



Submission from AMAN on the Exposure Draft of the Communications Legislation Amendment (Combatting Misinformation and Disinformation) Bill 2023

## **ABOUT AMAN**

AMAN works to prevent the harm resulting from racism, hatred and Islamophobia directed towards Muslims online. We contribute policy proposals that benefit all Australians and consider the complexities in implementation. Australian Muslims are disproportionately affected by misinformation and disinformation that portrays Muslims as a threat and lacking human qualities. Arabic-speaking people and Muslims worldwide are also subject to over-policing in relation to content moderation. We are very attuned to the need for precise definitions of harm that lift the burden of fighting misinformation from the community and also uphold freedom of expression.

AMAN notes the proposed legislation is complex and has prepared these submissions within time and resource constraints. We welcome any further queries from the Australian Government or other civil society.

## **OVERVIEW**

The proposed Communications Legislation Amendment (Combatting Misinformation and Disinformation) Bill 2023 (**'the Bill'**) seeks to extend the core regulatory framework applicable to traditional communication and broadcasting mediums to digital platforms for a specific purpose: to develop mechanisms for the reduction of misinformation and disinformation.

The Bill's underlying principle of holding accountable digital platforms and their providers for the content on their platforms is something AMAN supports wholeheartedly. We also congratulate the Australian Government for taking a consultative approach with this legislation, including releasing an exposure draft before it enters Parliament. The current drafting needs work to achieve the Bill's aims. We are concerned that the current drafting grossly overestimates

- The capability of ACMA to assess vast amounts of information and monitor emerging trends and patterns of harm.
- The willingness and commitment of digital platforms to moderate harmful content against their business interest and in the Australian public interest.

The human rights at stake include freedom of expression, freedom of opinion, privacy and security of persons. Social media is not a free marketplace for ideas. Certain ideas gain high engagement by being polarising, invoking ingroup fear or hatred towards outgroups, which gives those ideas greater reach and prominence online. Digital platforms are not transparent about how their business objectives influence the presentation and amplification of material online.

The European experience underscores that a self-regulatory or co-regulatory approach concerning these harms is ineffective and a waste of time.

The Bill should not assume that platforms will improve over time but set standards from the beginning that are defined well enough to have broad public support and a high impact on platforms. Platforms will not allocate resources in Australia to moderate harmful content unless it is specifically defined and clearly made their responsibility by law. Enforcement of the law must also provide adequate deterrence to companies that make inordinate profits from socialising misinformation and disinformation because of its tendency to generate high user engagement.

The Digital Services Act in Europe does not define categories of harm in detail because it leaves that to national laws. Our context is different because the categories of harm referred to in the Bill are not defined in other laws.

Online information operations designed to dehumanise outgroups to ingroup audiences is a core technique of racist nationalist and other violent extremist movements and state-sponsored foreign influence operations online. Our definitions of dehumanising materials (**see Schedule 3**) have been developed using a social psychology framework (Haslam), genocide prevention-related research (Leader and Maynard), studies on hate crime and incidents (Asquith), Canadian legal authorities on hate speech, and AMAN's practical ethnographic observations of information operations online.

The current lack of ACMA powers maintains a significant social burden on Australia's most affected communities. Currently, Australia's vilification laws are the only avenue available, which AMAN is testing through various complaints. Bringing a complaint under those laws is time-consuming, costly to victims and ineffective in achieving systemic and lasting change. Perpetrators also benefit from the publicity resulting from a complaint under vilification laws.

## RECOMMENDATIONS

To ensure that misinformation or disinformation that dehumanises groups based on protected characteristics is treated as a public harm rather than a private problem, this Bill must

- A. Review the ACMA Act and Broadcasting Services Act objectives to underline that a public information environment that supports diversity of opinion, veracity and accuracy of information is vital to Australia's obligations under various international instruments, including the ICCPR (freedom of expression, the right to non-discrimination, no advocacy of hatred), IESCR (the right to health) and environment. Preventing and moderating the advocacy of hatred enables greater freedom of expression by groups targeted by hatred. It also supports their fulfilment of the right to health by reducing their exposure to a social atmosphere that denies their human qualities.
- B. Ensure the overall design of the Bill supports, above all, immediate powers for ACMA to prevent well-defined harms effectively and, secondly, to handle community complaints more effectively. **Figure 1 below** outlines the underlying rationale for this approach.

