Human Rights Law Centre

Submission on the *Communications Legislation Amendment (Combatting Misinformation and Disinformation) Bill 2023*

Department of Infrastructure, Transport, Regional Development, Communications and the Arts

August 2023 |

Human Rights Law Centre

David Mejia-Canales
Senior Lawyer
Human Rights Law Centre Ltd
Level 17, 461 Bourke Street

T: [redacted]
F:
E:
W: www.hrlc.org.au

Alice Drury
Acting Legal Director
Human Rights Law Centre Ltd
Level 17, 461 Bourke Street
Melbourne VIC 3000

T: [redacted]
F:
E:
W: www.hrlc.org.au

Human Rights Law Centre

The Human Rights Law Centre uses strategic legal action, policy solutions and advocacy to support people and communities to eliminate inequality and injustice and build a fairer, more compassionate Australia. We work in coalition with key partners, including community organisations, law firms and barristers, academics and experts, and international and domestic human rights organisations.

The Human Rights Law Centre acknowledges the people of the Kulin and Eora Nations, the traditional owners of the unceded land on which our offices sit, and the ongoing work of Aboriginal and Torres Strait Islander peoples, communities and organisations to unravel the injustices imposed on First Nations people since colonisation. We support the self-determination of Aboriginal and Torres Strait Islander peoples.

Follow us at http://twitter.com/humanrightsHRLC

Join us at www.facebook.com/HumanRightsLawCentreHRLC/

Contents

Introduction **4**

1. The problem **4**

2. Recommendations **5**

3. Human rights are being undermined by misinformation and disinformation **5**

3.1 Human rights in the digital realm 6

3.2 Relevant human rights law 7

4. Regulator-drafted industry standards should be the norm **10**

5. What an ACMA standard should include **11**

5.1 Platforms should be subject to a comprehensive transparency, risk assessment, mitigation and audit framework 11

5.2 Users should have control over what they see and how their data is used 13

6. Additional, minimum necessary amendments to the draft Bill **14**

6.1 Strengthening the data access regime by extending it to civil society and expert researchers 14

6.2 Including the human rights to be protected in misinformation standards 15

Introduction

The Human Rights Law Centre thanks the Department of Infrastructure, Transport, Regional Development, Communications and the Arts for the opportunity to provide feedback on the *Communications Legislation Amendment (Combatting Misinformation and Disinformation) Bill 2023* (Cth) (**draft Bill**).

The Human Rights Law Centre supports a form of regulation not premised on industry self-regulation or co-regulation, given its demonstrable failure to adequately address disinformation and hate speech to date. We instead support the introduction of a regulatory regime modelled on the European Union's *Digital Services Act* (**DSA**).

Nonetheless, with some amendments, the Human Rights Law Centre welcomes the draft Bill as a positive step toward regulating digital platforms, which have profited from the spread of dangerous disinformation, misinformation and hate speech during Australian elections, a pandemic, and bush fires exacerbated by the climate crisis.

Suggestions by some that digital platforms should not be regulated at all because of concerns regarding the right to freedom of speech, fundamentally misunderstands the nature of the right, and ignores all other rights, including the right to freedom of conscience, the right to vote and, particularly in times of crisis, the right to health, the right to life and the right to a healthy environment.

Too often, a misguided interpretation of the right to free speech is weaponised by the powerful to avoid accountability for the harm they cause by their speech. The Australian Government must not be dissuaded from pursuing this reform proposal by such people.

1. The problem

Digital platforms are driving the spread of disinformation and hate speech on a level never before seen in history. Large digital platforms have an enormous level of influence over public discourse, with the power to amplify the information – and disinformation – that forms the basis for people's decisions and beliefs.

Around the world, calls to regulate the tech sector grow louder every day. Australia has been an early mover on innovative policy reform around online safety and digital media[1], yet it currently lags behind on key aspects of regulating digital platforms.

The Human Rights Law Centre has previously made a submission to the Senate's Standing Committee on Economics' Inquiry into the Influence of Large Online Platforms. That submission addressed similar matters that the Department is currently seeking feedback on. We include that submission as part of our feedback into this process.

In this submission we will expand further on the need to regulate digital platforms with a reference to human rights law and principles.

