To Whom It May Concern,

I am writing this submission to address my concerns regarding the above-named Bill.

Preamble

Misinformation and disinformation are not new problems but have increased in scope significantly in the age of the Internet. The trouble is that discerning fact from fallacy is irreducibly complex, and no centralised body can or should take on that responsibility. Our best resource for determining truth is the consensus formed by a general public who are individually free to perceive reality and express their interpretations as they see fit.

As the renowned social conformity researcher Solomon Asch stated so elegantly: "*Life in society requires consensus as an indispensable condition. But consensus, to be productive, requires that each individual contribute independently out of his experience and insight. When consensus comes under the dominance of conformity, the social process is polluted and the individual at the same time surrenders the powers on which his functioning as a feeling and thinking being depends."*[1]

Managing misinformation means to manage the voices of those who do not conform to the accepted narrative. Our society and individual freedoms are weakened in proportion to the extent that this consensus-forming process is manipulated. In the sense of Asch's description, I regard the eSaftey Commission as already being an unacceptable threat to individual freedom, and I do not support equipping it or the ACMA with additional powers.

For the health of our democracy Australians must continue to enjoy the freedom to appraise unfiltered information, the freedom to express themselves, the freedom to be wrong, and the freedom to associate with others who share their beliefs (including via social media connection).

The Australian Government, elected officials, and public servants must respect the autonomy and individual judgement of constituents, respect the consensus-forming process, and avoid taking on the paternalistic role of truth arbitration (directly or indirectly as that may be).

Responses to Specific Areas of Feedback Sought

1. The definitions of misinformation and disinformation

I find that the definitions of "misinformation", "disinformation" and "serious harm" are inadequate due to the high level of subjectivity involved.

For misinformation or disinformation to be identified and managed, somebody at some point needs to make a judgement about what is true and what is false. This right belongs exclusively to individual constituents, not to public servants, journalists, the employees of digital platforms, or anybody else.

[1] Asch, Solomon E. "Opinions and social pressure." *Scientific American* 193.5 (1955): 31-35

This legislation and these definitions have the potential to stifle genuine critique of Government actions, and to hinder debates on controversial topics like climate change and energy policy, the indigenous voice to parliament, the management of the pandemic, immigration policies, and appropriate medical care for children expressing gender dysphoria (to name a few).

The definition of "serious harm" in the Bill is also highly subjective. In this ideological age of words being deemed "literal violence" and unintended offence being deemed "microaggression", I do not trust public servants (or any third party) to make judgements on my behalf regarding what types of information might cause harm to me or the Australian people.

Example

As an example of where this type of Government process has already become subject to uncertainty, subjectivity, and ideology, I draw on the media release[2] made by the eSafety Commission on 22/06/2023.

This media release targeted Twitter (now X Corp), and quoted commissioner Julie Inman Grant:

> *"We are seeing a worrying surge in hate online … Twitter appears to have dropped the ball on tackling hate. A third of all complaints about online hate reported to us are now happening on Twitter … We are also aware of reports that the reinstatement of some of these previously banned accounts has emboldened extreme polarisers, peddlers of outrage and hate, including neo-Nazis both in Australia and overseas."*

In itself, "peddlers of outrage and hate" is already a highly subjective claim. I have been following online trends closely over the past 5-10 years. I have seen many social media personalities accused of hate speech and of being Nazis, and rarely have I agreed that this is the case. Most often, those accusations are made in bad faith, as *ad hominem* attacks, and could not be farther from the truth.

The second issue is the use of number-of-complaints as the indication that the Twitter/X has "dropped the ball on tackling hate". This assertion highlights the risk that the eSafety Commission may be misled/weaponised by a coordinated submission of vexatious complaints. Anyone who has used social media platforms long enough will realise that this is an inevitability, rather than merely hypothetical. Just because a third of complaints to the eSafety Commission relate to Twitter does not mean that the complaints are homogenous, reasonable, valid, or made in good faith.

After defensively stating "eSafety is far from being alone in its concern about increasing levels of toxicity and hate on Twitter", the media release goes on to cite sources:

[2] *Media Release: eSafety demands answers from Twitter about how it's tackling online hate*, eSafety Commission 22/06/2023 [https://www.esafety.gov.au/newsroom/media-releases/esafety-demands-answers-from-twitter-about-how-its-tackling-online-hate]

> *"… US advocacy group GLAAD designated Twitter as the most hateful platform towards the LGBTQ+ community as part of their third annual social media index … Research by the UK-based Center for Countering Digital Hate (CCDH) demonstrated that slurs against African Americans [had more than doubled and] … also found that those paying for a Twitter Blue Check seemed to enjoy a level of impunity when it came to the enforcement of Twitter's rules governing online hate, compared to non-paying users and even had their Tweets boosted by the platform's algorithms".*

My first issue here is the rhetorical use of "levels of toxicity and hate". What is the difference between the two? Does the eSafety Commission, as an official Government body, have specific definitions for what it means? Or is this simply impassioned hyperbole? This strikes me more as the sort of language an author with an agenda would use, not the language of an impartial Government agency.

