HUMAN RIGHTS LAW CENTRE - NOTE ON THE *COMMUNICATIONS LEGISLATION AMENDMENT (COMBATTING MISINFORMATION AND DISINFORMATION) BILL 2024*

1 Scope of note

This note compares the *Communications Legislation Amendment (Combatting Misinformation and Disinformation) Bill 2024* (Cth) ("**Bill**") with the principles outlined by the Human Rights Law Centre ("**HRLC**") for effective regulation of misinformation and disinformation ("**Principles**").[1] It includes our assessment of areas where the Bill and the Principles align and diverge, and our views on some amendments considering this comparison.

2 Summary of our assessment

Set out in the table below is our assessment of the degree of alignment between the Bill and the Principles. A more detailed analysis is included in section 4.

#	PRINCIPLE	DEGREE OF ALIGNMENT
1	Digital regulation must be based on human rights law and principles	Mild
2	Digital platforms should have a legal duty to make sure their products, systems, and services do not cause harm	Mild
3	Content removal powers have a role to play in limited circumstances but only as part of a broad, comprehensive regulatory framework	Mild
4	Users should have control over how platforms collect and use their data	Low
5	Court oversight is essential in a comprehensive regulatory framework	Moderate

3 Overview of the Bill

The Bill proposes to amend the *Broadcasting Services Act 1992* (Cth) ("**BSA"**) in response to growing concerns about the spread of misinformation and disinformation in Australia.[2]

3.1 Objectives of the Bill

The Bill has three key objectives:[3]

1 to empower the Australian Communications and Media Authority ("**ACMA**") to require digital communications platform providers to take steps to manage the risk that misinformation and disinformation on digital communications platforms poses in Australia;

2 to increase transparency regarding the way in which digital communications platform providers manage misinformation and disinformation; and

[1] David Mejia-Canales, Human Rights Law Centre, *Rights-First: Principles for Digital Platform Regulation*, 1 September 2024 ("**Principles**").

[2] *Broadcasting Services Act 1992* (Cth) ("**BSA**").

[3] *Communications Legislation Amendment (Combatting Misinformation and Disinformation) Bill 2024*, Explanatory Memorandum ("**EM**") 1.

3 to empower users of digital communications platforms to identify and respond to misinformation and disinformation on digital communications platforms.

3.2 Obligations imposed by the Bill

The Bill imposes core obligations on digital communications platform providers to:

1 assess risks relating to misinformation and disinformation on their platforms, and publish a report of the outcomes of that assessment;

2 publish their policy or policy approach in relation to managing misinformation and disinformation; and

3 publish a media literacy plan setting out the measures the provider will take to enable end-users of the platform to better identify misinformation and disinformation.[4]

4 Alignment between the Bill and the Principles

4.1 Principle 1: Digital regulation must be based on human rights law and principles

Summary: The Bill mildly aligns with this principle. While the Bill aligns with human rights law and principles in its general aim to reduce harmful misinformation and disinformation, it also arguably risks, in some of its measures, disproportionately curtailing those same principles.

(a) The right to freedom of thought and conscience

On the one hand, the Bill diverges from this right given it lacks any explicit consideration of it. On the other hand, given the reference in the Principles to the way in which misinformation may hinder this right, it is arguable that both the right to freedom of expression and the right to participate in public affairs including the right to vote, are related, both of which are considered by the Bill (see paragraphs 4.1(b) and 4.1(d)).

Conclusion: mild degree of alignment

(b) The right to freedom of expression

The Principles refer to Article 19 of the *International Covenant on Civil and Political Rights* ("**ICCPR**"), which allows people to seek, receive, and share information and ideas through any medium.[5] It includes the right to seek, receive and impart information that may be 'deeply offensive',[6] and the right to seek, receive and impart information irrespective of the truth or falsehood of the content (subject to the restrictions permitted by Articles 19(3) and 20 of the ICCPR).[7] In Australia, the right to freedom of expression is principally protected by the implied freedom of political communication, which is inferred from ss 7, 24 and 128 of the Constitution.[8]

The Explanatory Memorandum ("**EM**") notes that freedom of expression is a "human right potentially affected by the Bill" due to the restrictions the Bill places on it. Relevantly, the Principles state that:

> "*Under Article 19, any restriction on freedom of expression must meet the following criteria: it must be:*
>
> *(i) clearly defined by law,*
>
> *(ii) aimed at protecting the rights or reputation of others, or safeguarding national security, or public order, or public health, or public morals, and*
>
> *(iii) necessary and proportionate.*

[4] EM 5.

[5] Principles 10.

[6] Human Rights Committee ("**HRC**"), General Comment No 34: Article 19: Freedom of Expression, UN Doc CCPR/C/GC/34 (12 September 2011) para 11; HRC, Views: Communication No 736/97, UN Doc CCPR/C/780/D/736/1997.

[7] *Handyside v United Kingdom* (European Court of Human Rights, Application No 5493/72, 7 December 1976) para 49.