[1] See eg Australian Government, eSafety Commissioner, 'What is the Online Safety Act 2021?' January 2022 https://www.esafety.gov.au/newsroom/whats-on/online-safety-act; Australian Competition and Consumer Commission, 'News media bargaining code' (accessed 28 February 2023) https://www.accc.gov.au/focus-areas/digitalplatforms/news-media-bargaining-code/news-media-bargaining-code

2. Recommendations

1. The federal Government should move away from self-regulatory and co-regulatory models for digital platforms, by amending the draft Bill to empower ACMA to immediately create and enforce standards that are grounded in international human rights law, similar to the EU's *Digital Services Act*. The draft Bill should also be amended to include:

   (a) requirements for major platforms to undertake and publish risk assessments that identify human rights risks and other forms of harm, with corresponding obligations to develop and implement mitigation measures; and

   (b) measures to give users greater control over the collection and use of their personal data, including by making recommender systems opt-in and by limiting forms of profiling.

2. The federal Government should at minimum strengthen the draft Bill by:

   (a) including a data access regime to support civil society and expert research into the amplification of disinformation and hate speech;

   (b) explicitly including the human rights to be protected in misinformation standards, be they developed by industry or ACMA.

3. Human rights are being undermined by misinformation and disinformation

The internet and the rapid spread of information and communication has brought great potential to accelerate human progress. These technologies can, if their use is grounded in human rights law and principles, contribute to the protection and promotion of human rights. However, the rapid spread of disinformation and misinformation online poses significant challenges to the fulfilment and the enjoyment of those rights.

The amplification of disinformation and hate speech online turbo-charges discrimination, polarises society and distorts public debate on matters of critical importance. From disinformation campaigns undermining the right to health during a pandemic, to misleading material that can distort free and fair elections, to hate speech that stokes violence and threatens lives, the proliferation of disinformation, misinformation and harmful material online has a profound impact on human rights and democratic processes in Australia and beyond.

The United Nations Human Rights Council[2], the United Nations Secretary-General's High-Level Panel on Digital Cooperation[3], the United Nations' Special Rapporteurs on the Freedom of Expression, the Organization for Security and Co-operation in Europe, the Organisation of American States and the African

[2] United Nations Human Rights Council Resolution (2013) *The safety of journalists* UN Doc A/HRC/RES/39/6 (27 September 2018) 7

[3] UN Secretary-General's High-level Panel on Digital Cooperation (2019), The Age of Digital Interdependence

Commission on Human and People's Rights[4] have all expressed concern about the rapid spread of disinformation and misinformation and for the need to apply a human rights framework to limit its harm.[5]

3.1 Human rights in the digital realm

Australia, as a party to the major international human rights treaties[6], is required by law to guarantee and protect the rights enumerated in those treaties to all people in Australia. These obligations do not diminish or disappear online. The United Nations' Human Rights Council has issued resolutions requiring that 'the same rights that people have offline must also be protected online'.[7]

As businesses digital platforms do not have a legal obligation to guarantee and protect human rights in the way that nation states do. However, international human rights law is not silent on the role of businesses in the protection and promotion of human rights.

The United Nations' *Guiding Principles on Business and Human Rights* provides a framework to ensure that the activities of all businesses, regardless of their size, corporate structure, ownership or location; are human rights compliant.[8] The Principles require that all business enterprises respect human rights while also remedying any adverse human rights impacts that they may have created or contributed to.[9]

While digital platforms do not have the same legal duty as states to be bound by human rights law they do have an obligation to respect it. Furthermore, by virtue of them operating in Australia they recognise that all Australian governments have a duty to protect all people within our territory from human rights abuses, including on their platforms.[10]

---

[4] UN Office of the High Commissioner for Human Rights (2017), 'Freedom of Expression Monitors Issue Joint Declaration on 'Fake News', Disinformation and Propaganda', https://www.ohchr.org/EN/NewsEvents/Pages/DisplayNews.aspx?NewsID=21287&LangID=E (accessed 14 Oct. 2019).