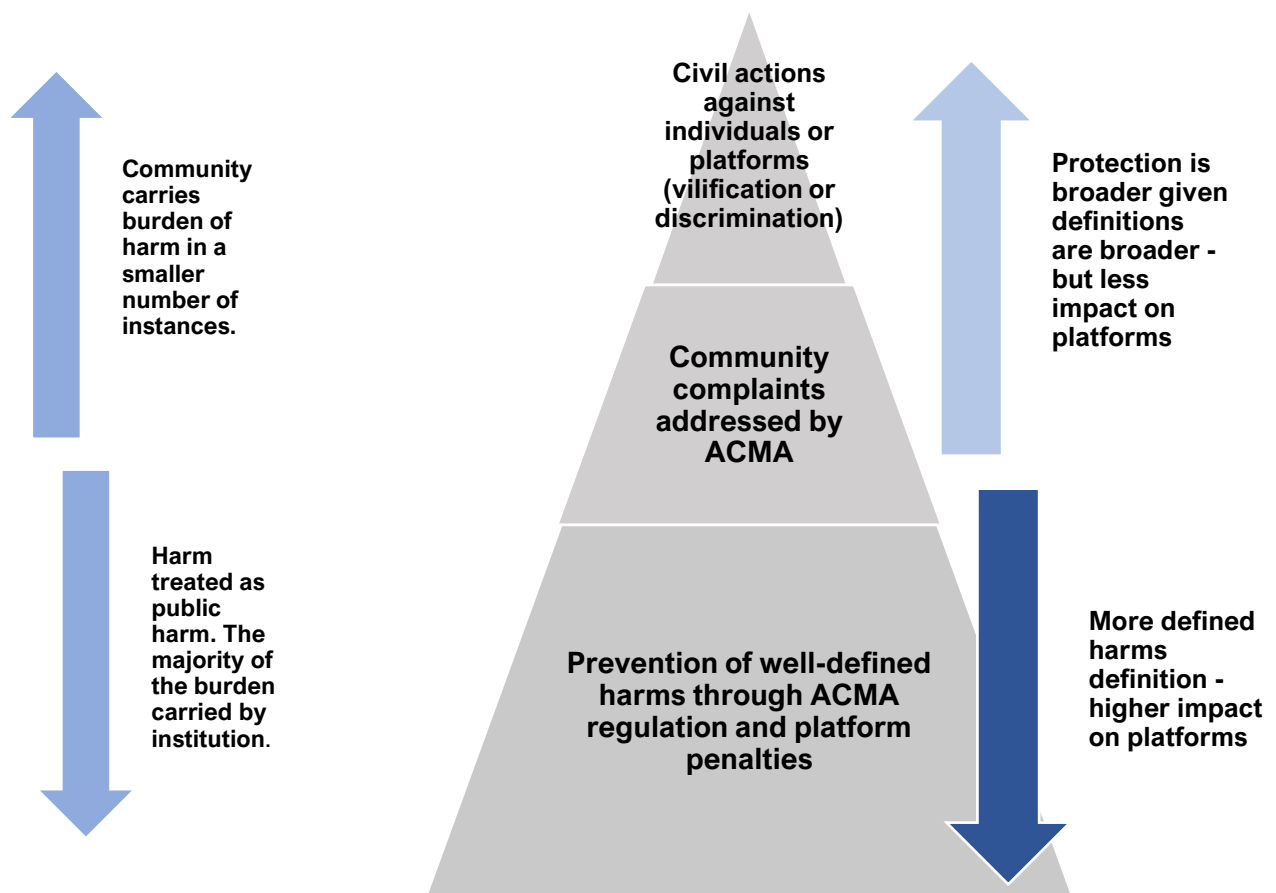


Figure 1 – Goal of regulation

- C. Immediately require clear standards to be set that define the harm of hatred in greater detail, using our definition of dehumanising material as a starting point (see **Schedule 3**).
- a. This definition outlines clearly hateful material, whether communicated through speech or words; the curation or packaging of information; images; and insignia.
  - b. This definition is universally applicable and resilient to cyclical changes in targeted groups.
  - c. It is also more capable of securing public support and consensus about what constitutes hatred across different contexts, noting that there is significant disagreement even within the Australian community about definitions for different forms of prejudice and racism.
  - d. Methodically, it links to one of the most dangerous forms of hatred – hatred that positions a group outside the human family, making them an easier and more deserving target for violence.
  - e. Similar standards should be set in relation to the other categories of harm (e.g. harm to the environment) following community consultation.
- D. Improve governance arrangements and capability by
- a. increasing culturally diverse representation within ACMA, including ACMA’s complaint-handling sections. Consider public reporting by ACMA about its cultural diversity within staff.
  - b. expanding the mechanisms for monitoring and assessment to include researchers and civil society, who are more equipped to identify emerging trends and patterns in misinformation and disinformation. As a starting point, consider Article 40 of the European Digital Services Act and the ‘crowdtangle provision’ supporting immediate access to aggregations of public data.
- E. Amend the definition of ‘professional news content’ or remove the exemption entirely. If amending the definition, do not include self-regulatory codes and make certain operational transparency and accountability requirements mandatory.
- a. It is in the public interest for this Bill to not allow well-resourced and far-reaching news outlets to continue misinformation. At the very least, the Bill must increase their requirements for transparency and accountability to benefit from that exemption.
  - b. AMAN recommends that the Australian Government work with Australian researchers, civil society and the Global Disinformation Index to formulate these requirements. AMAN provides proposed wording as a starting point for discussion in **Schedule 2**.
- F. Better explain and define ‘election and referendum matters’ given its exemption from the Bill and ensure that online content that causes harm defined by the Bill is not exempt from the Bill *unless* it is covered by AEC legislation. Then list the relevant AEC legislation and provisions in the Bill.