[8] See, e.g., *Monis v R; Droudis v R* (2013) 295 ALR 259.

> *A misguided interpretation of the freedom of expression (including by free speech absolutists) has been weaponised to avoid accountability for the harms caused by abuses of the freedom of expression. However, it is also important to recognise that over-regulation, especially when it is not grounded in human rights law, can itself pose a legitimate threat to freedom of expression. Excessive or poorly designed regulations may suppress free speech, stifle public debate, or lead to censorship that is neither necessary nor proportionate*."[9]

The Bill moderately aligns with this approach taken by the Principles, however, there is a reasonable degree of risk it may suppress free speech, stifle public debate, or lead to censorship that is neither necessary nor proportionate.

Limb (i): clearly defined by law

This limb is satisfied because the measures set out in Sch 9 to the BSA are either prescribed in the schedule itself, or will be prescribed in digital platform rules,[10] approved misinformation codes or misinformation standards.

Limb (ii): pursuit of a legitimate aim

The overarching aim of the measures in the Bill is to address "the risk posed by misinformation and disinformation; thus they aim to reduce the risk that content disseminated on digital communications platforms will cause or contribute to one of the harms listed in clause 14".[11] The EM states that these harms align with the purposes for which, pursuant to Article 19(3) of the ICCPR, the freedom of expression may be restricted, which are shown in the table below.

CLAUSE 14: TYPE OF HARM[12]	ALIGNED PURPOSE FOR WHICH FREEDOM OF EXPRESSION MAY BE RESTRICTED, PURSUANT TO ARTICLE 19(3)[13]
(a) Harm to the operation or integrity of a Commonwealth, State, Territory or local government electoral or referendum process	The rights of others. Specifically, the right to take part in the conduct of public affairs, and to vote and be elected at genuine periodic elections, enshrined in Article 25 of the ICCPR. (See EM, pp 8-10).
(b) Harm to public health in Australia	Public health. (See EM, pp 13-15).
(c) Vilification of a group in Australian society distinguished by race, religion, sex, sexual orientation, gender identity, intersex status, disability, nationality or national or ethnic origin, or vilification of an individual because of a belief that the individual is a member of such a group	The rights of others. Specifically, the right to be protected against discrimination, enshrined in numerous international human rights instruments. (See EM, pp 10-13).
(d) Intentionally inflicted physical injury to an individual in Australia	The rights of others. Specifically, the right to security of the person, enshrined in Article 9 of the ICCPR.[14] (See EM, pp 7-8).

[9] Principles 10.

[10] Bill cl 82 relevantly defines digital platform rules as rules, made by legislative instrument, prescribing matters necessary or convenient to be prescribed for carrying out or giving effect to Sch 1 of the Bill ("***Digital Platform Rules***").

[11] EM 19.

[12] EM 19-20.

[13] ICCPR Article 19(3).

[14] ICCPR Article 9.

CLAUSE 14: TYPE OF HARM[12]	ALIGNED PURPOSE FOR WHICH FREEDOM OF EXPRESSION MAY BE RESTRICTED, PURSUANT TO ARTICLE 19(3)[13]
(e) Imminent damage to critical infrastructure or disruption of emergency services in Australia	Public order. The Human Rights Committee has not explicitly defined public order for the purposes of Article 19(3) of the ICCPR; however, in relation to a similar clause describing permissible restrictions on the right of peaceful assembly, the Committee has said that 'public order' refers to 'the sum of the rules that ensure the proper functioning of society, or the set of fundamental principles on which society is founded'.[15] The protection of critical infrastructure and emergency services are necessary to 'ensure the proper functioning of society'.
(f) Imminent harm to the Australian economy	Public order. See preceding comment. Protection against imminent harm to the Australian economy, such as the destabilisation of the banking system or financial markets, is necessary to 'ensure the proper functioning of society'.

Limb (iii): Necessary and proportionate

Alignment

First, the EM describes the Bill as being "*focussed on systems and processes rather than the regulation of actual content*".[16] This is captured by both the core transparency obligations on platforms and the ACMA's regulatory powers. Clause 67 also reflects this, which provides that nothing in Pt 2 of Sch 9, or in a rule, approved code, or standard made, approved or determined pursuant thereto, can require a platform provider to remove content from a platform or prevent an end-user from using the platform, except in the case of disinformation that involves inauthentic behaviour (discussed at paragraphs 4.1(b)(vii) and 4.3(a)).

Second, there are a few other protections to avoid a "chilling effect" on the freedom of expression online, namely:

(i) clause 16 provides exemptions for: the dissemination of content that would reasonably be regarded as parody or satire; the dissemination of professional news content; and the reasonable dissemination of content for any academic, artistic, scientific or religious purpose.

(ii) subclauses 33(3), 34(4) and 30(3), and clauses 45 and 46, provide that the ACMA must not exercise its information-gathering powers, powers to require digital communications platform providers to make and retain records, or powers to approve a code or determine a standard, in relation to private messages or Voice over Internet Protocol ("**VoIP**") communications.