[5] Kate Jones, Online Disinformation and Political Discourse- Applying a Human Rights Framework, Research Paper, November 2019 https://www.chathamhouse.org/sites/default/files/2019-11-05-Online-Disinformation-Human-Rights.pdf 27

[6] International Covenant on Civil and Political Rights; International Covenant on Economic, Social and Cultural Rights; International Convention on the Elimination of All Forms of Racial Discrimination; Convention on the Elimination of Discrimination against Women; Convention against Torture and Other Cruel, Inhuman or Degrading Treatment or Punishment; Convention on the Rights of the Child, and Convention on the Rights of Persons with Disabilities

[7] UN Human Rights Council Resolutions (2012-2018), The promotion, protection and enjoyment of human rights on the Internet, UN Doc A/HRC/RES/38/7 (5 July 2018), A/HRC/RES/32/13 (1 July 2016), A/HRC/RES/26/13 (26 June 2014), A/HRC/RES/20/8 (5 July 2012) as referenced in Kate Jones, Online Disinformation and Political Discourse- Applying a Human Rights Framework, Research Paper, November 2019 https://www.chathamhouse.org/sites/default/files/2019-11-05-Online-Disinformation-Human-Rights.pdf 30

[8] United Nations Office of Human Rights, Guiding Principles on Business and Human Rights, HR/PUB/11/04, 2011, https://www.ohchr.org/sites/default/files/documents/publications/guidingprinciplesbusinesshr_en.pdf

[9] United Nations Office of Human Rights, Guiding Principles on Business and Human Rights, HR/PUB/11/04, 2011, https://www.ohchr.org/sites/default/files/documents/publications/guidingprinciplesbusinesshr_en.pdf at 13

[10] Kate Jones, Online Disinformation and Political Discourse- Applying a Human Rights Framework, Research Paper, November 2019 https://www.chathamhouse.org/sites/default/files/2019-11-05-Online-Disinformation-Human-Rights.pdf at 30

3.2 Relevant human rights law

Freedom of expression

The right to freedom of expression is a fundamental right that combined with other rights, like the right to vote (discussed further below), allows for the fulfilment of other rights.

The right to freedom of expression, as well as the permissible restrictions of the right, are contained in Article 19 of the International Covenant on Civil and Political Rights (**ICCPR**). Article 19 should also be read in conjunction with Article 20 of the ICCPR which outlaws all propaganda for war as well as the advocacy of national, racial or religious hatred that incites discrimination or violence.

The United Nations' Human Rights Committee has commented that the right to freedom of expression also guarantees a free, uncensored and unhindered press or other media as well as the 'free communication of information and ideas about public and political issues between citizens, candidates and elected representatives[11].

The right to freedom of expression is not absolute and can be restricted, however these restrictions are narrowly construed. The United Nations Human Rights Committee has observed that a restriction on the right is only permitted when the restriction is: a) provided by a law of sufficient precision, b) for one of the purposes outlined in Article 19(3)[12] , and c) conforms to strict tests of necessity and proportionality.[13] ' Furthermore, subject to a few exceptions (those contained in Article 20, for example) the right to disseminate information is *not* limited to information that is true.[14]

Responses to the problem of disinformation, whether on the part of governments, regulatory bodies or the platforms themselves, must not infringe upon the right to freedom of expression and the right to access information – two cornerstones of democratic discourse. Protecting the right to freedom of expression is not incompatible with preventing the harms caused by misinformation and disinformation online provided that any limitations on the right are human rights compliant.

The internet has allowed for a much greater and more diverse number of voices to take part in public debate. The internet has also enabled voters to be better informed during elections and referendums. However, there is also a growing appreciation both by digital platforms and governments to counter the potential harms of misinformation and disinformation online.

To protect and promote the right to freedom of expression online, digital platforms must be required to establish consistent, fair, effective and transparent guidelines grounded in human rights law. Digital platforms are not best placed to make and enforce these regulations without a directive to do so.

As commercial entities, digital platforms' main imperative is to maximise users and in turn their revenue and, as they are not public entities they are not required to properly consider human rights laws and principles over their decisions. If digital platforms are allowed to make and enforce their own regulations they risk being unfair, inconsistent, applying the wrong human rights standards, and not providing the necessary level of transparency over their decisions to track their compliance.[15]

---

[11] UN Human Rights Committee, General Comment No. 34 (2011), para. 15, 20, 22
[12] Article 19(3) provides that the right to freedom of expression can' be subject to certain restrictions, but these shall only be such as are provided by law and are necessary: (a) For respect of the rights or reputations of others; (b) For the protection of national security or of public order (ordre public), or of public health or morals.
[13] UN Human Rights Committee, General Comment No. 34 (2011), para. 22
[14] Kate Jones, Online Disinformation and Political Discourse- Applying a Human Rights Framework, Research Paper, November 2019 https://www.chathamhouse.org/sites/default/files/2019-11-05-Online-Disinformation-Human-Rights.pdf at 41
[15] Kate Jones, Online Disinformation and Political Discourse- Applying a Human Rights Framework, Research Paper, November 2019 https://www.chathamhouse.org/sites/default/files/2019-11-05-Online-Disinformation-Human-Rights.pdf at 47

Freedom of thought and conscience

The right to freedom of thought and conscience is contained in Article 18 of the ICCPR.