Secondly, the eSafety Commission seems to insinuate that it is somehow undue for paid users to have Tweets boosted by Twitter's algorithm. This implies an inappropriately opinionated stance from eSafety on the operation of Twitter's algorithms.

Finally, the media release references two activist organisations (GLAAD and CCDH). No opposing opinions or perspectives were included in the media release. X Corp has since filed legal action against CCDH, and I have included excerpts from the complaint[3] here as it highlights X Corp's alternative perspective and is exemplar of the complexity involved. The X Corp complaint states:

> *"… CCDH [are UK and US based] activist organizations masquerading as research agencies, funded and supported by unknown organizations … [which have] embarked on a scare campaign to drive away advertisers from the X platform. CCDH has done this by engaging in a series of unlawful acts designed to improperly gain access to protected X Corp. data, needed by CCDH so that it could cherry-pick from the hundreds of millions of posts made each day on X and falsely claim it had statistical support showing the platform is overwhelmed with harmful content."*
>
> *"CCDH's underhanded conduct is nothing new. It has a history of using similar tactics not for the goal of combating hate, but rather to censor a wide range of viewpoints on social media with which it disagrees. CCDH's efforts often rely on obtaining and intentionally mischaracterizing data in "research" reports it prepares to make it appear as if a few specific users (often media organizations and high profile individuals) are overwhelming social media platforms with content that CCDH deems harmful. CCDH uses those reports to demand that platform providers kick the targeted users off of their platforms, thus silencing their viewpoints on broadly debated topics such as COVID-19 vaccines, reproductive healthcare, and climate change. In this manner, CCDH seeks to*

[3] *X Corp. v. Center for Countering Digital Hate, Inc., et al*., Case No. 3:23-cv-03836, U.S. District Court for the Northern District of California, Filed 31/07/2023, [accessed at: https://cdn.arstechnica.net/wp-content/uploads/2023/08/X-Corp-v-Center-for-Countering-Digital-Hate-7-31-2023.pdf]

> *prevent public dialogue and the public's access to free expression in favor of an ideological echo chamber that conforms to CCDH's favored viewpoints."*

So who is right? The situation is very complex. And this is just the point.

Through this example I have shown that the eSafety Commission already appears to be behaving rashly and relying on the evidence of activist groups to guide its activities. This behaviour is enabled by the subjective definitions of misinformation, disinformation, and serious harm, illustrating my point that the definitions are inadequate. What protections are in place to prevent undue external/internal influence of the activities of ACMA and eSafety?

Finally, I note that the Commission seems to be taking a punitive and hostile approach toward the industry, using language in media release titles including "eSafety demands answers" and "Twitter, TikTok and Google forced to answer tough questions", and intimating "maximum financial penalties of nearly $700,000 a day [should Twitter fail to respond within 28 days]".

These are not government agencies I would feel comfortable granting more power than is currently held. If anything, I would like to see the existing powers repealed.

2. The definition of digital platform services and the types of services we propose be subject to the new framework.

It is alarming to me that "unauthorised electoral or referendum content" is specifically regarded as being within scope for this legislation. Australian citizens have a right to voice their political opinions about elections and referenda regardless of whether a public servant, politician, employee of a digital service, or anyone else thinks that opinion is valid.

This Bill provides protection to the voices of the Government, mainstream media, and educational organisations. This biased safeguarding diminishes equality in speech among entities in Australia. It risks fortifying the dominance of the political class and establishment media, and risks the sidelining of independent media and alternative political entities.

In this sense, this Bill seems to have been crafted to take us back to a time before social media, when mainstream media acted as gatekeepers for which political opinions would be broadcast, and which political topics would dominate the national discourse. Indeed, according to the Daily Mail[4], on 15/08/2023 in a hypothetical discussion Prime Minister Anthony Albanese told 3AW radio host Neil Mitchell that if he was a dictator that it "would be handy" to ban social media, noting "keyboard warriors who can anonymously say anything at all and without any fear". The time before social media may have been more convenient for politicians, but it was not a better time for democracy or accountability.

[4] Anthony Albanese reveals he would ban social media and abolish the state if he ever became dictator, *Daily Mail UK*, published online 16/08/2023 [https://www.dailymail.co.uk/news/article-12411539/Anthony-Albanese-dictator-ban-social-media-abolish-states.html]

I also note, by means of the definition of "professional news", that the ABC and SBS are explicitly excluded from the scope of the proposed Bill. For years now ABC News has taken to publishing all manner of journalist opinion pieces under the guise of the "analysis" label. How is it fair that the political opinions of tax-payer funded ABC staff are afforded a greater audience and legislative immunity than that of individual Australians posting on social media?

3. **How instant messaging services will be brought within the scope of the framework while safeguarding privacy.**

   Instant messaging services should be entirely excluded from this legislation. There is no role for government to infringe on the private communications of Australian citizens. This is the true for individual messaging and broadcast messaging.