- G. Strengthen existing complaint mechanisms by linking a standard that defines dehumanising material to existing provisions within the Broadcasting Services Act on incitement of hatred.
- a. The Bill currently does not expand ACMA’s remit in relation to handling complaints. ACMA cannot be expected to make judgements about what is vilifying without guidance and access to expertise.
  - b. This lack of expertise is underscored by AMAN’s complaints to ACMA concerning Sky News, which was not found to be vilifying. Currently, a news outlet can take an editorial stance to promote the view that a particular religious group is a threat and promote that view as a fact, to intentionally ignore counter views, and be protected under ‘opinion’ or ‘current affairs’ categories that do not require the same veracity of fact-checking.
  - c. While existing media regulation safeguards against vilifying material, ACMA does not have sufficient expertise to understand demographic invasion, Eurabian and great replacement conspiracy theories used to justify terrorist and genocidal violence against Muslims worldwide.
- H. By further amendment to the *Online Safety Act 2021* (Cth), prohibit the serial or systematic publication of dehumanising material, as defined in our working definition. Such amendments could include takedown powers for e-Safety and penalties for serial bad actors.
- I. By a further miscellaneous amendments bill, clarify that
- a. section 18C of the *Racial Discrimination Act 1975* (Cth) has extraterritorial application to foreign-based digital platforms;
  - b. discrimination provisions of various federal discrimination laws have extraterritorial application to foreign-based digital platforms; and
  - c. relevant entities can bring *discrimination* complaints on behalf of groups or communities based on protected characteristics.

## **DISCUSSION**

### **PART 3 of SCHEDULE 9 – MISINFORMATION CODES AND STANDARDS**

#### ***The co-regulatory approach will not protect Australians and Australian democracy***

However, the co-regulatory approach of misinformation codes is an inappropriate mechanism for achieving this as digital platforms pursue and model of business that is frequently at odds with community needs and public interest.

The European regulatory experience, which began with self-regulation and has ended with regulation, underscores this point. We should learn from the European experience rather than repeating their mistakes.

#### ***Misinformation codes – complaint and review mechanisms***

For a misinformation code to be registerable, ACMA must be satisfied with three elements pertaining to complaint and review mechanisms. Firstly, the code ‘requires participants in that section of the digital platform industry to implement measures to prevent or respond to misinformation or disinformation on the services’.<sup>1</sup> Secondly, the code enables the assessment of compliance with the prevention or response measures.<sup>2</sup> Thirdly, the code provides adequate protection from misinformation or disinformation.<sup>3</sup>

The digital platform industry and the content therein are highly expansive. Even under the proposal that the digital platform industry be divided into sections with representatives, the onerous task of monitoring review and complaints mechanisms is beyond the capability of any single regulator.

Access by vetted third parties to data pertaining to complaints and review mechanisms for analysis of their efficacy is essential to the successful evaluation of such mechanisms.

#### ***Complexity and vastness of digital technology content***

The Bill’s success is contingent on collecting and analysing data on the occurrence of misinformation and disinformation to determine the efficacy of response mechanisms and co-regulation. Timely identification of trends and evaluation of response mechanisms is essential in a field as labile as digital technology.

AMAN understands that ACMA has experience in data analysis and preparing reports on broadcasting, radiocommunications, telemarketing, and similar communication fields. This includes regular Regulation Impact Statements making specific policy recommendations.

---

<sup>1</sup> Ibid cl 37(1)(e)(i).

<sup>2</sup> Ibid cl 37(1)(e)(ii).

<sup>3</sup> Ibid cl 37(1)(e)(iii).

The breadth of content on digital platforms is, however, immeasurable in comparison to traditional methods of communication and broadcasting. So too, is the breadth of cultural and linguistic diversity in digital content, each requiring disparate methods of approach. The European Commission recognised this difficulty in October 2022 when creating the Digital Services Act. It included provisions permitting researchers to access relevant data for very specific purposes.

Similarly, this Bill needs to empower ACMA to involve researchers in gathering, analysing and assessing data.

### ***Platform to researcher data sharing***

The Digital Services Act to Article 40 of the digital services act (Europe) enables platform-to-researcher data sharing, with guardrails for privacy protection.

The Digital Services Act requires very large online platforms or very large online search engines to provide access to data to vetted researchers for the sole purpose of conducting research.<sup>4</sup> Researchers become vetted upon application if their application fulfils the following requirements:<sup>5</sup>

- (a) Affiliation with a scientific research organisation;
- (b) Independence from commercial interests;
- (c) Disclosure of the funding of their research;
- (d) Capability of fulfilling specific confidentiality and data security requirements in relation to protecting personal data and a description of their specific technical and organisational measures;
- (e) Their access to the data is necessary for the purposes of their research;
- (f) Their research is for the purpose of the detection, identification, and understanding of specific risks to the EU or for the assessment of the adequacy, efficiency and impacts of the risk mitigation measures of very large online platforms and very large online search engines; and
- (g) They make their research results public, free of charge, within a reasonable period after their research is completed, subject to the rights and interests of the recipients of the service concerned.

The procedure to underpin this law is being developed in subordinate legislation.

The specific risks to the EU to which the purpose of research may relate are:<sup>6</sup>

- (a) The dissemination of illegal content;
- (b) Actual or foreseeable negative effects on fundamental rights enshrined in the Charter of Fundamental Rights of the European Union;
- (c) Actual or foreseeable negative effects on civic discourse and electoral processes, and public security; and

---

<sup>4</sup> Single Market For Digital Services and amending Directive 2000/31/EC, Regulation (EU) 2022/2065 of the European Parliament and of the Council, 19 October 2022 art 40(4).

<sup>5</sup> Ibid art 40(8).

<sup>6</sup> Ibid art 34(1).



- (d) Actual or foreseeable negative effects in relation to gender-based violence, the protection of public health and minors, and serious negative consequences to the person's physical and mental well-being.