The United Nations Human Rights Committee has commented that the right to freedom of thought and conscience is the foundation for free and democratic societies, in part, because freedom of thought is a precondition for the full enjoyment of other human rights.[16] At a minimum, the right to freedom of thought gives people the right to not have their opinions 'unknowingly manipulated or involuntarily influenced'.[17]

Digital platforms are not benign, inert, or neutral repositories of data and content. Most are profit-driven services that derive their income from captivating their users' attention through the use of algorithms. To maximise user attention which in turn generates revenue, digital platforms use algorithms that favour content that provokes strong emotional responses, inspires emotion, or arouses anger or outrage.[18]

There is a real fear that these algorithms could be impacting the right to freedom of expression by suppressing or promoting different types of content. Currently, it is not possible to fully understand how digital platforms may be manipulating and influencing the opinions of their users as they do not have to be transparent over the algorithms that they use to serve, curate, and promote content on their service. This leaves digital platforms, and by extension their users, open to abuse by those who may wish to manipulate these algorithms to influence audiences for their own purposes.[19]

For example, a report found that Russian interference in the 2016 election for the presidency of the United States of America amounted to 'a sweeping and sustained social influence operation consisting of various coordinated disinformation tactics aimed at US citizens, designed to exert political influence and exacerbate social divisions in US culture.'[20]

Digital platforms, social media companies in particular, are in truth advertising companies. Their main motivation, despite their protestation, is to curate the content posted onto their platforms (as well as the personal data generated by their users) in a way that maximises their own profits instead of upholding human rights or democratic principles.[21]

We currently have few meaningful opportunities to understand how platforms and algorithms shape the information environment in which we form views and make decisions. It is largely thanks to industry whistle-blowers that we have achieved any scrutiny and accountability for the big tech companies.[22] Now

---

[16] UN Human Rights Committee, General Comment No. 34: Article 19 (Freedoms of Opinion and Expression), Human Rights Committee 102nd session, UN Doc CCPR/C/GC/34 (12 September 2011) para 2-9, in Kate Jones, Online Disinformation and Political Discourse- Applying a Human Rights Framework, Research Paper, November 2019 https://www.chathamhouse.org/sites/default/files/2019-11-05-Online-Disinformation-Human-Rights.pdf at 32.
[17] UN Secretary-General's High-level Panel on Digital Cooperation (2019), The Age of Digital Interdependence, p. 17.
[18] Williams, J. (2018), Stand out of our Light, Cambridge: Cambridge University Press, pp 33-35, 79-80
[19] Kate Jones, Online Disinformation and Political Discourse- Applying a Human Rights Framework, Research Paper, November 2019 https://www.chathamhouse.org/sites/default/files/2019-11-05-Online-Disinformation-Human-Rights.pdf at 36.
[20] Diresta, R. et al (2018), The Tactics and Tropes of the Internet Research Agency; Howard, P. et al (2018), The IRA, Social Media and Political Polarisation in the United States, 2012-2018 in Kate Jones, Online Disinformation and Political Discourse- Applying a Human Rights Framework, Research Paper, November 2019 https://www.chathamhouse.org/sites/default/files/2019-11-05-Online-Disinformation-Human-Rights.pdf at 36.
[21] Kate Jones, Online Disinformation and Political Discourse- Applying a Human Rights Framework, Research Paper, November 2019 https://www.chathamhouse.org/sites/default/files/2019-11-05-Online-Disinformation-Human-Rights.pdf at 36.
[22] Shirin Ghaffary, 'Big Tech's employees are one of the biggest checks on its power' Vox, 29 December 2021 https://www.vox.com/recode/22848750/whistleblower-facebook-google-apple-employees

and into the future, we should not need to rely on whistle-blowers in order to understand the ways we are tracked and targeted, and the systems that determine the information that is delivered to us.

Accordingly, regulation should be focused on increasing transparency and accountability. Transparency features of an appropriate regulatory model should place the onus on digital platforms to identify the risks posed by their platforms and the steps they will take in response, as well as providing information to users about their advertising and recommender systems.

These obligations should be reinforced by a well-resourced, independent regulator with the power to verify digital platforms' information, and hold them accountable for failures to report and act.

