20 August 2023

Submission in response to the Communications Legislation Amendment (Combatting Misinformation and Disinformation) Bill 2023

1. This Bill proposes *censorship by proxy*, in which the government (ACMA) would impose censorship on the Australian public, by coercing DPPs to prevent Australians from accessing and sharing 'misinformation'.

2. This submission is intended to demonstrate that:
    a. there is no mandate for the Bill,
    b. the Bill will undermine freedom of speech,
    c. censorship will result in material harm to the Australian public,
    d. there is an alternate solution (which does impose censorship), and
    e. the Bill cannot be repaired (and should be abandoned).

There is no mandate for the Bill

3. Under the title "Why these powers are needed", the Fact Sheet for the Bill[1] refers to misinformation as being a "major issue worldwide" and claims "a multitude of harms, specifically citing:
    a. "disrupted public health responses",
    b. "foreign interference in elections", and
    c. "the undermining of democratic institutions"

    These 'harms' are discussed below.

"Disrupted Public Health Responses"

4. The ACMA's *Fact sheet 1: key research findings* states that "4-in-5 Australian adults have seen misinformation about COVID-19, with 22% seeing 'a lot' or 'a great deal'" and points to "the propagation of anti-vaccine misinformation narratives within the Australian community".

5. If this COVID-19-related misinformation was as harmful as claimed by the ACMA, it would be reasonable to expect that a significant number of Australians - perhaps the 22% who saw 'a lot' or 'a great deal' of misinformation - would refuse the vaccine. The Australian Government reports, however, that 19,638,849 Australian adults - the vast majority of the adult population - received one or more doses of the vaccine.

6. The ACMA's own evidence, therefore, indicates that misinformation about COVID-19 had *minimal* impact on Australians' acceptance of the vaccine (which has been a remarkable success, not a "disrupted public health response").

---

[1] Page 3 of the Fact Sheet for the Bill

7. The remarkable success of Australia's vaccine takeup demonstrates that the Australian Government is capable of addressing misinformation by communicating information effectively with the Australian public and, therefore, has no mandate for *censorship by proxy* to avoid a disrupted public health response.

"Foreign interference in elections"

8. The most prominent claim of foreign interference in elections is the widely-reported claim that Russian bots used Twitter to interfere with the 2016 US election[2] and allowed Donald Trump to be elected as President. It has been revealed, however, that *the claim* was disinformation[3], and actual Russian disinformation had minimal impact on the election[4].

9. The supporting materials for the Bill have not demonstrated that the government has a mandate to impose censorship by proxy on the Australian public to avoid foreign interference in elections.

"The Undermining of Democratic Institutions"

10. The ACMA's *Fact Sheet 1* refers to misinformation "eroding trust in democratic institutions over time" and, specifically to #stopthesteal. While DPPs should certainly censor *illegal* speech (e.g. incitements to riot), it is difficult to believe that censorship of *legal* speech (e.g. questioning whether an election was valid or claiming that it was not) will *increase* trust in democratic institutions. A democratic institution which censors (even by proxy) the questions and claims that the public have against that democracy institution is giving legitimacy to the belief that the democratic institution is unable to answer those questions and/or is withholding information from the public.

11. The solution to misinformation (and to build trust in democratic institutions) is for the democratic institution to transparently provide *more information* which answers questions from the public and debunks the false claims made in public against it.

12. The government has not demonstrated that it has a mandate to impose censorship by proxy on the Australian public.

The Bill undermines freedom of speech

13. The Fact Sheet for the Bill claims that it contains "strong protections for [...] freedom of speech"[5]. The reality, however, is that the Bill would undermine freedom of speech by:
    ● censoring:
        ○ *legal* content,

---

2 Numerous sources, e.g.
https://www.npr.org/sections/alltechconsidered/2017/04/03/522503844/how-russian-twitter-bots-pumped-out-fake-news-during-the-2016-election
3 https://www.racket.news/p/move-over-jayson-blair-meet-hamilton
4 https://archive.is/abkZG
5 Page 2 of the Fact Sheet for the Bill

- ○ *true* information,
- ○ content provided in *good faith*,
- ○ all information which *contains* misinformation,
- ○ *disagreement* with protected minorities, and
- ○ Information which protects minorities from *actual harm*, and
- providing strong incentives for DPPs to overcensor,
- requiring the ACMA to determine truthfulness, and
- failing to provide oversight for the ACMA.