We note these categories are broad and not well-defined because the DSA, as European legislation, leaves that definitional work to member states.

Providers of very large online platforms and very large online search engines are required to give vetted researchers access without undue delay to data (known as the "crowdtangle provision").<sup>7</sup>

### ***Adaption to Australian context - Definitions***

For researchers to assist, they need a working definition of what constitutes hatred across various contexts, noting that hate campaigns tend to be cyclical and rotational and require expertise relevant to each context. However, the standards applied must be universal and identical across contexts to manage community disagreement about the definitions of various forms of racism and other prejudice. Australia's various definitions of vilification and offensive hate speech do not provide enough guidance to administrators or researchers applying this regulation, as those definitions are more generous, focused on effect and lend themselves to deliberative processes of evidence review involved in judicial consideration. Our work on a definition for dehumanising material attempts to respond to these policy needs.

### ***Hatred as a harm***

AMAN strongly supports the Bill's recognition of hatred against a group in Australian society based on protected characteristics as constituting harm without the need to prove incitement to hatred or violence. Online hatred, particularly in the form of dehumanising material, is a harm in itself that should be treated as a form of psychological violence. Its effects on targeted groups include:

- limits on the fulfilment of other fundamental human rights, including participation in public life and freedom of expression;
- the tendency to internalise dehumanising messages and harmful stereotypes; and
- the tendency to accept or normalise inferior status or rights.

### ***Vilification, dehumanisation, and hate speech***

While the proposed Bill is a welcome addition to the fight against misinformation and disinformation, it is also a systemic method of mitigating vilification, hate speech, and dehumanisation. Misinformation and disinformation are typical means through which dehumanising material is conveyed and disseminated.

---

<sup>7</sup> Ibid art 40(12).

Digital platforms have already expressed a view that they are not subject to Australia's vilification laws or standards without express provision, and this Bill contains insufficient standards and incentives for them to improve their systems.

The Bill must require ACMA to immediately set a standard of what constitutes hateful or dehumanising material to prevent harm online. Our recommendations also contain suggestions for E-Safety takedown powers and the Attorney Generals' Department in strengthening vilification laws.

### ***Severeness and severity***

Assessing the severity of harm to meet the element of seriousness necessarily requires including factors, or a combination of types of harm, to determine the severity and a robust definition of hatred. The Rabat Plan of Action, developed in collaboration with the Office of the High Commissioner for Human Rights in response to post-electoral violence, extremism, hate speech, and malicious portrayal of certain religions, notes that an overly broad definition of incitement to hatred 'opens the door for arbitrary application' of the law.<sup>8</sup> The Rabat Plan of Action recommends the Camden Principles on Freedom of Expression and Equality to guide states in creating a robust definition of hatred. The Camden Principles define hatred as 'intense and irrational emotions of opprobrium, enmity and detestation towards the target group.'<sup>9</sup> Exceptions are permitted to promote a positive sense of group identity and to criticise or debate about particular ideologies or religions unless it constitutes hate speech.

The Bill contains a list of factors determining seriousness and severity, which appear to align with the Rabat Plan of Action. However, the purpose of that clause could be better clarified and linked to the definition of harm.

### ***Harm category- 'disruption of public order or society in Australia.'***

The definition of harm is too broad in respect to the 'disruption of public order or society in Australia' type of harm.<sup>10</sup> The term disruption is unnecessarily vague and prone to various interpretations that could see the arbitrary enforcement of the proposed Bill to censor fringe groups within Australian society. Disruption is an underlying component of civil discourse and essential in societal reform. In this sense, the phrasing interferes with the implied freedom of political communication. While the proposed Bill contains safeguards against any such interference,<sup>11</sup> it is necessary to alter the phrasing to provide sufficient certainty.

### ***The exception for electoral and referendum matters***

---

<sup>8</sup> Report of the United Nations High Commissioner for Human Rights on the expert workshops on the prohibition of incitement to national, racial or religious hatred, 22<sup>nd</sup> sess, Agenda Item 2, A/HRC/22/17/Add.4 (11 January 2013) 8 [15].

<sup>9</sup> Article 19, *The Camden Principles on Freedom of Expression and Equality* art 10.

<sup>10</sup> Communications Legislation Amendment (Combatting Misinformation and Disinformation) Bill 2023 (Cth), Exposure Draft (definition of 'harm').

<sup>11</sup> *Ibid* cl 60.

A permissible exception to the registration of misinformation codes under the Bill is where they contain requirements relating to electoral and referendum content.<sup>12</sup> AMAN welcomes the protection of the implied right of political communication and promulgates freedom of speech principles where they do not enable discrimination or vilification. On this basis, we note that Article 19, a preeminent international human rights organisation specialising in promoting freedom of expression, has developed a set of guiding principles that seek to balance equality and freedom of expression and are worth consulting. The Camden Principles on Freedom of Expression and Equality note that states must impose upon public officials an obligation to avoid, as far as possible, ‘making statements that promote discrimination or undermine equality and intercultural understanding.’<sup>13</sup> It further recommends that politicians and leadership figures avoid making statements that could promote discrimination or undermine equality.<sup>14</sup> Recognition of the Camden Principles and these balances should be made in the Explanatory Memorandum.