Right to participate in public affairs and the right to vote

The right to take part in public affairs and to vote is contained in Article 25 of the ICCPR. The right to take part in public affairs provides someone the right to participate in public life either through their elected representatives or for running for election themselves. The right to vote also guarantees citizens the right to participate in fair and free elections without interference.

The United Nations' Human Rights Committee has observed that all states are required to ensure that voters are able to form their opinions independently without fear or threats of violence, inducement, compulsion or manipulative interference of any kind.[23]

Large digital platforms have an enormous level of influence over public discourse, with the power to amplify the information – and disinformation – that forms the basis for people's opinions, decisions and beliefs, particularly through their use of algorithms that manipulate what content people see.

These algorithms could be inconsistent with the right to participate in public affairs and to vote if, for example, they give a person a distorted impression of public debate, prioritise disinformation, or give the impression that a particular outcome or candidate is bound to be successful to dissuade voters from campaigning for that outcome or candidate.[24]

As noted above, without adequate transparency over the algorithms employed by digital platforms it is not possible to properly determine if or how they are impacting the right to take part in public life and to vote.

Online misinformation and disinformation has sadly become a fixture of elections across the globe, with profound implications for democracy in the digital age. Powerful false narratives can be quickly amplified to millions with the potential to confuse the public, distort outcomes and undermine public confidence in electoral processes and results.[25]

---

[23] 7 UN Human Rights Committee, General Comment No. 25: Article 25 (Participation in Public Affairs and the Right to Vote), The Right to Participate in Public Affairs, Voting Rights and the Right of Equal Access to Public Service, 57th session, UN Doc CCPR/C/21/Rev.1/Add.7 (1996), para. 8; UN OHCHR, Promotion, protection and implementation of the right to participate in public affairs in the context of the existing human rights law: best practices, experiences, challenges and ways to overcome them (23 July 2015), A/HRC/30/26, paras 9–11 in Kate Jones, Online Disinformation and Political Discourse- Applying a Human Rights Framework, Research Paper, November 2019 https://www.chathamhouse.org/sites/default/files/2019-11-05-Online-Disinformation-Human-Rights.pdf at 48

[24] Kate Jones, Online Disinformation and Political Discourse- Applying a Human Rights Framework, Research Paper, November 2019 https://www.chathamhouse.org/sites/default/files/2019-11-05-Online-Disinformation-Human-Rights.pdf at 48

[25] Centre for Media Transition, University of Technology Sydney, Information Disorder Lessons from Australia, 9 December 2022, p. 32.

4. Regulator-drafted industry standards should be the norm

Instead of a "graduated" approach in which the Australian Communications and Media Authority (**ACMA**) only creates and enforces a standard as a last resort, the draft Bill should be amended to empower ACMA to create a standard which is immediately enforceable.

Australia's reliance on rules written by the tech industry itself has led to demonstrably weaker protections against harm and is clearly at odds with the expectations of the community.[26] Co-regulation is inappropriate for such a powerful and high-risk sector, in which business models frequently come into conflict with human rights principles, community needs, and the public interest.

Self-regulation and co-regulation will never be able to compete with the commercial interests of digital platforms. These commercial interests are driven by serving users a steady stream of emotive, divisive content and detailed personal profiling generated from their data.[27] As Ms Kate Jones, former associate fellow with Chatham House notes:

> Although market pressures encourage some action by companies that mitigates abuse- such as some removal of egregious content- they are not sufficient to respect the human rights of users. Often, voluntary commitments on the part of companies have not been implemented, and/or lack of transparency means that implementation has been impossible to measure.[28]

In the European Union the introduction of the DSA was driven by a growing recognition that self- and co-regulatory models are inadequate and ineffective. When the DSA comes into full force in February 2024, it will apply to all digital services operating in the EU, including social media, online marketplaces, search engines and other online platforms. Its supervised, risk-based approach to harms caused by content on digital platforms, complements and reinforces the European Commission's *Strengthened Code of Practice on Disinformation*.[29]

The DSA marks a major shift away from co-regulation in Europe, Australia should follow suit.

Effective regulation of digital platforms should include a comprehensive transparency framework and measures to protect privacy and give people greater control over what they see.

The Australian Government should not leave it to digital platforms to determine how they will guarantee that all Australians will have their human rights protected and affirmed online. The rules and norms regarding the protection of our human rights either online or offline should not be made by businesses or private companies that do not have the democratic accountability or the institutional legitimacy to make these decisions.