Censorship of *legal* content

14. The *Code of Practice*[6], which (in a revised form) the Bill would enable the ACMA to enforce, describes the moral and legal basis which should prevent the Australian Government from compelling DPPs to legal content (bold emphasis mine):

> 2.1 *Protection of freedom of expression*: Digital platforms provide a vital avenue for the open exchange of opinion, speech, information, research and debate and conversation as well as creative and other expression across the Australian community. **Signatories should not be compelled by Governments or other parties to remove content solely on the basis of its alleged falsity if the content would not otherwise be unlawful.** Given its subject matter, the Code gives special attention to international human rights as articulated within the Universal Declaration on Human Rights, including but not limited to freedom of speech. Signatories are encouraged to, in developing proportionate responses to Disinformation and Misinformation to be cognisant of the need to protect these rights.

15. The same *Code of Practice*, however, also states that legal content can be misinformation which the Bill would compel DPPs to censor [emphasis mine]:

> 3.6 Misinformation means:
>   A. Digital Content (**often legal**) that is verifiably false or misleading or deceptive;
>   B. is propagated by users of digital platforms; and
>   C. the dissemination of which is reasonably likely (but may not be clearly intended to) cause Harm.

16. The Australian Government is aware of its moral and legal obligations with respect to legal speech and must uphold them. The definition of *excluded content for misinformation purposes* in the Bill must, therefore, be changed to include all legal content.

Censorship of *true* information

17. The definition of *misinformation* in the Bill includes the following statement:

> 7(1)(a) "the content contains information that is false, misleading or deceptive"

---

[6] *Australian Code Of Practice On Disinformation and Misinformation, February 22, 2021*

Which is logically equivalent to:

7(1)(a) "the content contains information that is false OR misleading OR deceptive"

18. The above definition of misinformation means that *true* information can be considered 'misinformation' if it is deemed 'misleading' or 'deceptive' by DPPs (and, ultimately, by the ACMA).

19. A government which knowingly and intentionally censors true information is, in effect, deploying a strategy of disinformation against its own people.

20. The Australian Government must ensure that truth is a defence against censorship, by *preventing* DPPs from censoring information which is true.

21. Twitter has already labelled true information about COVID-19 as *misinformation* on the advice of Stanford University's *Virality Project* (which itself was a source of misinformation about COVID-19)[7].

No Good-Faith Provisions

22. The sole reference to good-faith provisions in the Bill is the phrase "content produced in good faith for the purposes of entertainment, parody or satire"[8]. In this statement, the phrase "content produced in good faith" presumably refers to a genuine intent to entertain, parodise or to satirise.

23. It is worth noting that this phrase [above] is ambiguous and may incentivise DPPs to censor content which is genuinely intended to entertain, parodise or satirise but may, in the eyes of the DPP, have been created with a *broader purpose* which the DPP believes to *not* be "in good faith". This may provide DPPs with incentive to censor entertainment, parody or satire when it has a purpose which is contrary to the political values or biases of that DPP and, ultimately, the ACMA.

24. The Bill provides no other good-faith provisions. This is made clear in the Bill's distinction between *misinformation* and *disinformation*, in which *disinformation* is misinformation offered in bad faith (with the intent to deceive) and, therefore, *misinformation* is always offered in good faith.

25. It would seem likely, therefore, that a government mandate for DPPs to censor *misinformation* (which, by definition, is offered by and to Australians *in good faith*) is likely to be interpreted as the Australian Government *acting in bad faith* and, in turn, *increase* public distrust in the Australian Government.

---

[7] https://twitter.com/mtaibbi/status/1636729166631432195
[8] From the definition of *excluded content for misinformation purposes* in the Bill.

