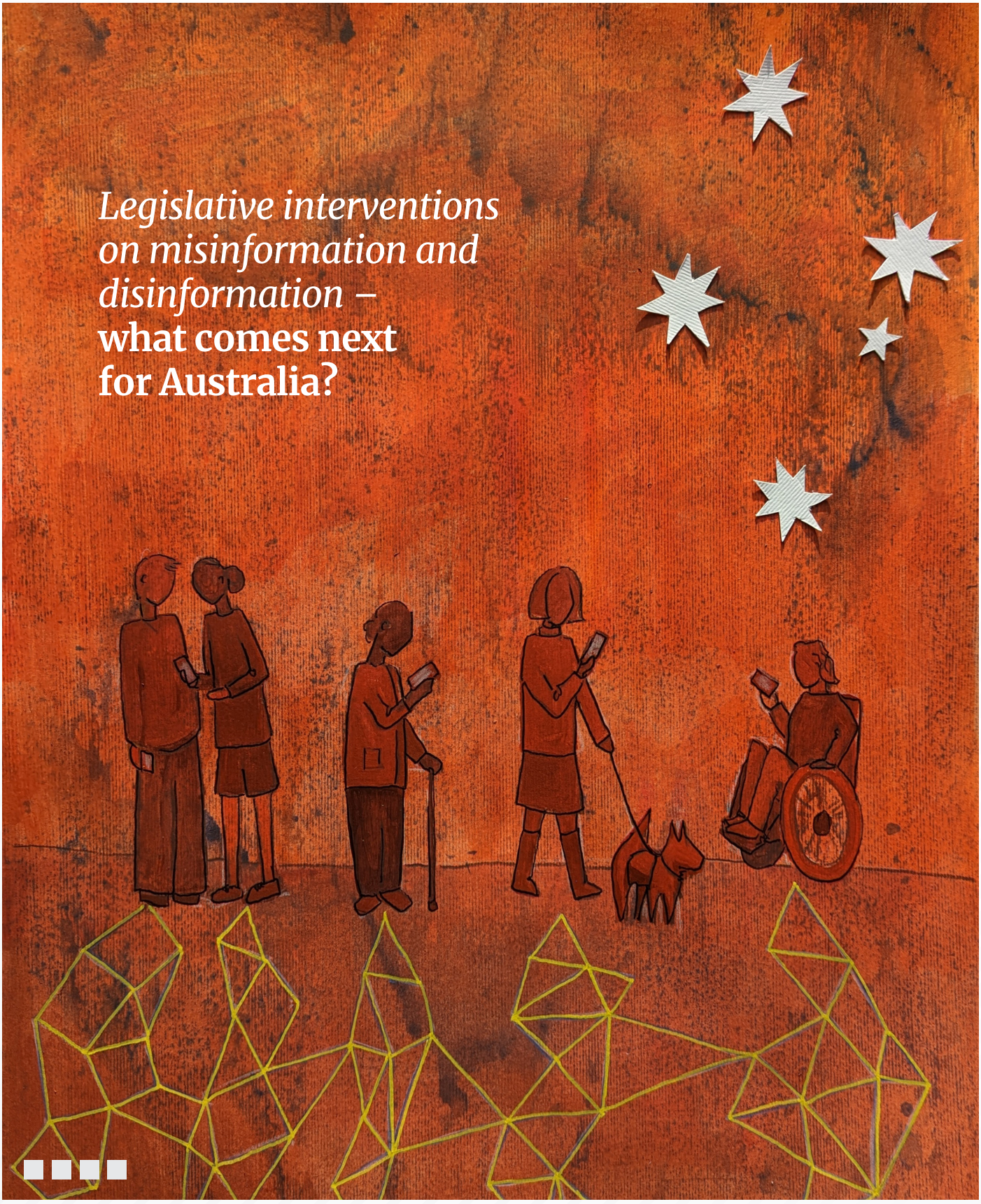
Reset. AUSTRALIA

MONASH University

Human Rights Law Centre

October 2023 | Policy briefing

Legislative interventions on misinformation and disinformation – what comes next for Australia?



Summary

This report summarises a policy roundtable held with 15 experts, unpacking legislative and regulatory interventions around misinformation and disinformation in the context of the *Combatting Misinformation and Disinformation Bill* (the Bill).

The discussion noted issues with the existing regulatory framework, based on Digi's *Australian Code of Practice on Disinformation and Misinformation* (the Code). The existing framework does not adequately create accountability, transparency nor comply with human rights principles. Fundamentally, because the Code is drafted and overseen by industry, there are insufficient incentives to improve practice.