However, it's important that the Australian Government be mindful that any regulatory efforts that are impose on digital platforms are themselves fully human rights compliant. This includes protecting the right

---

[26] Dr Rys Farthing, 'How outdated approaches to regulation harm children and young people and why Australia urgently needs to pivot', Reset Australia, ChildFund Australia & Australian Child Rights Taskforce, December 2022, https://au.reset.tech/uploads/report_-co-regulation-fails-young-people-final-151222.pdf

[27] Kate Jones, Online Disinformation and Political Discourse- Applying a Human Rights Framework, Research Paper, November 2019 https://www.chathamhouse.org/sites/default/files/2019-11-05-Online-Disinformation-Human-Rights.pdf at 52

[28] Kate Jones, Online Disinformation and Political Discourse- Applying a Human Rights Framework, Research Paper, November 2019 https://www.chathamhouse.org/sites/default/files/2019-11-05-Online-Disinformation-Human-Rights.pdf at 52

[29] European Commission, The 2022 Code of Practice on Disinformation, https://digital-strategy.ec.europa.eu/en/policies/code-practice-disinformation

to freedom of expression by not incentivising the removal of content where the context surrounding that content is ambiguous.

Regulation that relies on content moderation can lead to limiting human rights in circumstances where it was never intended. This is because penalising platforms for content-moderation failures incentivises platforms to err on the side of caution, resulting in restrictions on freedom of expression and the right to access information in circumstances well beyond what was contemplated by the regulatory model. For these reasons, content moderation ought to be seen as a small but important part of any comprehensive and effective framework for digital regulation that is directed by government. Instead, the focus should be on transparency and accountability for the business model of digital platforms, which prioritises misinformation and disinformation.

5. What an ACMA standard should include

Australia's laws should be guided by global best practice and research.

Today, people in Australia have weaker protections online than people in Europe and parts of the United States. The EU's recently-enacted DSA, in particular, provides a framework for comprehensive regulation of the digital space that may be considered a template upon which Australian legislators can build.

When the DSA comes into full force in February 2024, it will apply to all digital services operating in the EU, including social media, online marketplaces, search engines and other online platforms. Its supervised, risk-based approach to harms caused by content on digital platforms, complements and reinforces the European Commission's *Strengthened Code of Practice on Disinformation*.[30] The DSA also marks a major shift away from co-regulation in Europe.

Effective regulation of digital platforms to protect human rights should include a comprehensive transparency framework and measures to protect privacy and give people greater control over what they see. These are discussed below, with reference to the model provided in the DSA.

5.1 Platforms should be subject to a comprehensive transparency, risk assessment, mitigation and audit framework

The Australian public currently has few meaningful opportunities to ever understand how platforms and algorithms shape the information environment in which we form views and make decisions. It is largely thanks to industry whistleblowers that we have achieved any scrutiny and accountability for the big tech companies.[31] Now and into the future, we should not need to rely on whistleblowers in order to understand the ways we are tracked and targeted, and the systems that determine the information that is delivered to us.

Accordingly, regulation should be focused on increasing transparency and accountability. Transparency features of an appropriate regulatory model should place the onus on digital platforms to identify the risks posed by their platforms and the steps they will take in response, as well as providing information to users about their advertising and recommender systems. ACMA should be be well-resourced to enforce these standards, and immediately empowered to verify digital platforms' information and hold them accountable for failures to report and act.

---

[30] *Code of Practice on Disinformation 2022* (EU), https://digital-strategy.ec.europa.eu/en/policies/code-practice-disinformation

[31] Shirin Ghaffary, 'Big Tech's employees are one of the biggest checks on its power' Vox, 29 December 2021 https://www.vox.com/recode/22848750/whistleblower-facebook-google-apple-employees

It is critical that comprehensive information about a digital platform's services, actions and associated risks is available to governments and civil society. This allows evolving forms of harm to be debated and addressed, and obligations enforced.

Under the DSA, very large platforms are required to undertake annual risk assessments to identify, analyse and assess any significant systemic risks stemming from the functioning and use of their services, including their algorithms, recommender systems, content moderation systems, terms and conditions, advertising systems or data-related practices.[32] In their risk assessments, very large platforms are specifically required to consider the following systemic risks:[33]

- **Actual or foreseeable negative impacts on a range of fundamental rights**, including the right to dignity, to respect for private and family life, protection of personal data, freedom of expression and information, the prohibition on discrimination, the rights of the child, and to a high-level of consumer protection.
- Any foreseeable negative effects on **civic discourse and electoral processes, and public security**.
- Any 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**.
- Dissemination of **illegal content** through their services.