(iii) subclause 34(2) provides that in exercising its information-gathering powers, the ACMA must not require a person to provide information or documents relating to content posted by that person on a digital communications platform, other than in the person's capacity as a fact-checker, a content moderator, an employee of the provider of the platform, or a person providing services to the provider of the platform.

(iv) in exercising its power to approve, or approve variations of, misinformation codes or determine or vary misinformation standards, the ACMA must be satisfied that: the code or

[15] Human Rights Committee, *General Comment No 37 on Article 21 (Right of Peaceful Assembly),* UN Doc CCPR/C/GC/37 (17 September 2020) para 44. See also UN Economic and Social Council (ECOSOC), *Siracusa Principles on the Limitation and Derogation Provisions in the International Covenant on Civil and Political Rights*, UN Doc E/CN.4/1985/4, Annex (1995) para 22; Australian Government Attorney-General's Department (AGD), Right to Freedom of Opinion and Expression (Public Sector Guidance Sheet) <Right to freedom of opinion and expression | Attorney-General's Department (ag.gov.au)>.

[16] EM 20.

standard is reasonably appropriate and adapted to achieving the purpose of providing adequate protection for the Australian community from serious harm caused or contributed to by misinformation or disinformation on the platforms; and goes no further than reasonably necessary to provide that protection (subparas 47(1)(d)(iii) and (iv), subparas 50(1)(d)(iii) and (iv), clause 54, and subclause 60(2)).[17]

(v) clause 12 exempts certain digital communications platforms from the substantive requirements of Pt 2 of Sch 9, from Digital Platform Rules made for the purposes of that Part, and from approved misinformation codes and misinformation standards. Those provisions and instruments would not apply in relation to a digital communications platform to the extent that it is one of the following: an email service, a media sharing service that does not have an interactive feature, or a digital service the Minister determines is an excluded service for misinformation purposes under subclause 12(3).[18]

Divergence

Conversely, the Bill diverges with the right to freedom of expression due to the risks it poses to the right, summarised by the EM as follows:

> "*Schedule 9 to the BSA requires digital communications platform providers to be more transparent about the way in which they manage the risk of misinformation and disinformation on their platforms, and empowers the ACMA to require providers to take steps in relation to the risk. These measures could feasibly incentivise digital communications platform providers to take an overly cautious approach to the regulation of content that could be regarded as misinformation and disinformation – or in other words, they could have a 'chilling effect.' Thus, these measures could burden the freedom of expression.*"[19]

We suggest this risk could be reduced by amendments to the definitions of 'misinformation' and 'disinformation'.

(vi) **Misinformation**

Relevantly, *misinformation* is defined as content that:

(A) contains information that is reasonably verifiable as false, misleading or deceptive, and

(B) the provision of such content is reasonably likely to cause or contribute to serious harm.[20]

The term 'cause' implies a higher threshold of causality than the term 'contribute to', which raises a statutory interpretation question about the purpose of the latter.

Further, while the term 'information' is not defined in the Bill, the EM explains that it is intended to include "*opinions, claims, commentary and invective*".[21]

The High Court has emphasised the importance of balancing public interest with free political discourse, and that laws burdening political communication must be reasonably appropriate and adapted to serve a legitimate end.[22]

If a broad interpretation of 'contribute to' is taken, there is a risk of capturing speech that, while controversial or potentially misleading, still plays a role in robust political discourse and freedom of thought, which would disproportionately limit speech, as even indirect contributions to serious harm could trigger regulation under the Bill.

[17] EM 21.

[18] EM 43.

[19] EM 18.

[20] Bill cl 13.

[21] EM 44.

[22] *Lange v Australian Broadcasting Corporation* (1997) 189 CLR 520, 567.

There is also a strong argument that the intention to capture opinion will inevitably capture speech that plays a role in political discourse and freedom of thought, and thus disproportionately limit speech.

The current formulation would give rise to the risk of (a) a "chilling effect" on the right to freedom of expression, and (b) a challenge to the Bill in the High Court on the basis that the burden on the constitutionally-protected implied freedom of political communication is disproportionate.

The concept of 'misleading or deceptive' statements is also a legal threshold that can be challenging for commercial operators to apply without a high degree of subjectivity or excessive caution or to verify that a statement is misleading.

Suggested amendments:

(1) the words "contribute to" are deleted in the definition (and all other relevant places in the Bill) to align more with Principle 1 in relation to the right to freedom of expression.

(2) the words 'misleading or deceptive' could also be removed from (A) to limit the concept to statements that are actually false.

(3) the EM be amended to remove reference to 'opinions' in relation to what 'information' is intended to include.

(vii) **Disinformation**

Disinformation is defined in the Bill as, relevantly, misinformation where "(A) there are grounds to suspect that the person disseminating, or causing the dissemination of, the content intends that the content deceive another person; or (B) the dissemination of content on a digital service involves *inauthentic behaviour*".[23]

Considering (A) first, as with the definition of *misinformation*, the term 'grounds to suspect' appears insufficiently rigorous given the potential burden imposed on freedom of expression. The low bar would give rise to the same risks identified above in relation to the definition of *misinformation*.