The Bill must explain and define ‘election and referendum matters’ given its exemption from the Bill and ensure that online content that causes harm defined by the Bill is not exempt from the Bill *unless* it is covered by AEC legislation. Then list the relevant AEC legislation and provisions in the Bill. The AEC does not contemplate harms included in this Bill, and therefore, electoral and referendum matters should not be excised wholesale. Consider the re-emergence of anti-Islam political parties that curate and publish misinformation and disinformation about Australian Muslims’ beliefs, intentions and conduct, creating hatred against Muslims. This online content should not be excluded from the Bill. The interaction of the Bill with AEC legislation should be determined with reference to harms first.

### ***Implied freedom of political communication***

Section 60 of the proposed Bill seeks to reconcile the Bill with the implied freedom of political communication by limiting the effect of the bill to the extent that it is incompatible with the implied freedom. To accord with the implied freedom most effectively, including specific phrasing harkening to the Lange test may be pertinent. Section 45(b) could be rephrased from ‘whether the burden would be reasonable and not excessive ...’<sup>15</sup> to ‘whether the burden would be reasonably appropriate and adapted’.

## **EXEMPTION OF PROFESSIONAL NEWS CONTENT**

### ***Exemption/definition of ‘professional news content’ significantly undermines objects of this Bill.***

---

<sup>12</sup> Ibid cl 35.

<sup>13</sup> Article 19, *The Camden Principles on Freedom of Expression and Equality* art 8.1.

<sup>14</sup> Ibid art 10.1.

<sup>15</sup> Communications Legislation Amendment (Combatting Misinformation and Disinformation) Bill 2023 (Cth), Exposure Draft cl 45(b).

The proposed Bill provides, in section 7, an exemption for the offences of disseminating content using a digital service that is misinformation or disinformation where that content is excluded content for misinformation purposes. Excluded content includes professional news content,<sup>16</sup> being news content subject to some standards of practice, codes of conduct, and rules.

These rules are entirely inadequate at addressing issues of accountability and transparency in news reporting, except in the fairly universal requirement to disclose any conflicts of commercial interests. The existing rules are not an appropriate substitute for the protections provided in the proposed Bill as they do not consider misinformation and disinformation and are far too old to adequately respond to the new digital platform and online media landscape.

### ***The current framework regulating misinformation, disinformation, and the transparency and accountability of news publishers and sources***

The Australian Press Council (**APC**), being the principal body with responsible for responding to complaints about material published by Australian newspapers, magazines, and online media outlets, has its own Standards of Practice to guide news publishers.<sup>17</sup> These standards of practice do not address transparency and accountability of news reporting and journalism except by disclosure of conflicts of interest. Given that the only explicit requirement for news publishers to become constituent bodies that are protected from the Bill by the APC's Standards of Practice is the provision of funding to the APC, there are very concerning prospects that paid constituency with the APC could be exploited to avoid liability under the Bill.

Use of the Wayback Machine indicates that the Independent Media Council's (**IMC**) Code of Conduct has not been updated since at least 23 May 2014.<sup>18</sup> The IMC has not released any findings since 14 December 2022, and no annual reports since 31 January 2022. There is a serious concern as to whether the IMC is even currently active and, given its own lack of transparency as to its operations and owners, is an extremely inappropriate choice for a body maintaining a rule that should exempt a news source from the application of the Bill.

The Commercial Television Industry Code of Practice commenced on 1 December 2015.<sup>19</sup> It makes minor reference to concepts of accuracy as to material errors of fact, misrepresentation of viewpoints, and rectification of these errors. Rectification can be done up to 30 days from the error and may be made on the Program's website where the error was aired on the Program. Publication on the website will inevitably not target

---

<sup>16</sup> Communications Legislation Amendment (Combatting Misinformation and Disinformation) Bill 2023 (Cth), Exposure Draft s 1 (definition of 'excluded content for misinformation purposes').

<sup>17</sup> Australian Press Council, 'Statement of Principles', *Statement of Principles* (Web Page) <<https://presscouncil.org.au/standards/statement-of-principles/>>.

<sup>18</sup> Independent Media Council, 'Code of Conduct', *Way Back Machine* (Historical Web Page) <[https://web.archive.org/web/20131201000000\\*/http://www.independentmediacouncil.com.au/](https://web.archive.org/web/20131201000000*/http://www.independentmediacouncil.com.au/)>.

<sup>19</sup> Commercial Television Industry: Code of Practice (pdf file) <[https://www.freetv.com.au/wp-content/uploads/2019/07/Free\\_TV\\_Commercial\\_Television\\_Industry\\_Code\\_of\\_Practice\\_2018.pdf](https://www.freetv.com.au/wp-content/uploads/2019/07/Free_TV_Commercial_Television_Industry_Code_of_Practice_2018.pdf)>.

the same audience as the Program's viewers. The Code only lightly addresses issues of misinformation, disinformation, and journalistic transparency and accountability, making it an inadequate replacement for the Bill.

The Commercial Radio Code of Practice commenced on 15 March 2017 and suffers inadequacies tantamount to the Commercial Television Industry Code of Practice.<sup>20</sup>

The Australian Subscription Television and Radio Association's Subscription Broadcast Television Codes of Practice have not been revised since 2013 and do not address misinformation, disinformation, and journalistic transparency and accountability issues.<sup>21</sup>

The *Australian Broadcasting Corporation Act 1983* (Cth)<sup>22</sup> and the *Special Broadcasting Service Act 1991* (Cth)<sup>23</sup> are only binding on the ABC and SBS, respectively.