The Bill requires meaningful adjustment in order to achieve its admirable ambition. This paper recommends:

- Reframing the Bill, focusing on how the Bill will enhance public oversight. This includes oversight over any measures deployed by social media platforms that may affect freedom of speech.
- Amending the Bill to:
- Require proactive risk assessments for larger platforms that include consideration of human rights. These could be Australian versions of the risk assessment requirements that are already produced under the EU's *Digital Services Act*, to reduce regulatory burden.
- Require larger platforms to publish routine transparency data, to be set by Ministerial discretion and without Australian Communications and Media Authority (ACMA) requests needing to be made. This would help improve both public trust and transparency, as well as reduce the burden on ACMA's investigative team.
- Require researcher access to public interest data, enabling independent researchers to request relevant data from platforms. These requirements could mimic requirements established under the EU's *Digital Services Act* (DSA), which means large platforms would not have to establish new systems to comply.
- Empower the ACMA to intervene and substitute the Code with a regulatory standard before a 'total failure' of the Code occurs. Where substantial deficiencies are evident, ACMA should be able to act.
- Considering the implications of the existing *Privacy Act* Review and upcoming *Online Safety Act* Review, especially with regard to consumer opt-outs and risk assessments.
- Ensuring ACMA is adequately resourced to deliver effective changes.
- Ensuring ACMA has adequate powers to deliver accountability, especially through the levels of civic penalties resulting from breaches.

Table Of Contents

Introduction

Australia's current approach to managing misinformation and disinformation rests on a belief that co- and self-regulatory mechanisms work. For example, the *Australian Code of Practice on Disinformation and Misinformation* – written by and overseen by industry group Digi – governs platforms' responses to misinformation and disinformation content. Other aspects, such as the management of political advertising, is entirely self-regulatory. Even the proposed *Combatting Misinformation and Disinformation Bill*, the Commonwealth Government's attempt to develop a more rigorous solution, rests on the assumption that these self-regulatory practices are working.

Evidence collected in the lead-up to the Voice referendum highlighted systemic failures in the scope and efficacy of the Digi Code.[1] Despite various policies and processes, digital platforms fail to adequately mitigate the spread of electoral misinformation, both in 'feeds' and paid-for advertising, leaving a risky situation for public trust. Where trust in democracy is eroded, civil and political rights face risks. The potentially severe consequences of these failures raises serious questions about the impact of this "light touch", co-regulatory approach.[2] There is clearly a pressing and urgent need for the Government to strengthen requirements for platforms to effectively address misinformation and disinformation on their platforms.

Critically evaluating Australia's approach to misinformation and disinformation is timely. As the Government reckons with the next steps for the *Combatting Misinformation and Disinformation Bill*,[3] there are relevant policy development processes running with the expedited *Online Safety Act* Review and the ongoing *Privacy Act* Review. Australia is somewhat unique in splitting safety and misinformation into separate legislative agendas,[4] and without some focus on systemic risks regarding misinformation, this separation could fail to effectively mitigate risks.

Comprehensive systemic regulation that places duties of care on platforms to mitigate risks, both individual and societal, might be the best path forward. The *Combatting Misinformation and Disinformation Bill* provides a unique vehicle to embed some of these systemic principles and comprehensive protections. Against this backdrop, Reset.Tech Australia convened a policy roundtable of 15 experts with expertise across digital regulation, platform transparency and human rights. The discussion is summarised below.

---

1. Reset.Tech Australia 2023 *How do platforms respond to user-reports of electoral process misinformation? An experimental evaluation from the lead-up to Australia's referendum* https://au.reset.tech/uploads/Reset.Tech-Electoral-Misinformation-Report.pdf,

Reset.Tech Australia 2023 *How do platforms handle electoral misinformation in paid-for advertising? An experimental evaluation using the Voice referendum* https://au.reset.tech/uploads/Reset.Tech-Australia-Paid-For-Advertising.pdf, Reset.Tech Australia 2023 *How do platforms' recommender systems promote political content? An experimental investigation using the Voice referendum* https://au.reset.tech/news/report-recommender-systems-and-political-content/ *Reset.Tech Australia 2023 Is content over- or under-moderated in the Voice referendum debate? An experimental evaluation* https://au.reset.tech/uploads/Reset.Tech-Over-Under-Moderation-2.pdf.

2. Reset.Tech Australia 2022 *How outdated approaches to regulation harm children and young people* https://au.reset.tech/news/how-outdated-approaches-to-regulation-harm-children-and-young-people-and-why-australia-urgently-needs-to-pivot/

3. Exposure Draft Communications Legislation Amendment (Combatting Misinformation and Disinformation) Bill 2023 https://www.infrastructure.gov.au/sites/default/files/documents/communications-legislation-amendment-combatting-misinformation-and-disinformation-bill2023-june2023.pdf

4. Other approaches, such as the EU's *Digital Services Act*, take a systemic, comprehensive approach to tackling risks in general, including individual safety and societal safety like misinformation and disinformation. Canada too, which is in the process of redrafting a proposed online safety bill, has moved away from an initial proposed bill focused solely on content and individual harms to a more systemic approach that aims to capture societal harms such as misinformation and disinformation. The UK's *Online Safety Bill* also catalyses the UK's process to tackle misinformation through their regulator, Ofcom.

The existing framework is flawed

Australia's social media information ecosystem is governed by a voluntary, industry-led Code, the *Australian Code of Practice on Disinformation and Misinformation* (the Code).[5] The Code documents how platforms should address disinformation and misinformation, by broadly suggesting that platforms must have policies and processes in place to address misinformation and disinformation. It is written and overseen by an industry representative group, Digi.