In relation to (B), the dissemination of content on a digital service involves *inauthentic behaviour* if:

(1) the dissemination uses an automated system in a way that is reasonably likely to mislead an end-user about one of the following:

(I) the identity, purpose or origin of the person disseminating the content;

(II) the popularity of the content on the digital service;

(III) the motive or intention of an end-user; or

(IV) the source or origin of the content; or

(2) there are grounds to suspect the dissemination is part of coordinated action that is reasonably likely to mislead a user about one of (I)-(IV) above; or

(3) there are grounds to suspect that the dissemination uses an arrangement for the purpose of avoiding action by the provider of the digital service to either (a) comply with the Bill or another law; or (b) enforce compliance with the terms of use for the digital service; or

(4) the content is disseminated in the circumstances specified in the Digital Platform Rules.

Whilst the definition of *inauthentic behaviour* is complex, the intention is to capture 'troll farms' and social media bots and the like. It should also be observed that included in this

[23] Bill cl 13(2).

definition are terms of art requiring an element of judgment such as 'reasonably likely' and 'grounds to suspect' which, whilst well defined in other areas of the law, will likely require new consideration in this context.

Operatively, the only distinction between misinformation and disinformation in the Bill is the implied obligation on platform providers to both remove content that is disinformation that involves inauthentic behaviour; and to prevent an end-user from using the platform where the end-user is engaged in disinformation that involves inauthentic behaviour.[24]

Suggested amendments: a higher bar than 'grounds to suspect' is inserted where that standard appears, such as '*reasonable* grounds to *believe*', as this would more closely align with the proportionality principles in Australian case law, where the Courts have demanded a closer fit between the legislative aim and the means of achieving it.[25]

(viii) **Digital Platform Rules**

While Digital Platform Rules in the Bill typically relate to mechanical matters such as record-keeping and complaints handling procedures, limb (4) of *disinformation* at paragraph 4.1(b)(vii) above enables Digital Platform Rules to specify circumstances in which dissemination of content constitutes *inauthentic behaviour.* Given misinformation codes and misinformation standards must not contravene the constitutionally-implied freedom of political communication,[26] it seems logical that Digital Platform Rules do the same.

Suggested amendment: add a new clause 83 substantially the same as clause 54 limiting Digital Platform Rules from contravening the implied freedom of political communication.

Conclusion: mild degree of alignment & degree of divergence

(c) The right to privacy

Divergence

The Principles refer to Article 17 of the ICCPR, which guarantees the right to privacy, protecting individuals from arbitrary or unlawful interference with their privacy, family, home, or correspondence, and from unlawful attacks on their reputation.[27] However, the EM flags the right to privacy as a human right potentially affected by the Bill. It notes that three aspects of the Bill particularly engage Australia's obligations on the right to privacy:

(i) the conferral of power on the ACMA to make Digital Platform Rules requiring digital communications platform providers to make and retain records relating to misinformation and disinformation on their platforms (subclause 30(1)). In exercising this power, the ACMA could require digital communications platform providers to make and retain records relating to misinformation or disinformation posted by individual end-users, or relating to complaints made by end-users about misinformation or disinformation. It is possible that in making such a rule, the ACMA could effectively require digital communications platform providers to make and retain records that include personal information.

(ii) the conferral of power on the ACMA to require digital communications platform providers and other persons to provide it with information or documents relating to misinformation

[24] EM cl 67; while the obligation is expressed in the negative, i.e. "*[n]othing in this Part…requires a digital communications platform provider to remove…content…; or prevent an end-user from using a …platform…*", the EM confirms it has the capacity to operate as a positive obligation: see, e.g. pp 100-101: "*[r]esponding to misinformation or disinformation could include requiring platforms to remove content where it is disinformation that involves inauthentic behaviour*".

[25] *Australian Capital Television Pty Ltd & New South Wales v Commonwealth* [1992] HCA 45.

[26] Above n 17.

[27] Principles 10.

or disinformation on a digital communications platform (clauses 33 and 34). Again, it is possible that this information could include personal information.

