Human Rights Law Centre

Submission to the inquiry into the influence of international digital platforms

Senate Economics References Committee

March 2023 | Pre-Committee Version

Human
Rights
Law
Centre

Scott Cosgriff
Senior Lawyer
Human Rights Law Centre Ltd
Level 17, 461 Bourke Street
Melbourne VIC 3000

T: + 61 3 8636 4450
F: + 61 3 8636 4455
E: scott.cosgriff@hrlc.org.au
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

1. Introduction and recommendations 4

2. Human rights are being undermined by disinformation and hate speech 5

3. Current approaches are not working 8

3.1 Measures to address online harm must look beyond content moderation 8

3.2 Self-regulation and co-regulation are ineffective 9

4. Better approaches to regulating digital platforms 9

4.1 Platforms should be subject to a comprehensive transparency framework 10

4.2 Users should have control over what they see and how their data is used 12

1. Introduction and recommendations

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 welcomes the opportunity to contribute to this inquiry. This submission addresses the inquiry's terms of reference relating to international models for regulation of digital platforms and the need for greater transparency in relation to the use of algorithms, in order to address the significant human rights implications of disinformation and hate speech online.

The Federal Government has indicated that it will introduce legislation in 2023 to strengthen its powers to address disinformation on digital platforms, and is currently engaged in a review of Australian privacy law with significant implications for the digital space.[2] In the process, it is vital we learn from the regulatory models of our global counterparts. The EU's recently-enacted *Digital Services Act* represents an instructive example of a comprehensive and durable governance framework for the digital space.

It is crucial we recognise, as they have in Europe, that self-regulation is not working and should not be expected to work. Digital platforms are currently operating with both unprecedented reach and strong incentives to amplify disinformation. Meanwhile, current arrangements in Australia allow these platforms, some of the world's most powerful companies, to externalise the risks and social costs of their business models. This needs to change.

Regulation that makes digital platforms more transparent and accountable, while giving users greater control over their data, can put Australia on the path to a safer, more democratic internet.

[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/digital-platforms/news-media-bargaining-code/news-media-bargaining-code

[2] See eg Attorney-General's Department, *Privacy Act Review Report 2022*; The Hon Michelle Rowland MP, 'New ACMA powers to combat harmful online misinformation and disinformation' 20 January 2023, https://minister.infrastructure.gov.au/rowland/media-release/new-acma-powers-combat-harmful-online-misinformation-and-disinformation

Recommendations

1. The Federal Government should move away from self-regulatory and co-regulatory models for digital platforms, by replacing existing co-regulatory codes and ensuring new regulations are written by legislators or regulators.

2. The Federal Government should introduce a comprehensive digital regulatory framework for Australia, focused on transparency and risks arising from platforms' systems and processes.

   The design of this framework should include:

   - 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;
   - a data access regime to support civil society research and enhance transparency;
   - 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;
   - broad information-gathering and enforcement powers for an independent, well-resourced and integrated regulator.

3. Parliament should consider mechanisms for enhancing parliamentary scrutiny of digital regulation, such as the establishment of a dedicated committee for digital affairs, and for improving the coordination of tech policy across government.

2. Human rights are being undermined by disinformation and hate speech

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 harmful material online has a profound impact on human rights and democratic processes.

Disinformation and hate speech are always simmering beneath the surface, in some cases with the potential to peak and form violent movements where it is amplified to millions. To use one recent example, the videos of misogynist Andrew Tate, who frequently promotes violence against women to his young male audience, had been viewed 11.6 billion times on TikTok before August last year.[3] Teachers are now coming forward to say that Tate's misogynistic "ideology" has taken hold of the teenage boys in their classrooms, and that it is fuelling abuse against female teachers and students.[4]

[3] Shanti Das, 'Inside the violent, misogynistic world of TikTok's new star, Andrew Tate', The Guardian, 7 August 2022, https://www.theguardian.com/technology/2022/aug/06/andrew-tate-violent-misogynistic-world-of-tiktok-new-star

[4] Anna Fazackerley, '"Vulnerable boys are drawn in": schools fear spread of Andrew Tate's misogyny', The Guardian, 8 January 2023, https://www.theguardian.com/society/2023/jan/07/andrew-tate-misogyny-schools-vulnerable-boys;

Regulation of digital platforms needs to be able to address societal risks of this nature and scale, without relying on measures that threaten freedom of expression and the right to access information.