Current frameworks do not create adequate accountability

The Code does not create effective improvements for Australians. Commitments under the Code are poorly realised in practice. For example, recent research tested the efficacy of a number of systems on TikTok, X (formerly Twitter) and Facebook, as outlined in the platforms' own policies and in accordance with the Code. It found:

- **Evidence of insufficient moderation of misinformation and disinformation:** Platforms' content moderation processes were not effective at removing electoral misinformation.[6]
- **Evidence of insufficient advertising moderation:** Platforms' advertising approval systems were not sufficiently sensitive to detect and prevent paid-for electoral misinformation in advertising.[7]
- **Evidence of insufficient transparency reports:** Platforms are required to prepare annual transparency reports on the efficacy of their misinformation and disinformation efforts. None of the issues documented in recent research appeared in the reports.[8]

Platforms do not always realise their commitments under the Code, and there are limited mechanisms to require compliance in a timely fashion. For example, X removed its user-reporting channel for electoral misinformation a few weeks before the Voice referendum in Australia, despite being a signatory to relevant requirements in the Code. X were contacted directly, and a complaint was made to Digi, but this was not resolved before the referendum, leaving Australian voters in a vulnerable position. [9]

Further, the Code is not comprehensive enough to provide the accountability required. For example, important aspects of platforms' systems are not addressed by the Code, such as content recommender systems. Other important issues, such as political advertising, are beyond the scope of the Code.

---

5. Digi 2022 *Australian Code of Practice on Disinformation and Misinformation* https://digi.org.au/wp-content/uploads/2022/12/Australian-Code-ofPractice-on-Disinformation-and-Misinformation-FINAL-_-December-22-2022.docx.pdf

6. Reset.Tech Australia 2023 *Is content over- or under-moderated in the Voice referendum debate? An experimental evaluation* https://au.reset.tech/uploads/Reset.Tech-Over-Under-Moderation-2.pdf

7. Reset.Tech Australia 2023 *How do platforms handle electoral misinformation in paid-for advertising? An experimental evaluation using the Voice referendum* https://au.reset.tech/uploads/Reset.Tech-Australia-Paid-For-Advertising.pdf

8. See ibid, and, Reset.Tech Australia 2023 *Is content over- or under-moderated in the Voice referendum debate? An experimental evaluation* https://au.reset.tech/uploads/Reset.Tech-Over-Under-Moderation-2.pdf

9. Byron Kaye 2023 'Musk's X disabled feature for reporting electoral misinformation - researcher', *Reuters* https://www.reuters.com/technology/musks-x-disabled-feature-reporting-electoral-misinformation-researcher-2023-09-27/ Josh Taylor 2023 'X/Twitter scraps feature letting users report misleading information' *The Guardian* https://www.theguardian.com/technology/2023/sep/27/xtwitter-scraps-function-letting-users-report-misleading-information.

Current frameworks do not create effective transparency

There is a global problem of deteriorating transparency from platforms, against which the Code provides ineffective protection. In recent years, platforms have either closed or paywalled their key transparency tools, such as APIs. Researchers and regulators now face significant challenges to observe and interpret important outcomes emerging within platforms. For example, platforms have stopped providing information about state-backed operations,[10] alongside removing or reducing their API functionalities.

X have placed their API behind a prohibitively expensive paywall,[11] and removed information about state-backed operations. Facebook's API—Crowdtangle—is being sunsetted.[12] Notably, Facebook is currently the only platform still making state-backed authenticators. TikTok's Research API is currently only available to American and European academics,[13] and does not make attributions about state-backed actors. This makes it incredibly difficult for regulators and researchers to make attributions about foreign state-backed misinformation and disinformation. Recent research, for example, around Spamouflage activities backed by the Chinese Government,[14] are no longer able to be attributed.

Many of the moves to paywall APIs, from X to Reddit, may be driven by commercial considerations around the use of data to train AI, rather than to restrict access for public interest researchers.[15] Regardless of the intent, paywalling creates a significant financial barrier, limiting independent and public interest oversight.

---

10. See e.g. Australian Strategic Policy Institute 2023 Feedback on Communications Legislation Amendment (Combatting Misinformation and Disinformation) Bill 2023 Exposure Draft https://ad-aspi.s3.ap-southeast-2.amazonaws.com/2023-09/Ryan%20Zhang%20Aug%2023.pdf?VersionId=wbruzS6xDoNhqNgUqYEm5v8L4sKE3afl

11. Justin Hendrix 2023 'Twitter API Changes Set to Disrupt Public Interest Research', Tech Policy Press, https://techpolicy.press/twitter-api-changes-set-to-disrupt-public-interest-research/

12. Richard Lawler 2022 'Meta reportedly plans to shut down CrowdTangle, its tool that tracks popular social media posts', The Verge https://www.theverge.com/2022/6/23/23180357/meta-crowdtangle-shut-down-facebook-misinformation-viral-news-tracker

13. TikTok nd Research API https://developers.tiktok.com/products/research-api/

14. Australian Strategic Policy Institute 2023, 'Gaming public opinion: The CCP's increasingly sophisticated cyber-enabled influence operations', *Policy Brief* 71/2023 https://ad-aspi.s3.ap-southeast-2.amazonaws.com/2023-05/Gaming%20public%20opinion.pdf?VersionId=QYkBIWncbBU0E1KAhg9mX3TD7kwlWcWj

15. There was some discussion about the possible role of tax deductions on the donations of data to researchers, which may work in the US context.