Internal rules and editorial standards are self-regulatory and fundamentally inadequate means of ensuring transparency and accountability. Contrastingly to the other rules applicable to professional news content, internal rules and editorial standards are often not publicly available. It will be excessively unclear whether internal standards are in place when determining whether the exemption for disseminating content using a digital service is misinformation or disinformation. Without a body to oversee adherence to internal rules, they lack the enforceability paramount in creating quality journalism standards. The requirement that internal rules be analogous to the abovementioned rules is also not affirming, given their own substantial shortcomings.

### ***Sky News example***

AMAN queries the applicability of the definition of 'professional news content' to existing news sources. Sky News Australia, for example, is not a constituent body to which the Australian Press Council Standards of Practice rules apply, nor are they funding bodies to which the Independent Media Council's Code of Conduct apply. We still cannot confirm whether Sky News Australia has an editorial code. If they do, it is not public.<sup>24</sup> While the rules of the Commercial Television Industry Code of Practice or the Subscription Broadcast Television Codes of Practice may apply to news broadcasts on television, no rules apply to news disseminated over the Internet. An internal

---

<sup>20</sup> Commercial Radio Australia, 'Commercial Radio Code of Practice', (pdf file)

<<https://www.commercialradio.com.au/CR/media/CommercialRadio/Commercial-Radio-Code-of-Practice.pdf>>.

<sup>21</sup> Australian communications and Media Authority, 'Subscription broadcast television codes of practice 2013', (Web Page) <<https://www.acma.gov.au/publications/2019-10/rules/subscription-broadcast-television-codes-practice-2013>>.

<sup>22</sup> *Australian Broadcasting Corporation Act 1983* (Cth); ABC, 'Code of Practice 2023', (Web Page, 8 May 2023) <<https://about.abc.net.au/reports-publications/code-of-practice/>>.

<sup>23</sup> *Special Broadcasting Service Act 1991* (Cth); SBS, 'SBS Code of Practice' (Web Page, 4 April 2022) <<https://www.sbs.com.au/aboutus/sbs-code-of-practice>>.

<sup>24</sup> Sky News (UK and Ireland) have public Editorial Guidelines from 2015: [https://news.sky.com/docs/sky\\_news\\_editorial\\_guidelines.pdf](https://news.sky.com/docs/sky_news_editorial_guidelines.pdf)

editorial standard analogous to the discussed rules may be applicable. However, it is not publicly viewable if so. Without digital platform rules in place at the enactment of the Bill, news content providers may be vulnerable to committing an offence under the Bill.

### ***The need for operational transparency and accountability and the importance of the Bill's misinformation and disinformation provisions***

In the ACCC's Digital Platform Services Inquiry from February 2022, the ACCC recommended a potential new regulatory framework to address digital platform services in Australia. The advantage of such a framework was described as 'sufficient legal certainty for market participants and ... flexible enough to adapt to the dynamic and fast-moving nature of digital platform services ...'<sup>25</sup> This, too, is the strength of the Bill and its adaptable misinformation codes. This strength is not present in existing rules applicable to news sources and publishers, evident by their stagnation and failure to adapt to ongoing concerns regarding transparency and accountability. These rules are incapable of addressing misinformation and disinformation issues, as well as the ever-changing nature of digital platforms.

AMAN understands that the intention here was to extend the definition of professional news content used in the News bargaining code. However, we submit that the exemption of professional news content, as currently defined, will result in the most extensive operations of misinformation and disinformation being left untouched by this legislation. Consider the misinformation spread by Peta Credlin about the Uluru Statement of the Heart on Sky News under the guise of 'current affairs' and 'opinion', which spread a significant amount of fear and confusion among Australian audiences.

The exemption for professional news content should be removed. The existing regulatory framework is incapable of addressing the issues the Bill addresses. Alternatively, the definition of professional news content should be expanded to include operational transparency and accountability requirements, including transparency regarding editorial standards, complaints handling, and funding sources. Such requirements can be measured and sourced from the Global Disinformation Index pillars and indicators,<sup>26</sup> attached as **Schedule 1**. A proposed definition for professional news content is included in **Schedule 2**.

### **Private messaging and freedom of expression**

AMAN supports the proposed limitation of digital platform rules, which would restrict ACMA from creating rules requiring the retention of the content of private messages<sup>27</sup> and from requiring the production of anything that would reveal the content of a private

---

<sup>25</sup> ACCC, *Digital Platform Services Inquiry* (Interim Report No 5) 77.

<sup>26</sup> Solya Glazunova, Ehsan Deghan and Katherine FitzGerald, *Disinformation Risk Assessment: The Online News Market in Australia* (Report, September 2021) 19 item 2.

<sup>27</sup> Communications Legislation Amendment (Combatting Misinformation and Disinformation) Bill 2023 (Cth), Exposure Draft cls 14(3), 34.

message.<sup>28</sup> It is still within ACMA's powers to require the provision of information or documents from private messaging services that do not reveal the content of private messages. This appears to offer a good balance between freedom of expression and reducing misinformation and disinformation.

### **Miscellaneous**

Clause 19 (1)(b) provides that ACMA may obtain information and documents from other persons if ACMA considers that it requires the information, document or evidence for the performance of the ACMA's function under paragraph 10(1)(md) of the *Australian Communications and Media Authority Act 2005*. No such paragraph is found within that Act.

---

<sup>28</sup> Ibid cls 18(4), 19(4).