(iii) more broadly, the ACMA's power to request, approve, or approve variations to, misinformation codes, and to determine or vary misinformation standards. It is possible that these codes or standards may include provisions relating to the use, storage or disclosure of the personal information of individual end-users by digital communications platform providers.[28]

Alignment

In seeking to protect against what the ICCPR considers 'unlawful' interferences with the right to privacy, the EM refers to the test formulated by the Human Rights Committee, which is similar to the test relating to the implied freedom of political communication, namely, that restrictions on the right must not be arbitrary, and, for restrictions not to be arbitrary, they must be imposed in pursuit of a legitimate objective, and must be 'necessary and proportionate' to the achievement of that objective.[29]

The Bill also contains a few protections:

(i) the most substantive protection is in subclause 30(2), regarding the ACMA's power to make rules requiring digital communications platform providers to make and retain records relating to misinformation and disinformation, which provides that before making a digital platform rule for this purpose, the ACMA must consider the privacy of end-users and whether the rule is required for the performance of the ACMA's functions. Subclause 30(3) provides moreover that Digital Platform Rules relating to records must not require digital communications platform providers to make or retain records of the content of private messages or VoIP communications. Subclause 33(3) also provides that in exercising its powers to obtain information from digital communications platform providers, the ACMA must not require a person to provide information or documents that would reveal the content of a private message or VoIP communication.[30]

(ii) subclause 34(2) provides that in exercising its powers to obtain information from other persons, the ACMA must not require a person to provide information or documents relating to content posted by that person on a digital communications platform, other than in the person's capacity as a fact-checker, content moderator, an employee of the provider of the platform, or person providing services to the provider of the platform. Subclause 34(4) provides moreover that the ACMA cannot require a person to reveal the content of a private message or VoIP communication.

(iii) clauses 45 and 46 provide that the ACMA must not approve a code (or part of a code) or determine a standard that contains requirements related to, respectively, the content or encryption of private messages or VoIP communications. These protections for personal information reflect the fact that the ACMA's regulatory powers are not designed to address the behaviour of individual end-users, but rather, the way in which digital communications platform providers manage the risk of misinformation and disinformation on their platforms. The protections described above ensure that to the extent that the ACMA's new regulatory powers restrict the right to privacy, such restrictions are reasonable, necessary and proportionate to the achievement of a legitimate objective.[31]

---

[28] EM 16.

[29] EM 15, referring to Human Rights Committee, *Views: Communication No 633/1995*, UN Doc CCPR/C/65/D/633/1995, Annex ('*Gauthier v Canada*') para 13.6 (interpreting the term 'arbitrary' in relation to the right to freedom of expression). The UN High Commissioner for Human Rights has affirmed the relevance of this test to an assessment of restrictions on the right to privacy: UN High Commissioner for Human Rights, *The Right to Privacy in the Digital Age*, UN Doc A/HRC/51/17 (4 October 2022) para 48.

[30] Except for a private message or VoIP communication relating to the internal operations of the platform sent by an employee of, or person providing services to, the provider: Bill cl 33(3).

[31] EM 17.

(iv) clause 12 exempts certain digital communications platforms from the substantive requirements of Part 2, from Digital Platform Rules made for the purposes of that Part, and from approved misinformation codes and misinformation standards. Those provisions and instruments would not apply in relation to a digital communications platform to the extent that it is one of the following: an email service, a media sharing service that does not have an interactive feature, or a digital service the Minister determines is an excluded service for misinformation purposes under subclause 12(3).[32]

Our view is that these protections are modest compared with the degree of divergence with the right to privacy outlined above.

Conclusion: moderate degree of divergence

(d) The right to participate in public affairs and the right to vote

The Bill directly addresses this right. To guard against the risk to free and fair elections identified in the Principles, the EM notes that,

> *"…pursuant to the regulatory scheme set out in Schedule 9 to the BSA, the dissemination of content on a digital service is considered misinformation or disinformation if, in addition to meeting other criteria specified in clauses 13 and 14, it is reasonably likely to cause or contribute to serious 'harm to the operation or integrity of a Commonwealth, State, Territory or local government electoral or referendum process'. The inclusion of this type of harm in clause 14 means that digital communications platform providers have a responsibility to be transparent about the way in which they handle that type of content, and that the ACMA has regulatory powers in relation not such content.*"[33]

The EM further notes that, so far, the spread of misinformation and disinformation has not significantly damaged the Australian electoral process, and thus has not undermined the right of Australians to vote, to be elected and to participate in the conduct of public affairs. Schedule 9 to the BSA, however, seeks to address the risk that is currently posed by misinformation and disinformation – that it will radically disrupt an electoral process in Australia - and in doing so, seeks to safeguard the right of all Australians to participate in the conduct of public affairs and to vote and be elected at genuine periodic elections.

Conclusion: moderate degree of alignment

(e) Compatibility with human rights generally

It is a point of alignment that an assessment of whether Digital Platform Rules[34] and misinformation standards[35] are compatible with human rights must be prepared and included in the explanatory statement for the rules (see section 9 of the *Human Rights (Parliamentary Scrutiny) Act 2011* (Cth) and section 15J of the *Legislation Act 2003* (Cth)).

It is a point of divergence that an assessment of compatibility with human rights required under Digital Platform Rules is not also required for misinformation codes (noting that such codes are required not to contravene the constitutionally-implied freedom of political communication).[36]

Suggested amendments: to address the divergence under this paragraph 4.1(e),

(i) we agree with the HRLC's submission that a new clause 47(1)(h) should be introduced, requiring that the ACMA be satisfied that the Australian Human Rights Commission has been consulted in the development of any code under the Bill;[37]

[32] EM 43.