The Code may contain some (optional) language around platform collaboration with academic researchers, but empirical evidence suggests otherwise. External scrutiny is often actively undermined by the digital platforms, and it is an increasingly risky environment for independent researchers. For example, digital platforms are engaging in 'lawfare' against researchers, such as when X launched an action against the Centre for Countering Digital Hate for engaging in common research practices.[16] Researchers have experienced first hand harassment from platforms and their legal teams. In some cases, this has left people reconsidering their ability to work in Australia, or undertake research on Australian issues due to insufficient protections against corporate harassment—some even consider relocating to other countries for these necessary protections[17] (different legal contexts can be either protective or weaponised). Currently, the Australian tech accountability sector is unprepared and underprotected for these emerging threats.[18]

Under the current framework, regulators have limited powers to require additional information to scrutinise platforms' annual transparency statements. This creates an environment where platforms themselves get to decide what to be transparent about, and the levels of detail and nuance they provide to the public. With reputation risks, legal implications and 'trade secrets' at stake, it is unclear if they have strong incentives towards meaningful transparency.

Beyond this, there is also a skills and resources gap in regulatory bodies and civil society. It is unclear who has the authority or funding to interrogate platforms' performance.

Current frameworks do not comply with human rights principles

The systems and processes employed by digital platforms affect a number of human rights. These include:

- The right to freedom of thought and conscience. People's opinions should not be involuntarily manipulated or influenced, however, platforms' advertising and content recommender systems can interfere with these rights.
- The right to vote in free and fair elections. This involves a complex set of requirements including a free and uncensored press, and fair communication of political issues between citizens, elected representatives and candidates.
- The right to health, the right to life, the right to a healthy environment. All of these can be affected by the inappropriate spread of fact-checked inaccuracies.
- The right to freedom of opinion and expression. This can be affected by both under- and over-moderation by platforms.

16. Sheera Frenkel & Ryan Mac 2023 'Twitter Sues Nonprofit That Tracks Hate Speech', *The New York Times* https://www.nytimes.com/2023/07/31/technology/twitter-x-center-for-countering-digital-hate.html

17. Such as provided by the US' first amendment or article 40.4 of the EU's *Digital Services Act*. Article 40 states that 'Upon a reasoned request … providers of very large online platforms or of very large online search engines shall, within a reasonable period, as specified in the request, provide access to data to vetted researchers who meet the requirements in paragraph 8 of this Article, for the sole purpose of conducting research that contributes to the detection, identification and understanding of systemic risks in the Union, as set out pursuant to Article 34(1), and to the assessment of the adequacy, efficiency and impacts of the risk mitigation measures'

18. See, e.g., the work of the Coalition Independent Technology Research: https://independenttechresearch.org/

19. See for example, Human Rights Law Centre 2023 *Submission on the Communications Legislation Amendment (Combatting Misinformation and Disinformation) Bill 2023* https://www.hrlc.org.au/news/2023/8/22/aust-trailing-big-tech-disinformation-spreads

20. Susie Alegre 2023, *Freedom to Think: Protecting a Fundamental Human Right in the Digital Age*, London: Atlantic Books

These rights are the foundation for free and fair democracies, and many international human rights bodies have highlighted the risks misinformation and disinformation pose to them. As the Human Rights Law Centre[21] notes:

> ***"The United Nations Human Rights Council, the United Nations Secretary-General's High-Level Panel on Digital Cooperation, the United Nations' Special Rapporteurs on the Freedom of Expression, the Organization for Security and Co-operation in Europe, the Organisation of American States and the African Commission on Human and People's Rights have all expressed concern about the rapid spread of disinformation and misinformation and for the need to apply a human rights framework to limit its harm."***

While the Australian Government is a signatory to all the major human rights treaties—including those that have relevance to digital activities—these obligations do not 'flow down' into the operations of companies where they are allowed to self- or co-regulate. **Digital platforms do not have the same obligations to respect human rights, and by allowing them to draft their own codes, human rights principles can be easily overlooked.** Human rights decisions should not be made by private companies that lack democratic and institutional accountability.

Many of the issues regarding misinformation and disinformation are complex. Digital platforms are not best placed to balance these complexities. However, **the process of self- and co-regulation places industry representatives into key decision-making roles, largely without clear guidelines or directives on how to do so.** This often results in a narrow focus on freedom of speech that overlooks the complexity of the full set of human rights involved.

21. Human Rights Law Centre 2023 Submission on the Communications Legislation Amendment (Combating Misinformation and Disinformation) Bill 2023 https://www.hrlc.org.au/news/2023/8/22/aust-trailing-big-tech-disinformation-spreads

Self-regulatory and co-regulatory frameworks do not mitigate risks of misinformation and disinformation

Industry-drafted codes are an inappropriate tool for mitigating the serious risks that emerge on large digital platforms.