## SCHEDULE 1

Pillar	Indicator	Sub-indicators	Unit of analysis	Definition	Rationale
Content	Headline accuracy	None	Article	Rating for how accurately the story's headline describes the content of the story	Indicative of clickbait
	Byline information			Rating for how much information is provided in the article's byline	Attribution of stories creates accountability for their veracity
	Lede present			Rating for whether the article begins with a fact-based lede	Indicative of fact-based reporting and high journalistic standards
	Common coverage			Rating for whether the same event has been covered by at least one other reliable local media outlet	Indicative of a true and significant event
	Recent coverage			Rating for whether the story covers a news event or development that occurred within 30 days prior to the article's publication date	Indicative of a newsworthy event, rather than one which has been taken out of context
	Negative targeting			Rating for whether the story negatively targets a specific individual or group	Indicative of hate speech, bias or an adversarial narrative
	Article bias			Rating for the degree of bias in the article	Indicative of neutral, fact-based reporting or well-rounded analysis
	Sensational language			Rating for the degree of sensationalism in the article	Indicative of neutral, fact-based reporting or well-rounded analysis
	Visual presentation			Rating for the degree of sensationalism in the visual presentation of the article	Indicative of neutral, fact-based reporting or well-rounded analysis
Operations	Attribution	None	Domain	Rating for the number of policies and practices identified on the site	Assesses policies regarding the attribution of stories, facts and media (either publicly or anonymously); indicative of policies that ensure accurate facts, authentic media and accountability for stories
	Comment policies	Policies		Rating for the number of policies identified on the site	Assesses policies to reduce disinformation in user-generated content
		Moderation		Rating for the mechanisms to enforce comment policies identified on the site	Assesses the mechanism to enforce policies to reduce disinformation in user-generated content
	Editorial principles and practices	Editorial independence		Rating for the number of policies identified on the site	Assesses the degree of editorial independence and the policies in place to mitigate conflicts of interest
		Adherence to narrative		Rating for the degree to which the site is likely to adhere to an ideological affiliation, based on its published editorial positions	Indicative of politicised or ideological editorial decision making
		Content guidelines		Rating for the number of policies identified on the site	Assesses the policies in place to ensure that factual information is reported without bias
		News vs. analysis		Rating for the number of policies and practices identified on the site	Assesses the policies in place to ensure that readers can distinguish between news and opinion content
	Ensuring accuracy	Pre-publication fact-checking		Rating for the number of policies and practices identified on the site	Assesses policies to ensure that only accurate information is reported
		Post-publication corrections		Rating for the number of policies and practices identified on the site	Assesses policies to ensure that needed corrections are adequately and transparently disseminated
	Funding	Diversified incentive structure		Rating for the number of revenue sources identified on the site	Indicative of possible conflicts of interest stemming from over-reliance on one or few sources of revenue
		Accountability to readership		Rating based on whether reader subscriptions or donations are identified as a revenue source	Indicative of accountability for high-quality information over content that drives ad revenue
		Transparent funding		Rating based on the degree of transparency the site provide regarding its sources of funding	Indicative of the transparency that is required to monitor the incentives and conflicts of interest that can arise from opaque revenue sources
	Ownership	Owner-operator division		Rating based on the number of distinct executive or board level financial and editorial decision-makers listed on the site	Indicative of a separation between financial and editorial decision making, to avoid conflicts of interest
Transparent ownership		Rating based on the degree of transparency the site provides regarding its ownership structure	Indicative of the transparency that is required to monitor the incentives and conflicts of interest that can arise from opaque ownership structures		

Solya Glazunova, Ehsan Deghan and Katherine FitzGerald, *Disinformation Risk Assessment: The Online News Market in Australia* (Report, September 2021) 19 item 2.



## SCHEDULE 2

### Possible improved wording for 'professional news content' definition

Professional news content produced by a news source who

(a) Is subject to

- i. The rules of the Commercial Television Industry Code of Practice, the Commercial Radio Code of Practice or the Subscription Broadcast Television Codes of Practice; or
- ii. Rules of code of practice mentioned in paragraph 8(1)(e) of the Australian Broadcasting Corporations Act 1983 or paragraph 10(1)(j) of the Special Broadcasting Services Act 1991; and

(b) Is subject to internal editorial standards that

- i. Relate to the provision of quality journalism;
- ii. Ensure that factual information is reported without bias;
- iii. Implement labels that assist readers and audiences in distinguishing between news and opinion content;
- iv. Require diversity of opinion on controversial issues;
- v. Require pre-publication fact-checking and post-publication corrections that are adequately and transparently disseminated;
- vi. Prohibit material that is hateful or incites hatred against individuals or groups on the basis of protected characteristics;
- vii. Are published on its website and easily accessible; and
- viii. Provide an electronic email address and postal address for complaints.

(c) Publishes current information on their website that

- i. Provides transparency as to its sources of funding; and
- ii. Provides transparency as to the number of executive or board-level financial and editorial decision-makers so that the public can identify possible conflicts of interest.

(d) Has editorial independence from the subjects of the news source's news coverage.

### **SCHEDULE 3**

#### **AMAN's working definition of dehumanising material, updated 15 July 2023**

Note this definition is subject to ongoing revision until it is formally published.