[33] EM 8.

[34] Bill cl 54.

[35] Bill cl 82(1).

[36] Bill cl 47(1)(d)(iii)-(iv).

[37] David Mejia-Canales, Human Rights Law Centre, *Submission on the Communications Legislation Amendment (Combatting Misinformation and Disinformation) Bill 2024* (Cth), 30 September 2024 ("***HRLC Submission***").

(ii) alternatively, or in addition to the addition of new clause 47(1)(h), we suggest that the requirement of an assessment of the compatibility of Digital Platform Rules with human rights being prepared and included in the explanatory statement for the Rules, be extended to misinformation codes.

Conclusion: Principle 1 mildly aligns and mildly diverges from the Bill given the presence of both protections and restraints on human rights and principles

4.2 Principle 2: Digital platforms should have a legal duty to make sure their products, systems, and services do not cause harm

Assessment: mild degree of alignment

We note the reference in the HRLC's submission to a "legislated duty of care".[38] There is currently no explicit "duty of care" contained in the Bill, as is the case in the United Kingdom's *Online Safety Act 2023*, Part 3. There is also no implicit duty of care on platform providers through any legislative intention that the Bill confers on a user a private civil cause of action for breach of that duty[39] (which, in this case, would likely be a duty not to cause serious harm through allowing the dissemination of misinformation or disinformation on platforms). Similarly, we are not aware of any Australian court having found a common law duty of care to exist between users and platforms.

Short of a legal duty of care, the key obligations placed on digital platforms in the Bill (assessing risks, publishing policies, and publishing a media literacy plan on misinformation and disinformation), together with the focus on systems rather than content and ACMA's enforcement powers, operate as mechanisms by which digital platforms must ensure their products, systems, and services do not cause serious harm. We consider Principle 2 in terms of the duty of care requirements outlined in the Principles as follows:[40]

(a) Uphold and protect the fundamental human rights of their users

The Bill mildly aligns with this obligation - refer to discussion at paragraph 4.1.

(b) Undertake comprehensive risk assessments to identify and analyse risks stemming from their products and systems

The Bill aligns considerably with this obligation. Subparagraph 17(1)(a) of the Bill places an obligation on a digital communications platform provider to make publicly accessible a report on the outcomes of an assessment by the provider of risks relating to misinformation and disinformation on their platform.[41] The report is to include the outcomes of an assessment of risks arising from both the design or functioning of the platform and risks arising from the use of the platform by end-users.[42]

Subparagraph 19(1)(a) provides that the Digital Platform Rules may require platforms to update their assessments of risks relating to misinformation and disinformation, at times (e.g. at certain frequencies), or in circumstances specified in the rules (e.g. in response to an event that has triggered a marked increase of misinformation and disinformation on platforms such as a pandemic, or outbreak of war).[43]

These risk mitigation provisions accord with the HRLC's call for platforms to identify risks stemming from their products and systems, marking a point of alignment.

(c) Address identified risks though effective risk mitigation measures

[38] HRLC Submission 16.

[39] Roplh, D, et. al., *Balkin & Davis Law of Torts*, LexisNexis Australia (2021) 631, citing *Jane Doe v Fairfax Media Publications Pty Ltd* [2018] NSWSC 1996 at [111].

[40] Ibid.

[41] Bill cl 17(1)(a).

[42] EM 66.

[43] EM 75.

To comply with subparagraph 17(1)(a), it is implicit that a digital communications platform provider will need to engage in a process to identify and assess risks relating to misinformation and disinformation on its platform.[44]

Subclauses 19(c), 19(d) and 19(e) of the Bill provide the ACMA with the discretion to make Digital Platform Rules to: (1) require digital communications platform providers to have risk management plans for risks relating to misinformation and disinformation; (2) require those risk management plans to be prepared at times, or in circumstances, specified in the rules; and (3) require those risk management plans to state the steps (if any) being taken by digital communications platform providers in relation to risks (in respect of misinformation or disinformation) identified by providers or specified in the rules (paragraph (e)).[45] These obligations are enforceable by the ACMA through civil penalties.[46]

The risk assessment measures in the Bill align with the HRLC's call for platforms to not only identify risks but also take active steps to address them. While the Bill establishes a solid foundation for risk mitigation, its success would depend on how robustly platforms implement and disclose these measures. Principle 2 implies that platforms should take stronger steps to demonstrate that risk mitigation efforts are comprehensive and effective.

(d) Open up their assessment and mitigation measures for scrutiny by third parties to enable independent testing and verification

Subparagraph 17(1)(a) does not require digital communication platform providers to make accessible their complete risk assessments (although they may choose to do so). In the case of large providers in particular, a risk assessment may be a lengthy document, identifying threats and including confidential information or damaging information about system vulnerabilities that providers would be unlikely to want to make public. Hence digital communications platform providers are only required to publish a report on the outcomes of their risk assessment. At the discretion of the provider, the report could be an in-depth analysis of the key misinformation or disinformation related risks that the provider has identified; or it could be a high-level summary of those risks.[47] The provision for a triennial review in clause 70 is another avenue through which further third-party assessment could be made. The Minister would be required to set up a review as soon as possible after the third anniversary of the Bill commencing and afterwards at intervals of no longer than 3 years (subclause 70(1)).