Research indicates that where industry drafts codes in the technology domain, they are insufficiently incentivised to counterbalance the overwhelming commercial interests facing digital platforms.[22] They are not appropriate given the scale of the risks faced, do not meet the public's legitimate expectations, and are not suitable for an industry with systemic compliance issues. We have seen this play out in the recent online safety codes, where due to industry's drafting rights, requirements for children's protections were set far lower than in Codes drafted by regulators or legislators, and lower than contemporary industry practice.[23] In this drafting process, industry representatives were incentivised and able to draft a Code with the lowest possible standard and then 'negotiate' with the regulator until sufficient changes were made for registration.[24] This approach clearly placed regulators on the back foot, as industry was able to effectively test for and secure the lowest possible bar of standards for registration.

It remains unclear if the process for developing the *Australian Code of Practice on Disinformation and Misinformation*, which did not require any regulator registration, delivered any better standards for Australian digital users.

The deficiencies of self- and co-regulation models for addressing systemic digital risks are well known overseas.

While the evolution of European regulation is often mischaracterised as having 'jumped straight into' regulative action, Europe also travelled through self-regulation in the form of a *Code of Practice on Disinformation*.[25] Self-regulation was found to be ineffective, and voluntary Codes have largely been supplanted by regulatory requirements outlined in the *Digital Services Act*. The voluntary code produced two issues. First, it was a constant 'cat-and-mouse game' of independent researchers and civil society undertaking small tests and analysis of effectiveness, at their own expense and risk, and platforms responding with either piecemeal changes or rejections of the research findings.[26] Second, there was no accountability or requirements for proactive measures to prevent harms from happening in the first place. The *Digital Services Act* (DSA) remedied this by both establishing risk assessments and mitigation as a requirement to do business in the EU, and providing researchers with some protection through rights to access data which makes independent oversight more authoritative. The DSA is currently in the implementation process.

23. See Reset.Tech Australia 2023 *Response to the revised Online Safety Codes consultation* https://onlinesafety.org.au/wp-content/uploads/wpforms/31-9e10405917e4c106ebe4ec5e69a7bc86/Reset.Tech-Australia-Revised-Codes-Reset-Submission-Google-Docs-869c0da1775d8d037a97bb1a1db860d5.pdf

24. Ibid.

25. See Reset.Tech Australia and Susan McKinnon Foundation 2023 *Response to the Exposure Draft Communications Legislation Amendment (Combatting Misinformation and Disinformation) Bill*, also indexed in Appendix.

26. Brandi Geurkink & Helena McDonald 2020 *Congratulations, YouTube... Now Show Your Work* https://foundation.mozilla.org/en/blog/congratulations-youtube-now-show-your-work/

The *Combatting Misinformation and Disinformation Bill* needs revising

Australia needs legislation that addresses the pressing issue of misinformation and disinformation and corrects the flaws in the existing framework. The *Combatting Misinformation and Disinformation Bill* is admirable in aim but needs significant revision to achieve this ambition.

Creating an effective response by focusing on systems and processes

The Bill lacks sufficient requirements for platforms to create effective responses to misinformation and disinformation. There are no requirements for platforms to undertake and publish risk assessments to proactively reduce risks in systems and processes.

Misinformation and disinformation content has a 'life cycle', and there are many levers platforms can pull to reduce its prevalence or resulting harms.[27] Platforms have various options for content moderation, including removal, demotion, warnings before posting, and limiting content spread. They can also implement measures like restricting the number of forwards a post can receive or taking action against unusual account usage. The current Bill gives platforms autonomy to choose and implement these measures without mandatory disclosure. Requirements for risk assessments would introduce proactive expectations around disclosing these systems, and introduce oversight to this process.

Improving interventions by amplifying independent transparency

As currently drafted, the Bill does not create processes for stronger or more effective public transparency. There are no requirements for access to data for independent researchers, and while ACMA has the authority to request information, there is no guarantee that these powers will be used to make public the kind of consistent, routine, and detailed data necessary to shed light on the matter. For example, simple requirements like a 'live list'[28] of every post that achieves an audience of 10k views or more could create a database that regulators and researchers could use to understand social media content without incurring privacy violations.[29]

There was discussion around the need for something similar to a Freedom Of Information (FOI) mechanism for very large digital platforms, to enable oversight on procedural discussions and decisions, given their importance in shaping Australian discourse.

---

27. In the military defence domain, there is a notion of 'kill chains' which is a parallel way of thinking about these life cycles

28. See, for example, Reset.Tech Australia 2021 Research Memo: Anti-vaccination & vaccine hesitant narratives intensify in Australian Facebook Groups https://au.reset.tech/uploads/resetaustralia_social-listening_report_screen-reader-1.pdf

29. The rationale being that once a post hits a reach of several thousand views, it is evidently public in nature.

Improving interventions through accountability

It is unclear if in its current form, the Bill provides ACMA with the powers or resources to hold some of the world's largest companies to account for their products. For example, the proposed regime for remedial directions, which includes issuing directions and civic penalties at the level of 100 units for incorrect records (around $31,300) will not incentivise change. The penalty for contravening an order from ACMA is greater, at 2% of annual turnover, or at most for 5% for failure to comply with misinformation standards. Regulations regarding digital platforms impose penalties that are double this across Europe and the UK, with 10% appearing to become a regulatory 'norm'.[30] Both the ACCC[31] and ASIC[32] have powers to issue civic penalties up to 10% of turnover, so there is precedent in Australia.[33]

The size of penalties matters. We have recently seen, for example, how X has failed to engage with a serious non-compliance notice from the eSafety Commissioner regarding child sexual exploitation and abuse materials, of $610,500.[34] At this level, these sorts of penalties are simply 'the cost of doing business' and may not incentivise change.