(1) Dehumanising material is the material produced or published, which an ordinary person would conclude, portrays the class of persons identified on the basis of a protected characteristic ("class of persons") as not deserving to be treated equally to other humans because they lack qualities intrinsic to humans. Dehumanising material includes portraying the class of persons:

(a) to be or have the appearance, qualities, or behaviour of

(i) an animal, insect, filth, form of disease or bacteria;

(ii) inanimate or mechanical objects; or

(iii) a supernatural alien or demon.

(b) are polluting, despoiling, or debilitating an ingroup or society as a whole;

(c) have a diminished capacity for human warmth and feeling or to make up their own mind, reason or form their own individual thoughts;

(d) homogeneously pose a powerful threat or menace to an in-group or society, posing overtly or deceptively;

(e) are to be held responsible for and deserving of collective punishment for the specific crimes, or alleged crimes of some of their "members";

(f) are inherently criminal, dangerous, violent or evil by nature;

(g) do not love or care for their children;

(h) prey upon children, the aged, and the vulnerable;

(i) was subject as a group to past tragedy or persecution that should now be trivialised, ridiculed, glorified or celebrated;

(j) are inherently primitive, coarse, savage, intellectually inferior or incapable of achievement on a par with other humans;

(k) must be categorised and denigrated according to skin colour or concepts of racial purity or blood quantum; or

(l) must be excised or exiled from public space, neighbourhood or nation.

(2) Without limiting how the material in section (1) is presented, forms of presentation may include,

(a) speech or words;

(b) the curation or packaging of information;

(c) images; and

(d) insignia.

### ***Intention component***

If the above definition was used as a standalone civil penalty, it should be complemented by an intention component:

*in circumstances in which a reasonable person would conclude that the material was intended to portray the class of persons as not deserving to be treated equally to other humans or to incite hatred, serious contempt or severe ridicule toward the class of persons.*

Adding an intention element may make enforcement more difficult and may not be necessary, especially if the definition is used as part of a legal framework where there are already intention components or exceptions available.

### **How did we develop this working definition?**

AMAN developed this working definition after spearheading a study of five information operations online (Abdalla, Ally and Jabri-Markwell, 2021). The first iteration of this definition was published in a joint paper with UQ researchers (Risius et al, 2021). It continues to be developed with input received from researchers, lawyers and civil society.

Possible dehumanising conceptions are surfaced through research and then tested against [Haslam](#)'s frame of whether it deprives a group of qualities that are intrinsic to humans.

If a subject is dehumanised as a mechanistic form, they are portrayed as 'lacking in emotionality, warmth, cognitive openness, individual agency, and, because [human nature] is essentialized, depth.' A subject that is dehumanised as animalistic, is portrayed as 'coarse, uncultured, lacking in self-control, and unintelligent' and 'immoral or amoral' (258).

Some conceptions are found to fall outside the frame of dehumanisation but could still qualify as vilification or discrimination, for example, using anti-discrimination laws.

The three categories of dehumanising comparisons or metaphors in Clause (a) are drawn from [Maynard and Benesch](#) (80), and fleshed out with further examples from tech company policies (refer to Meta for example).

Clause (b) is derived from Maynard and Benesch (80).

Clause (c) is derived from [Haslam](#) (258).

Clauses (d) and (e) are elements of dangerous speech that Maynard and Benesch refer to as 'threat construction' and 'guilt attribution' respectively (81). However, [Abdalla, Ally and Jabri-Markwell's](#) work shows how such conceptions are also dehumanising, as they assume a group operates with a single mindset, lacking independent thought or human depth (using Haslam's definition), and combine with ideas that Muslims are inherently violent, barbaric, savage, or plan to infiltrate, flood, reproduce and replace (like disease, vermin)(15). The same study found that the melding and flattening of Muslim identities behind a threat narrative through headlines over time was a dehumanisation technique (17). Demographic invasion theory-based memes (9) or headlines that provided 'proof' for such theory (20) elicited explicit dehumanising speech from audiences.

Maynard and Benesch write, 'Like guilt attribution and threat construction, dehumanization moves out-group members into a social category in which conventional moral restraints on how people can be treated do not seem to apply' (80).

Clauses (f), (h), (i) are drawn from the "Hallmarks of Hate", which were endorsed by the Supreme Court of Canada in *Saskatchewan (Human Rights Commission) v. Whatcott* 2013 SCC 11, [2013] 1 S.C.R. 467. These Hallmarks of Hate were developed after reviewing a series of successful judgements involving incitement of hatred to a range of protected groups. These clauses were tested using Haslam's definitional frame for the denial of intrinsic human qualities.

Clauses (f) ('criminal') and (g) are drawn from harmful characterisations cited in the Uluru Statement of the Heart.

Clauses (j) and (k) are drawn from AMAN's observations of online information operations generating disgust toward First Nations Peoples. Disgust is a common effect of dehumanising discourse. These clauses were tested using Haslam's definitional frame for the denial of intrinsic human qualities.

Clause (l) was drawn from Nicole Asquith's Verbal and Textual Hostility Framework. (Asquith, N. L. (2013). The role of verbal-textual hostility in hate crime regulation (2003, 2007). Violent Crime Directorate, London Metropolitan Police Service.) The data and process used to formulate this Framework is exceptional. Reassuringly, this research had surfaced examples that were already captured by this Working Definition of Dehumanising Material.

This working definition is a work in progress. AMAN welcomes feedback as it continues to be developed.

*Updated 15 July 2023*