4. **The scope of the information-gathering and recording keeping powers, which includes the prevalence of false, misleading, or deceptive information on digital platform services.**

   The information-gathering powers proposed by this Bill are unreasonable and represent clear Government overreach. Any requests for information from digital service agencies, third parties, or individuals should at least be required to be made through the court system to safeguard against abuse of agency power.

   It is not reasonable to expect digital service providers, often processing millions of pieces of information daily, to be able to classify each and every piece of information as being either true, false, misleading, or deceptive. This is what would be required of them under this Bill.

5. **The preconditions that must be met before the ACMA can require a new code, register a code and make an industry standard**

   The preconditions are inadequate to prevent the AMCA from abusing its power and using the prospect of an ACMA-mandated code as a coercive tool to manipulate the behaviour of digital service providers or other industry stakeholders.

6. **How the digital platforms industry may be able to operationalise the Bill and various content exemptions (e.g. professional news, satire, authorised electoral content)**

   I think that the operationalisation of this Bill would unfairly burden digital services providers with unreasonable levels of complexity and expense. It is not reasonable to expect that a digital service provider, often processing millions of pieces of information per day, be able to classify all content as exempt or non-exempt from the proposed Bill.

   There is a significant risk that digital service providers will opt to save costs by overcorrecting and restricting speech of users much more harshly than is necessary, causing significant harm to Australians.

7. Appropriate civil penalties and enforcement mechanisms for non-compliance.

The proposed penalties are far too high. The magnitude of these penalties risks creating an atmosphere of self-censorship amongst digital service providers to minimise the risks associated with inadvertent penalties. The risk is that there would be an overcorrection that restricts speech of users much more harshly than is necessary, causing significant harm to Australians.

Unintended Harms of this Bill

The proposed Bill and its supporting documentation comment extensively on the possible harms of misinformation/disinformation, but make almost no statements about the possible harms of censorship or restriction of individual expression.

To mirror Table 1 from the Department's guidance note[5] for this Bill, which outlines a series of hypothetical harms that may be caused by misinformation, I have prepared the below table of harms that might come about due to the Bill's implementation.

Type of Harm	**Hypothetical Example of Serious Harm**
Suppression of Minority Voices	A marginalised group are unable to share their experiences or grievances as they have been labelled as disseminators of misinformation.
Stifling of Political Dissent	An unauthorised political group critical of the Government is unable to share their views as their content has been suppressed by digital service providers in accordance with an industry code.
Undermining of Public Trust	The public becomes wary of what they read online, not due to misinformation, but due to a belief that content is being manipulated or censored through Government coercion of digital service providers.
Hindered Academic and Scientific Progress	Researchers and experts hesitate to share preliminary findings or alternative theories, fearing they might be misconstrued as misinformation. (This example is particularly topical as it has been shown to have happened on Twitter during the pandemic[6]).
Reduction in Artistic and Creative Expression	Artists and content creators cannot grow their audiences because their satire has been unfairly flagged as mis/disinformation. Despite being excluded from the Bill, false positives are common amongst digital service providers who struggle to classify millions of posts per day.
Erosion of a Free and Open Society	Over time, the general public self-censor more broadly, leading to a society where dissenting political opinions are rare. The stagnant political *status quo* results in mediocre performance of politicians and economic decline. Standard of living falls, morale falls, and national security is weakened.

[5] Communications Legislation Amendment (Combatting Misinformation and Disinformation) Bill 2023—guidance note, *Department of Infrastructure, Transport, Regional Development, Communications and the Arts*, Published 24/07/2023 [https://www.infrastructure.gov.au/department/media/publications/communications-legislation-amendment-combatting-misinformation-and-disinformation-bill-2023-guidance]

[6] The Twitter Blacklisting of Jay Bhattacharya: The social-media platform revealed that many had been censored and shadow-banned, *Wall Street Journal*, Published 09/12/2022, [https://www.wsj.com/articles/the-twitter-blacklisting-of-jay-bhattacharya-medical-expert-covid-lockdown-stanford-doctor-shadow-banned-censorship-11670621083]

Conclusion: An Alternative Approach to Misinformation and Disinformation

I do not support the proposed Bill, and I do not believe that the ACMA and eSafety Commission are fit for purpose in their current formulation. I propose the following alternative approach to the management of misinformation and disinformation by the Government.

I encourage the Government to:

- take a corrective approach to misinformation and disinformation, whereby the relevant agency dynamically publishes evidence-based public awareness and education material to counter misinformation, and
- refrain from "combatting" misinformation through restrictive, censorial, or coercive means, including the use of industry codes and penalties that apply these means indirectly.

Through this alternative approach Australian citizens remain the fundamental unit of discernment, and the Government fulfils its role in harm reduction by maintaining the availability of high-quality information. This approach avoids the top-down, punitive measures proposed in this Bill and instead places trust in the wisdom of the Australian people.

Thank you for the opportunity to make this submission. I implore the relevant representatives to consider its content carefully.

Yours Sincerely,

Dr Thomas Everingham
MD B.Eng (hons) B.Sci