Creating a human rights focused approach

Regulation needs to comply with human rights principles but these are currently missing from the proposed Bill. As described above, misinformation and disinformation affect a broad range of human rights, but as it currently stands the Bill does not address these human rights considerations. This is a missed opportunity.

Moving away from voluntary frameworks and co-regulation

If the *Combatting Misinformation and Disinformation Bill* rests on the assumption that the Digi Code is functional, it will simply introduce the flaws of the Code into legislation. The Bill introduces powers for ACMA to determine its own standard where there is no Code or the 'registered misinformation code is deficient or there are exceptional and urgent circumstances'.

It states that these powers can only be exercised if the registered Code if there is a total failure of the Code, meaning it 'is totally deficient if, and only if, the code is not operating to provide adequate protection for the community from misinformation or disinformation on the services.'[35] The requirements for deficiencies to be 'total' before the ACMA can intervene are excessively high and may lead to substandard protections. **In plain language, the ACMA would presumably have no powers to intervene, while Australians suffer through a 99% deficient Code.**

---

30. For example the UK's *Digital Markets, Competition and Consumers Bill* (2023) which provides for fines of up to 10% of global turnover, alongside an additional 5% for every day a breach continues. Likewise the EU's DSA provides for fines of up to 10% of global turnover, or 20% for repeat offenders.

31. ACCC nd *Fines and penalties* https://www.accc.gov.au/business/compliance-and-enforcement/fines-and-penalties

32. Capped at $782.5million. ASIC 2023 *Fines and Penalties* https://asic.gov.au/about-asic/asic-investigations-and-enforcement/fines-and-penalties

33. For violations of ASIC administered legislation, albeit capped at $782.5million.

34. Georgie Hewson 2023 'Australia's eSafety commission fines Elon Musk's X $610,500 for failing to meet anti-child-abuse standards' *ABC* https://www.abc.net.au/news/2023-10-16/social-media-x-fined-over-gaps-in-child-abuse-prevention/102980590

35. Clause 48.6 Exposure Draft Communications Legislation Amendment (Combatting Misinformation and Disinformation) Bill 2023 https://www.infrastructure.gov.au/sites/default/files/documents/communications-legislation-amendment-combatting-misinformation-and-disinformation-bill2023-june2023.pdf

Suggestions to improve the Bill

Reframing the narrative

Currently, the public debate around the *Combatting Misinformation and Disinformation (Exposure Draft) Bill* has been positioned as a 'free speech vs censorship' debate. This is a mischaracterisation of what is at stake. The Bill, if recalibrated to include better measures around transparency and publication of detailed risk assessments, allows the public (and regulators and researchers) to understand the impact of platform systems and processes on free speech. **This Bill actually enhances free speech, and brings visibility to any measures platforms have that may bias or interfere with political speech. By maximising transparency, rather than fostering censorship, the Bill offers protection against inappropriate restrictions on freedom of expression.**

Requirements for proactive risk assessments

Digital platforms should be required to identify risks posed by their platforms in terms of misinformation and disinformation, and how this affects a broader set of human rights. These risk assessments also need to identify the steps that platforms will take in response across all their systems, including content moderation, advertising and content recommender system transparency. These should be made publicly available.

The EU's *Digital Services Act* (DSA) provides one model for what these provisions and protections could look like.[36] The DSA requires companies to release annual risk assessments that 'diligently identify, analyse and assess any systemic risks in (Europe) stemming from the design or functioning of their service and its related systems, including algorithmic systems, or from the use made of their services'.[37]

Risk assessments need to consider various systemic risks, such as those posed by the dissemination of illegal content; negative effects on human rights; negative effects on civic discourse and electoral processes and public security, and; impacts on gender-based violence, public health and minors and serious negative consequences to the person's physical and mental well-being. Assessments need to consider how content moderation systems operate; platforms' terms of service and how they are enforced, and; advertising approval systems. All of the large platforms operating in Australia are already required to produce these risk assessments, but with a focus limited to Europe.

Introducing similar requirements in Australia would help advance the regulatory response beyond mere notice and take down. While industry often talks up the role of 'notice and take down' in addressing misinformation and is keen to highlight the complex free speech implications of this approach, there are actually multiple approaches to reducing misinformation and disinformation across its life cycle. These should be foregrounded in risk assessments, and requirements to produce these should help to rebalance the discussion.

The publication of these risk assessments also go some way into improving the emerging body of literature around taxonomies of harm (i.e. understanding in more detail what the risks are) and categories of interventions across the harm life cycle (i.e. what platforms can do to mitigate these risks).