Censorship of content which 'contains' misinformation

26. In the Bill's definition of *misinformation*, the word "contains" dramatically expands the scope of what DPPs must identify and censor as *misinformation:*

    7(1)(a) the content *contains* information that is false, misleading or deceptive;

27. DPPs will, therefore, be required to remove content which contains *any* misinformation, even when the misinformation forms a minor part of the content or is incidental to it.

Censorship of content which is likely to "contribute to" serious harm

28. The Bill's definition of *misinformation* includes content which is reasonably likely to *contribute* to serious harm":

    7(1)(d) (d) the provision of the content on the digital service is reasonably likely to cause or contribute to serious harm.

29. The phrase "cause [...] serious harm" indicates a high threshold for labelling content as *misinformation*.

30. The inclusion of "or contribute to", however, effectively negates the term "cause" by *minimising* the threshold for labelling content as misinformation (as the likely contribution of the content to *serious harm* may be incredibly small but still meet the criterion of "contribute to serious harm").

31. An pertinent example may be the anti-scientific belief that:
    a. any barrier to 'affirmative' models of care for trans-identified children constitutes *serious harm* to those children, therefore
    b. any information which explains the severity and limited effectiveness of 'affirmative' models of care[9] should be censored (as *misinformation*), because
    c. that information may limit access to 'affirmative' models of care (*especially* if the information is *true*).

32. The Bill should maintain a high threshold for defining content as *misinformation* by removing the phrase "or contribute to".

The Bill contains no definition for *serious harm*

33. The Bill provides definitions for *misinformation* and *disinformation* which include the following statement [emphasis mine]:

    (d) the provision of the content on the digital service is reasonably likely to cause or contribute to *serious harm*.

---

[9] https://segm.org/studies

34. While the *Fact Sheet for the Bill* provides a definition for *serious harm*, the Bill does not (the Bill merely provides a definition for *harm*).

35. This omission may have the effect of allowing the definition of serious *harm* (and, in turn, the definitions of *misinformation* and *disinformation*) to have greater scope for interpretation than was intended.

Censorship to protect the special interests of minorities

36. The *Fact* Sheet *for the Bill* describes *serious harm* as something which would impact a "significant portion" of the Australian and, therefore, implies that the definition of *misinformation* (and censorship of the same) would be limited to major threats to the Australian populace at large:

    "Serious harm is harm that affects a significant portion of the Australian population, economy or environment, or undermines the integrity of an Australian democratic process." [Page 1]

37. The Biil's definition of *harm*, however, specifically refers to *minority* groups in its definition of *harm*:

    "*harm* means any of the following:

    (a) hatred against a group in Australian society on the basis of ethnicity, nationality, race, gender, sexual orientation, age, religion or physical or mental disability;"

    [cont.]

38. This dramatically increases the scope of *harm* to the special interests of a small minority of Australians, which is contrary to the expectations set for the Australian public in the *Fact Sheet for the Bill* (that harm relates to "a significant portion of the Australian population").

Censorship of *disagreement* with the claimed interests of minorities

39. The inclusion of *hatred* against protected minorities in the definition of *harm* provides another mechanism by which the scope of *harm* would be dramatically expanded.

40. Under the internal logic and standpoint epistemology of identity politics, there is no objective standard for *hatred* against protected minorities, only a subjective standard which is defined by members of that protected minority. It follows, therefore, that members of those groups have significant scope to claim *hatred* in situations that would not meet any objective or legal threshold for *hatred* or *harm*. The most obvious and common example is the claim that *disagreement* with members of a protected minority constitutes *hatred* against them (e.g. that disagreement with the statement 'trans women are women' constitutes 'transphobia').

41. More than this, identity politics mandates that other people act as 'allies' to protected minorities by proactively looking for instances (or *suspected* instances) of *hatred* and *harm* and then police them on behalf of members of protected groups. This process increases the scope of *hatred* and *harm* even further (to *suspected* instances of *hatred* and *harm*).

42. It follows, therefore, the inclusion of protected minorities in the Bill will dramatically increase the scope of *misinformation* and *disinformation* and, in turn, dramatically increase the amount of censorship experienced by Australians who use these digital platforms (not least because the majority of DDPs are currently and demonstrably supportive of identity politics and, therefore, are highly likely to act as 'allies' by detecting *harm* and censoring content where objective standards for *harm* would detect none and require no censorship).