It is notable that risk assessments are required to consider the impact of moderation systems on the right to freedom of expression and opinion. It is also significant that this scope extends to the impact of platforms' activities on a wide range of human rights, and encompasses risks associated with content that is legal but may pose serious societal risks. Further work should be done, however, to extend the risks beyond gender-based violence, to other forms of discriminatory violence, including racist, transphobic and other hate speech.

In response to systemic risks identified in risk assessments under the DSA, very large platforms are required put into place reasonable, proportionate and effective mitigation measures.[34] As a starting point, the DSA leaves decisions as to how to mitigate systemic risks to digital platforms. Regulators have significant power to step in where digital platforms do not take sufficient action, or where there is a need to ensure that they take a consistent or coordinated approach to systemic risks.[35] Risk assessments and mitigation measures are also subject to annual independent audits at digital platforms' expense.[36]

This framework for rights-focused risk assessments and mitigation is supported by robust information-gathering powers for the regulator, extending to all data necessary to assess compliance with the DSA.[37] The European Commission has also recently launched the European Centre for Algorithmic Transparency to support its enforcement efforts.

---

[32] Regulation (EU) 2022/2065 of the European Parliament and of the Council of 19 October 2022 on a Single Market For Digital Services and amending Directive 2000/31/EC, Article 34. Under the DSA, escalating obligations and requirements apply to platforms with the greatest influence and resources. *Very Large Online Platforms* are currently defined as those with 45 million or more average monthly users in the EU. Violations of the DSA can result in fines of up to 6% of a platform's annual worldwide turnover for a financial year, representing meaningful incentives for firms with the greatest influence and resources.

[33] Digital Services Act (EU), Article 34. By requiring entities to carry out human rights due diligence of this kind, risk assessments can support platforms to meet the standards set out in the UN Guiding Principles on Business and Human Rights, which is the authoritative international standard on the corporate responsibility to respect human rights.

[34] Digital Services Act (EU), Article 35.

[35] Digital Services Act (EU), Article 35(3).

[36] Digital Services Act (EU), Article 37.

[37] Digital Services Act (EU), Article 40.

Overall, this risk-based transparency framework is designed to provide layers through which systemic risks and shortfalls in addressing them can be identified and rectified. This is a strong and flexible model for ensuring that the primary obligation for avoiding and addressing harm is borne by platforms themselves, rather than individuals or sections of society.

5.2 Users should have control over what they see and how their data is used

People should have genuine choice in relation to the data they provide to platforms, how it is used and how information is delivered to them.[38] By giving people control over the use of their data and restricting the circumstances in which platforms can engage in intrusive tracking and profiling, regulation can reduce the scope for recommender systems to target and amplify harmful material.

Profiling refers to the platforms' practice of building a 'profile' of a person's personal attributes and interests through tracking their behaviour over time, which can then be used for targeted advertising and personalised recommender systems. Under the DSA, very large platforms are required to provide users with clear, accessible and comprehensible information on the parameters used in their recommender systems, and options to influence these parameters, with at least one option that is not based on profiling.[39]

The DSA prohibits targeted advertising based on types of sensitive data, including a person's ethnicity and political views,[40] and prohibits any form of profile-based advertising that targets children.[41] It also requires platforms to give users clear information about why they have been targeted with a particular piece of advertising. These rules complement and reinforce the right to object to profiling and targeted advertising based on profile data under the EU's General Data Protection Directive.[42] The Attorney-General's Department's recent review of the *Privacy Act 1988* recommends a right to opt-out of targeted advertising, and a prohibition on targeting based on sensitive information.[43] Under this proposal, platforms would still be able to collect personal information without consent, provided it is not sensitive information and users have the ability to opt out.

A regulatory model for digital platforms in Australia should ensure that people have genuine options to avoid tracking and profiling, and that platforms cannot rely on complex consent processes that nudge users towards an outcome that benefits the platform. Australia should move toward the prohibition of surveillance-based advertising by establishing restrictions in relation to both (i) the categories of data that can be processed for targeting purposes, and (ii) the categories of data that can be disclosed to third parties to facilitate targeted advertising.