---

36. See Article 34, Digital Services Act *Data access and scrutiny* https://eur-lex.europa.eu/legal-content/EN/TXT/HTML/?uri=CELEX:32022R2065

37. Ibid.

Requirements for routine data publication, without requiring explicit ACMA requests

A Bill that creates requirements for routine, meaningful data transparency about content could help address the 'lifecycle' of misinformation and disinformation, and also move the debate beyond notice and takedown. For example, the Bill could require the production of public libraries of:

- All ads approved in Australia, alongside ads rejected for approval,
- Content that achieves 10K plus views in Australia,
- Meta-data about viral content that had over 1k views and was subsequently labelled, removed or demoted,
- Documentation regarding the efficacy of content moderation processes and decision-making around these systems.

These sorts of libraries and repositories could create the conditions that allow regulators and researchers to understand the impact of platforms' systems.

Currently, the Bill gives ACMA powers to request specific information regarding compliance with the legislation, which may or may not become publicly available. However, **the Bill could go one step further and require that some standard evidence and data is publicly released to both enable public oversight and reduce burden on ACMA's investigative team.** Requirements to produce public, routine and meaningful transparency data would create public accountability which could lever concerns around brand reputations as an additional incentive for platforms to improve. To remain responsive, these requirements could be set by Ministerial discretion, similar to the Basic Online Safety Expectations in the *Online Safety Act*.[38] However, unlike the *Online Safety Act*, compliance with these requirements should be mandatory, and violations should lead to penalties.

---

36. eSafety Commissioner 2023 *Basic Online Safety Expectations: Reglatory Guidance* https://www.esafety.gov.au/sites/default/files/2023-09/Basic-Online-Safety-Expectations-Regulatory-Guidance-updated-September-2023_0.pdf

Requirements for access to public interest data and researcher access

Independent oversight is an important tool for accountability, and to furnish regulators with evidence. The Bill needs to be modified to ensure researcher access to public interest data. The EU's *Digital Services Act* (DSA) provides one model for what these provisions and protections could look like.[39] The DSA model places an obligation on platforms to provide regulators 'within a reasonable period specified in that request, access to data that are necessary to monitor and assess compliance with this Regulation'.[40] It dovetails this with three key initiatives for researchers:[41]

- **Reactive data sharing: A framework for accredited researchers and civil society to request data from very large platforms and search engines via the relevant regulator**
- **Proactive data sharing: Industry promises under the Strengthened Code of Practice on Disinformation. Note this scheme now links to the DSA as an example of a 'risk mitigating measure', meaning that companies can point to their performance under the Code of Practice to decrease the risk of regulatory retaliatory action**
- **A draft framework for an independent, third-party intermediary body for vetting data access requests by the European Digital Media Observatory.[42] The existence of a future intermediary body is explicitly mentioned in the DSA and the Code of Practice includes a co-funding commitment from companies.**

Notably, large online platforms operating in the EU already have in-house systems and processes developed to enable this kind of researcher access. Note also, data access interpreted as merely access to APIs alone will be insufficient.

Replace the Digi Code with a regulator standard earlier

Currently, the floor of protections in the Bill is defined by the Digi Code. The Code is weak, ineffective and creates neither adequate accountability nor transparency. Instead of a graduated approach, where ACMA is considered the last resort and the Code must be a 'total failure' before the ACMA can step in, the Bill must be built around a regulator drafted standard. This could be achieved through either ACMA drafting a Standard as part of the Bill process, or by lowering the threshold which allows ACMA to exercise its powers under Clause 48 from 'total failure' to 'significant failures'.

A regulator drafted standard has the capacity to better balance commercial interests and human rights, including the right to freedom of thought and freedom of speech. The development of this standard will require extensive consultation with industry as well as human rights advocates, tech researchers and users themselves. This approach has a much better likelihood of achieving a balance that is pro-user. ACMA would need to be adequately resourced and equipped to do this.

42. European Digital Media Observatory 2022 *Report of the European Digital Media Observatory's Working Group on Platform-to-Researcher Data Access* https://edmoprod.wpengine.com/wp-content/uploads/2022/02/Report-of-the-European-Digital-Media-Observatorys-Working-Group-on-Platform-to-Researcher-Data-Access-2022.pdf

Harmonisation with both the Privacy Act and Online Safety Act Reviews

Some of the measures that may help to address misinformation and disinformation are also reflected in the *Privacy Act* Review and upcoming *Online Safety Act* Review. For example, **allowing people to opt-out of targeted advertising and targeting (i.e. content recommender systems) remains one of the most effective ways of empowering users to avoid popular or monetised misinformation.** A right to opt-out to targeted advertising – or indeed, a more robust approach such as the 'opt-in' model, would also go some way to empowering users to have control over 'what they see' online, including exposure to misinformation and disinformation.[43] Likewise, systemic risk assessment which includes human rights considerations may go some way into shifting the focus of the *Online Safety Act* from notice and take down towards proactive systemic reforms.

Stronger enforcement regimes

The success of the Bill will require regulators to work closely with platforms while providing extensive levels of independent oversight. This will require capacity building and resourcing across the regulatory system.