43. These entirely subjective, expansive and limitless definitions of *hatred* and *harm* under identity politics are, therefore, at odds with the objectivity, proportionality and clarity in Australian law and case law (and under which the concept of *harm* is reasonably well-defined, objective, proportionate and clear) and threatens the individual rights and moral freedoms of Australians (including freedom of speech and freedom from censorship).

44. The Australian Government has a mandate to ensure that new bills, such as this Bill, abide by the safeguards and principles of Australian law and, therefore, not to succumb to political pressure to embed identity politics and its radical agenda into Australia's laws.

45. The following statement should, therefore, be removed from the definition of *harm* in the Bill:

> (a) hatred against a group in Australian society on the basis of ethnicity, nationality, race, gender, sexual orientation, age, religion or physical or mental disability;"

The ACMA *will* determine truthfulness (and harm)

46. The *Fact Sheet for the Bill* states that[10]:

   1. "The ACMA would have no role in determining truthfulness, nor will it have a role in taking down or requesting action regarding individual pieces of content.", and

   2. "the proposed ACMA powers will focus on ensuring digital platform providers have systems and measures in place to combat misinformation and disinformation on their services which pose a risk of serious harm".

47. It is impossible, however, for the ACMA to achieve the latter without doing the former. That is, the ACMA *must* form a view on truthfulness (and the definition of "serious harm") in order to determine whether DPPs have *effective* systems and measures for combating misinformation and to determine whether the *misinformation code* has failed in part or in

---

[10] Ibid.

full. It follows that the ACMA will *enforce* its view on truthfulness and harm on DPPs through hard and soft power.

48. The claim that the ACMA will not have a "role in taking down or requesting action regarding individual pieces of content" is largely moot, as it will set the policies which will determine the action on *all* pieces of content. Moreover, it would be absurd for the ACMA to compel DPPs to perform censorship by proxy and then give DPPs free reign to set their own standards for that censorship and assess their own efficacy.

No oversight for the ACMA

49. The Bill contains no provision for oversight of the ACMA as it compels DPPs to censor the information available to Australians and exercises significant discretion in doing so.

50. The Bill contains no constraints on the ACMA's use of soft power and alternate lines of communication in influencing the behaviour of DPPs which may allow the ACMA to act against the interest of the public without detection in the short term (as US Federal Agencies did in pressuring Twitter to do its bidding[11]).

Strong incentives for DPPs to overcensor

51. The Bill imposes severe penalties for DPPs which fail to satisfy the ACMA that they have built systems which censor misinformation from the Australian public. This is, effectively, a powerful incentive for DPPs to provide a minimum level of censorship on their platforms.

52. While the Bill provides a definition for excluded content for misinformation purposes, the Bill does not prevent DPPs from censoring this excluded content. And, as stated earlier, the definition of excluded content does not include legal or true information. The bill does not, therefore, mandate any limits to the censorship performed by DPPs (which would protect the free speech of Australians).

53. The Bill, therefore, provides DPPs with strong incentive to overcensor information, to 'play it safe' by denying Australians access to information which falls short of the definition of misinformation.

54. A Bill which actually provided "strong protections for [...] freedom of speech"[12] would *limit* DPP censorship and provide *penalties* for unwarranted censorship of content created by and made for the Australian public.

Material harms from censorship

55. Given that many other submissions are likely to point to the harms of censorship, I have not attempted to provide a comprehensive review of such harms (I have limited my response in this section to two harms which may not receive coverage elsewhere).

---

[11] https://twitter.com/mtaibbi/status/1606701397109796866
[12] Page 2 of the Fact Sheet for the Bill

Censorship of information which protects minorities from material harm

56. Another unintended consequence of including "hatred" against protected minorities in the definition of *harm* in the Bill is that censorship of *disagreement* with protected minorities (in the belief that this disagreement constitutes 'hatred') can cause *material harm* to the Australian public (and, in particular, the members of those protected minorities) by denying them access to vital information about material reality.