The review must provide for public consultation (subclause 70(3)) and would need to include an assessment of the impact of Part 2 (which contains most operative provisions) on freedom of expression (subparagraph 70(2)(a)) and consider whether Part 2 should be amended (subparagraph 70(2)(b)).

The first time the review is conducted, an amendment to include a scheme for third party data access in Part 2 must be considered (clause 70(2)(c)), which would involve platform providers being required to give accredited independent researchers access to data regarding misinformation or disinformation on platforms.[48]

While these provisions stop short of wholly meeting the call for transparency in Principle 2(d), they must be considered in the broader context of the Bill's overall transparency framework. Specifically, in addition to the obligations in clauses 17 and 19 discussed above, the following provisions bolster transparency in the Bill by ensuring key information is either accessible to the public or provided to the ACMA:

[44] EM 66.

[45] EM 76.

[46] Bill cl 20.

[47] EM 66.

[48] EM 125.

- clause 25 allows Digital Platform Rules to require publication, or provision to the ACMA, of information regarding misinformation complaints and dispute handling[49]
- clause 30 allows Digital Platform Rules to require digital communications platform providers to give reports to the ACMA on misinformation and disinformation on their platforms, and on the measures they have implemented to prevent and respond to misinformation or disinformation[50]
- clause 33 enables the ACMA to require digital communications platform providers to provide information and documents relevant to misinformation or disinformation on their platforms[51]
- clause 38 enables the ACMA to publish on its website certain information about misinformation and disinformation including statements of reasons for why a provider has not published certain information obtained under subparagraph 17(4)(b) and information that ACMA has gathered from digital communications platform providers under clauses 30 and 33.[52]

On balance, the Bill diverges from Principle 2(d) in terms of transparency. Without a stronger ability for independent oversight of risk assessments, third-party scrutiny may be limited, diluting alignment between the Bill and Principle 2.

Conclusion: mild degree of alignment

4.3 Principle 3: Content removal powers have a role to play in limited circumstances but only as part of a broad, comprehensive regulatory framework

Assessment: mild degree of alignment

Content moderation is principally the responsibility of digital platform providers under the Bill. Although it is one of a few key obligations placed on platforms (among the requirements to assess risks, publish policies, and publish a media literacy plan on misinformation and disinformation), it appears that content moderation and removal would be the primary mechanisms by which platform providers would be required to comply with the Bill on a day-to-day basis.

Clause 44 states that industry-written misinformation codes (which are approved by the ACMA and become legislative instruments)[53] and misinformation standards (which are determined by the ACMA and also become legislative instruments)[54] may deal with "*preventing or responding to misinformation or disinformation on digital communications platforms*", including through "*using technology*".[55] If platforms are found by the ACMA not to be compliant with their misinformation codes or misinformation standards, the ACMA may exercise its enforcement powers (being warnings, remedial directions, infringement notices and civil penalty notices),[56] suggesting a degree of divergence with Principle 3.

To temper the reliance on content moderation, the Bill contains two key provisions that attempt to limit platform providers' obligations to remove content:

(a) First, clause 67 provides that nothing in Part 2 of the Bill, or in a rule, approved code, or standard made, approved or determined pursuant thereto, can require a platform provider to remove content from a platform or prevent an end-user from using the platform, except in the

[49] Bill cl 25; EM 39.

[50] Bill cl 30; EM 40.

[51] Bill cl 33; EM 40.

[52] Bill cl 38; EM 40.

[53] Bill cl 47.

[54] Bill cl 55.

[55] Bill cl 44(3)(a)-(b).

[56] Bill Sch 1, Pt 3, Div 1.

case of disinformation that involves inauthentic behaviour (as to which see the discussion at paragraph 4.1(b)(vii)).

In our view, this limitation is only mildly effective in relation to Principle 3. While subclause 67(2) states that the removal requirements do not apply to misinformation, the EM explains in relation to clause 44 that:

> "*The term 'preventing' is intended to have its ordinary meaning. This could include specific measures to minimise the spread of misinformation or disinformation, such as altering recommendation algorithms to prevent misinformation or disinformation content being propagated, or closing or limiting the reach of accounts that repeatedly disseminate misinformation or disinformation…*
>
> *The term 'responding' is also intended to have its ordinary meaning. For example, a code or standard could require platforms to put in place measures to 'nudge' users to consider whether to on-share the information.*"[57]

As for the use of technology, the EM states that:

> "*Digital communications platform providers could be required to use automated processes and technology to detect and act appropriately on misinformation and disinformation under a misinformation code or misinformation standard. For example, they could be required to use technology or algorithms to 'downrank' or reduce the spread of misinformation or disinformation, or to 'nudge' a person to reflect on the information before sharing it.*"[58]

We suggest 'nudging' is a proportionate content moderation measure that would align with both Principle 3 and the right to freedom of expression (discussed at paragraph 4.1(b)).