Alongside this, regulators will need to be given powers to issue meaningful penalties where platforms fail to comply. We note that European and British legislation sets penalties at higher levels—specifically 10% of global turnover—[44] and that other Australian regulations have these powers.[45]

---

43. See for example, Human Rights Law Centre 2023 *Submission on the Communications Legislation Amendment (Combatting Misinformation and Disinformation) Bill 2023* https://www.hrlc.org.au/news/2023/8/22/aust-trailing-big-tech-disinformation-spreads.

44. Such as the EU's *DSA* and the UK's *Digital Markets Bill*.

45. Such as the ACCC for franchising violations (See ACCC nd *Fines and penalties* https://www.accc.gov.au/business/compliance-and-enforcement/fines-and-penalties) and ASIC for violations of ASIC administered legislation, albeit capped at $782.5million (see ASIC 2023 *Fines and Penalties* https://asic.gov.au/about-asic/asic-investigations-and-enforcement/fines-and-penalties).

Recommendations

- Reframe the narrative around the Bill, and focus on how the Bill should enhance public oversight over any censorship measures deployed by the Platforms
- Amend the *Combatting Misinformation and Disinformation Bill* to:
    - Require proactive risk assessments for larger platforms that include consideration of human rights. These could be Australian versions of the risk assessment requirements that are already required and being produced under the EU's *Digital Services Act*, to reduce regulatory burden.
    - Require some routine transparency data from larger platforms be published, to be set by Ministerial discretion and without ACMA requests needing to be made. This would both help improve public transparency and reduce the burden on ACMA's investigative team.
    - Require researcher access to public interest data, to allow independent researchers to be able request relevant data from platforms. These requirements could mimic requirements established under the DSA, which means large platforms would not have to establish new systems to comply.
    - Allow ACMA to step in and replace the Digi Code with a regulator drafted standard earlier, when there is evidence of significant failures rather than total failures. Requirements around total failures could see Australians enjoy largely defunct protections but prevent ACMA from acting because the failures might not be 'total'.
- Consider the implications of the existing *Privacy Act* Review and upcoming *Online Safety Act* Review, especially with regards to consumer opt-outs and risk assessments
- Ensure ACMA is adequately resourced to deliver effective changes
- Ensure ACMA has adequate powers to deliver accountability, especially through the levels of civic penalties resulting from breaches. Penalties of up to 10% are becoming the 'global norm' in digital regulation, and there is no clear reason why the Bill should cap Australia's penalty rates at half this level. Further, other Australian regulators such as the ACCC and ASIC are afforded these powers.

Appendix

Extract from Reset.Tech Australia and Susan McKinnon Foundation's submission to the Exposure Draft consultation, summarising Europe's transition from voluntary codes to the *Digital Service Act*.

From voluntary code to comprehensive regulation: the European experience

This timeline summarises the European experience and shows how legislators gradually responded to the shortcomings of the voluntary industry code with a more comprehensive package. Notably, requirements for data access were consistently invoked to ensure that there were mechanisms for independent assessments of what was otherwise mere platform self-reporting.

March 2018	April 2018	September 2018	January 2019	March 2019	April 2019
Final report of the High Level Expert Group on Fake News and Online Disinformation	European Commission responds with a 'Code of Practice on Disinformation' which would commit online platforms and the advertising industry to provide academia with "access to platform data"	Version 1 of the Code of Practice is released	The European Commission expresses concern on the platforms' failure to benchmark and meaningfully measure progress.	The European Commission remarks platforms "didn't provide access to more granular data to assess the effectiveness of their activities to counter disinformation"	The European Commission calls for independent data access to ensure that the platforms are "not just marking their own homework"

2019-2020	September 2020	2020-2021	June 2022	November 2022	September 2023
An independent assessment by EU Media Regulators (ERGA) notes no sufficient progress was made on platform commitments under the Code.	Findings from the European Commission on the first 12 months of the Code of Practice released, noting "shortcomings mainly due to the Code's self-regulatory nature".	Draft *Digital Services Act* provisions construct a data access regime with a legal basis to force VLOPs/VLOSE to provide access to data to third Parties, including regulators, vetted researchers, and civil society organisations.	Roll-out of the 'Strengthened' Code of Practice on Disinformation.	The *Digital Services Act* enters into force, including risk mitigation duties on platforms and mandated data access for regulators, civil society organisations, and accredited researchers.	First risk mitigation reporting from platforms expected under the *Digital Services Act*.

Acknowledgements

This briefing paper reflects the expertise of those who contributed to the roundtable and paper. Contribution does not necessarily mean endorsement. This includes:

- Aruna Anderson, Reset.Tech Australia
- Emma Briant, Associate Professor at Monash University
- Alice Dawkins, Reset.Tech Australia
- Alice Drury, Human Rights Law Centre
- Rys Farthing, Reset.Tech Australia
- Brandi Geurkink, Mozilla Foundation and Coalition for Independent Tech Research
- Frances Haugen, Beyond the Screen
- David Mejia-Canales, Human Rights Law Centre
- Joaquin Muslera, Beyond the Screen
- Matt Nguyen, incoming 2024 Ford Foundation Technology Fellow
- Racheline Tantular, Reset.Tech Australia
- Bryson Voirin, Beyond the Screen
- Albert Zhang, Australian Strategic Policy Institute

All errors and omissions rest with Reset.Tech Australia.

hello@au.reset.tech