57. An example of the above is the treatment of children who identify as transgender. There is a cultural expectation to censor information which disagrees with the 'affirmative' model of care, driven by the belief that this protects trans-identified children from *harm* (by limiting their access to 'affirmative' care). This censorship, however, has resulted in *material harm* to children by masking the severity of 'affirmative' treatments and perpetuating the false belief that they are highly effective[13].

Scientific progress requires anomalies (a.k.a. 'misinformation')

58. The seminal book "The Structure of Scientific Revolutions" by Thomas S. Kuhn explains science makes advances through discontinuous and revolutionary shifts in understanding which are triggered when *anomalies* accumulate against the current scientific paradigm.

59. If *anomalies* are censored as *misinformation*, science cannot progress and limit our ability to understand and solve problems.

60. A pertinent example is the scientific consensus on climate change and the mandate to maintain the *scientific* consensus in order to mandate a *political* consensus which, activists claim, is the only solution to climate change. This *political* goal provides incentive for *anomalies* (information which is contrary to the scientific consensus on climate change and/or may lead to solutions which are *technological* - rather than *political* - in nature) to be labelled as *misinformation* and censored. In turn, this may:
    a. prevent humanity from understanding climate change and addressing it technologically (e.g. adopting nuclear power at scale), and *also*
    b. enable political interventions (e.g. 'degrowth' in capitalism) which may be both ineffective in addressing climate change and harmful to humanity at scale[14].

An Alternate Solution

Mandating friction in DPPs (rather than censorship by proxy)

61. Former Facebook employee and algorithmic expert Frances Haugen (also known as 'The Facebook Whistleblower') gave testimony to the UK Parliament which describes an alternate solution for addressing misinformation which is more effective than censorship (emphasis mine)[15]:

---

[13] https://segm.org/studies
[14] https://www.wired.com/story/opinion-why-degrowth-is-the-worst-idea-on-the-planet/
[15] https://youtube.com/watch?v=E-ITxQk6Xio&t=1460s

"I have been mischaracterized repeatedly, in certain parts of the internet that I'm here as, like, a plant to get more censorship. One of the things that I saw over and over again in the [Facebook] docs was that **there are lots and lots of solutions that don't involve picking good and bad ideas** - they're about designing the platform for safety, slowing the platform down and that when you focus - when you give people more content from their family and friends you get - for free - less hateful, divisive content. You get less misinformation. You get - because the biggest part that's driving misinformation is these hyper distribution nodes, these groups where it goes out to 500,000 people.

Some examples of non-content-based interventions are things like ... let's imagine Alice post something and Bob re-shares it, and Carol re-shares it and now it lands in Dan's newsfeed. If Dan had to copy and paste that to continue to share it, so his share button was greyed out - that's a two-hop re-share chain **- that [UI change] has the same impact as the entire third party fact checking system**, only it's going to work in the global south [where Facebook is not investing money on fact-checking]. It doesn't require us to have a language-by-language system, it just slows the platform down.

Moving to systems that are human scaled, instead of having AI tell us where to focus, is the safest way to design social media. And I want to remind people - we liked social media before we had an algorithmic feed. And Facebook said like, if you move to a chronological feed you won't like it. And it's true with groups that are 500,000 people where just like a spraying content people, you're not going to like it.
But facebook has choices that it could do in different ways - if you have groups that were designed like these things called Discord servers, where it's all chronological but people break out into different rooms as it gets too crowded. That's a human intervention - a human scale solution - not an AI-driven solution. And so slowing the platform down, content-agnostic strategies, human scale solutions - that's the direction we need to go."

Concluding Statement

62. The large number of problems in the Bill demonstrate the counterproductivity of censoring 'misinformation' to improve public safety.

63. This may be driven by an unstated belief that the government, the ACMA and DPPs have (or can obtain) a neutral and authoritative viewpoint from which 'misinformation' can be defined, identified and removed without causing significant harm to the Australian public and our democratic institutions. The road to hell is paved with this kind of folly.

64. While the Bill could be improved (to become *less harmful* to the Australian public than this Exposure Draft would be), it cannot be fixed.

65. The Australian Government must abandon this Bill and all other attempts to censor legal information.