In relation to particular content, the balance of the suggested misinformation measures, being altering recommendation algorithms to prevent its propagation, closing or limiting the reach of accounts that repeatedly disseminate it, and using technology, algorithms and automated processes to 'downrank' or reduce the spread of it, are all strict methods of content moderation. While short of the express removal obligation in relation to disinformation that involves inauthentic behaviour, *preventing* and *responding* to misinformation (and disinformation that doesn't involve inauthentic behaviour) would arguably operate in a manner closely akin to content removal.

Although they may be successful at limiting the spread of *misinformation* and *disinformation,* there is a risk the content removal and moderation powers in the Bill, together with the strong penalties (a maximum of the greater of 25,000 penalty units or 5% of annual turnover during the turnover period),[59] would act as the primary force in the Bill's regime in a manner divergent from Principle 3.

As the definitions of *misinformation* and *disinformation* are currently formulated, this regime would also arguably give rise to the risk raised in the Principles - that it would "*inadvertently suppress the freedom of speech or restrict access to legitimate information*"[60] (see further discussion at paragraph 4.1(b)).

(b) Second, in approving, or approving variations of, misinformation codes or determining or varying misinformation standards, the ACMA must be satisfied they do not contravene the implied freedom of political communication.[61] This would likely limit content moderation powers in alignment with Principle 3 (however we would refer to the discussion at paragraph 4.1(b)).

Suggested amendments & conclusion: the Bill would align more closely with the Principles if it instituted a regime that prevented the amplification of misinformation and disinformation, rather than a

[57] EM 100-101.

[58] EM 101.

[59] Bill Sch 2, Pt 1, item 21(5H).

[60] Principles 16.

[61] Above n 17.

regime that proposes to use practices akin to shadow banning and preventing any propagation of the authentic posts of ordinary individuals and organisations.

The dissemination of *disinformation* that involves inauthentic behaviour may be one of the limited circumstances in which the Bill should encourage the use of content removal powers (subject to the discussion at paragraph 4.1(b)(vii)).

Finally, in general, we agree with the HRLC's submission that the Bill should clarify the concepts of *misinformation* and *disinformation*. At present, the interpretation of the distinction between *misinformation* and *disinformation* is done by implication in the Bill and with respect to one operative provision only (clause 67), which we would agree is unclear and may create confusion.[62]

Conclusion: mild degree of alignment

4.4 Principle 4: Users should have control over how platforms collect and use their data

Assessment: low degree of alignment

The Bill only indirectly aligns with this principle, in two respects:

(a) first, clause 70 provides for a triennial review of the operation of Part 2, which is discussed at paragraph 4.2(d).

(b) second, the Bill contains some educational provisions, such as the requirement for platform providers to publish a media literacy plan,[63] and the provision that misinformation codes and standards may include sections seeking to improve the media literacy of the end-user.[64]

While these measures may assist users understand how their information is collected and used, they do not give users any degree of control over how platforms collect and use their data. Given this, our view is there is a low degree of alignment between the Bill and this Principle.

Conclusion: low degree of alignment

4.5 Principle 5: Judicial oversight is essential in a comprehensive regulatory framework

Assessment: moderate degree of alignment

The Bill provides that applications may be made to the Administrative Review Tribunal ("**ART**") to review a decision by the ACMA (a) to approve or vary a misinformation code;[65] and (b) under any Digital Platform Rules, so long as those rules provide that the decision is a reviewable decision for the purposes of the relevant section.[66] The structure of the system of review with the ART as the first body of review strikes the same balance the Principles refer to between judicial oversight and avoiding "*well-resourced litigants, like large social media companies, from weaponizing the courts through endless litigation*".[67]

It is unclear why ART review is not present in the Bill for misinformation standards. Given misinformation standards are determined by ACMA, and that they prevail over inconsistent misinformation codes[68] and are subordinate to inconsistent Digital Platform Rules,[69] it would be logical that they be subject to the same review process as misinformation codes and Digital Platform Rules. The inconsistent review framework marks a point of divergence from Principle 5.

Should a party wish to appeal a question of law in relation to a decision by the ART, an application may be made to the Federal Court. The Federal Court is also able to determine the amount of civil

[62] Principles 11.

[63] Bill cl 17(1)(c).

[64] Bill cl 44(3)(l).

[65] Bill Sch 2, Pt 1, item 14.

[66] Bill Sch 2, Pt 1, items 15 and 35.

[67] Principles 20.

[68] Bill cl 65.

[69] Bill cl 66.

penalties for breaches of approved misinformation codes and misinformation standards (up to amounts specified in the Bill).

Thus, the Bill appears to balance regulatory powers, quasi-judicial review, and judicial oversight, and therefore contains a moderate degree of alignment with Principle 5.

Conclusion: moderate degree of alignment