Instead of permitting profiling by default and allowing users to opt out, regulation in Australia should go further by requiring default settings to not be based on profiling. This would ensure that users who are less aware of the operation of recommender systems will not be treated less favourably, and would limit the role of personalised content recommendation systems in amplifying disinformation and hate speech.

---

[38] The systematic collection of behavioural data and targeted advertising can violate the right to freedom of opinion; and the lack of transparency around platforms amplification of content online 'points towards an unacceptable level of intrusion into individuals' right to form their ideas free from manipulation and right to privacy.': UN Special Rapporteur on the promotion and protection of the right to freedom of opinion and expression, *Disinformation and freedom of opinion and expression*, 13 April 2021, UN Doc A/HRC/47/25, 14-15.

[39] Digital Services Act (EU), Article 38. Regulation (EU) 2016/679, Article 4(4).

[40] Digital Services Act (EU), Article 26(3). Sensitive data includes personal data referred to in Article 9 of the Regulation (EU) 2016/679.

[41] Digital Services Act (EU), Articles 28, 39; Regulation (EU) 2016/679, Article 4(4).

[42] General Data Protection Directive (EU), Articles 5, 6 and 9.

[43] Attorney-General's Department, *Privacy Act Review Report 2022*.

Recommendation 1

The federal Government should move away from self-regulatory and co-regulatory models for digital platforms, by amending the draft Bill to empower ACMA to immediately create and enforce standards that are grounded in international human rights law, similar to the EU's *Digital Services Act*. The draft Bill should also be amended to include:

(a) requirements for major platforms to undertake and publish risk assessments that identify human rights risks and other forms of harm, with corresponding obligations to develop and implement mitigation measures; and

(b) measures to give users greater control over the collection and use of their personal data, including by making recommender systems opt-in and by limiting forms of profiling.

6. Additional, minimum necessary amendments to the draft Bill

6.1 Strengthening the data access regime by extending it to civil society and expert researchers

The draft Bill should be amended to include a mandatory data access regime. The nature and impacts of disinformation, political polarisation, and harmful content ought to be open to scrutiny from experts and civil society researchers who can communicate their findings to Government policymakers and the public. Under the draft Bill, research undertaken internally by platforms will only become known to the public if the platforms choose to release it.

Under the DSA, on request from the relevant regulator, platforms are required to provide external researchers with access to data for the purposes of conducting research on detection, identification and understanding of the systemic risks to which risk assessments apply, and assessment of the adequacy, efficiency and impacts of platforms risk mitigation measures.[44] This data access regime is another significant feature of the DSA, and reinforces commitments made by signatories to the EU's *Strengthened Code of Practice on Disinformation*.

Effectively implemented, a robust data access regime can support civil society to operate as an 'early warning system' for emerging risks, as well as an additional avenue for accountability in relation to platforms' risk-mitigation commitments. During the pandemic, for example, an effective data access regime could have allowed public health officials, researchers and journalists timely access to comprehensive, anonymised data about COVID-19-related content on major platforms, including the role played by algorithmic amplification.

Concerns raised by digital platforms about the implications of sharing their data with governments and researchers – such as privacy and exploitation concerns – are legitimate, but they are also surmountable. Rather than allowing these concerns to outweigh the critical value of transparency, they should be addressed through appropriate safeguards for protecting sensitive data.

An immediately enforceable transparency regime that fosters the role of civil society will allow better insights into the impacts digital platforms have on individuals, social groups, and society as a whole. It will

[44] Digital Services Act (EU), Article 40(4).

allow the impact of digital platforms and their use to spread harmful material to be studied, reported on and debated.

6.2 Including the human rights to be protected in misinformation standards

Assessing and balancing competing human rights is not something the industry is equipped to do. This should remain the purview of the democratically accountable Australian Government.

The draft Bill should be amended to clearly state the human rights that should be considered and protected by misinformation standards, be they developed by industry or ACMA. This will help empower ACMA to ensure all relevant rights are adequately considered in misinformation standards, and competing rights are proportionately balanced against one another.

Recommendation 2

The federal Government should at minimum strengthen the draft Bill by:

(a) including a data access regime to support civil society and expert research into the amplification of disinformation and hate speech; and

(b) explicitly including the human rights to be protected in misinformation standards, be they developed by industry or ACMA.

Should the Department wish to discuss anything in this submission further, please do not hesitate to get in touch with Alice Drury or David Mejia-Canales.