Australia has recently witnessed major, real-world examples of how seriously disinformation can affect human rights and democratic processes.

Misinformation and disinformation threatened public health in the pandemic

The COVID-19 pandemic highlighted both the rapidly evolving nature of online mis- and disinformation, as well as its potential to undermine public health and fuel discrimination. Misleading content about the origin and nature of the virus spread rapidly across digital platforms in Australia.[5] This disinformation combined with hate speech online to fuel discrimination and stoke violence against Asian people in Australia, threatening their safety.[6] Australian health authorities pointed to the spread of online mis- and disinformation contributing to a spike in cases during a critical period in 2000.[7] Platforms, including signatories to the voluntary, industry-drafted code on mis- and disinformation, saw a rapid increase in engagement with groups peddling 'anti-vaxx' and vaccine hesitant content in Australia.[8]

Disinformation risks undermining democratic elections in Australia

Online disinformation has 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.[9]

The amplification of mis- and disinformation has contributed to a deterioration of democracy in the US since 2016. The 2020 US presidential election was defined by disinformation campaigns, undermining confidence in the election, and shaping the conditions that led to the storming the US Capitol building in an attack that left 5 dead and 140 police officers injured.[10]

Lola Okolosie, 'Parents, talk to your sons about Andrew Tate – we teachers can't take him on alone', The Guardian, 14 February 2023, https://www.theguardian.com/commentisfree/2023/feb/14/parents-sons-andrew-tate-teachers-toxic-influencers

[5] See eg ACMA, 'Misinformation and news quality on digital platforms in Australia: A position paper to guide code development', June 2020, *Appendix A: Case study – misinformation during COVID-19*.

[6] See eg Centre for Media Transition, University of Technology Sydney, *Information Disorder Lessons from Australia,* 9 December 2022, p. 21.

[7] Melissa Davey and Matilda Boseley, 'Coronavirus Victoria: experts warn against blaming migrant communities for spreading misinformation', The Guardian, 28 June 2020, https://www.theguardian.com/australia-news/2020/jun/28/coronavirus-victoria-experts-warn-against-blaming-migrant-communities-for-spreading-misinformation; AFP Australia, 'Fake graphic misrepresents government data on Covid-19 vaccines in Australia' 2 June 2021, https://factcheck.afp.com/fake-graphic-misrepresents-government-data-covid-19-vaccines-australia

[8] Reset Australia, 'Anti-vaccination & vaccine hesitant narratives intensify in Australian Facebook Groups', 2021 https://au.reset.tech/uploads/resetaustralia_social-listening_report_100521-1.pdf

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

[10] Lois Beckett, 'Facts won't fix this: Experts on how to fight America's disinformation crisis', The Guardian, 1 January 2021, https://www.theguardian.com/us-news/2021/jan/01/disinformation-us-election-covid-pandemic-trump-biden

Disturbingly similar narratives emerged in the Australian 2022 federal election.[11] For example, ahead of the election, an animated cartoon posted by Pauline Hanson spreading false claims about election fraud in Australia was viewed almost 100,000 times.[12]

Facebook identified 2.2 billion fake accounts as engaging in "coordinated inauthentic behaviour" in the lead-up to the 2019 election.[13] Local disinformation campaigns, such as Mediscare in 2016 and Death Tax in 2019, are becoming a common feature of Australian elections as campaigns and news consumption move further online.[14]

While elections and the pandemic have provided clear warnings, disinformation and hate speech are having an impact on a spectrum of human rights that the Australian Government has legal obligations to protect. These obligations extend to the government's role in regulating digital platforms and the activities of big tech companies, and involve specific rights in the case of children.[15] Internationally, there is increasing recognition that governments must place human rights at the centre of bold action to combat online harm.[16]

In addition to the role of the Australian Government, companies themselves are expected to respect human rights in their activities and operations in line with the *UN Guiding Principles on Business and Human Rights*. At a minimum, this involves regular human rights impact assessments of their products, operations and policies and due diligence processes aimed at identifying, preventing or mitigating actual or potential adverse impacts on human rights.[17]

---

[11] Josh Butler, 'From 'voter fraud' to Albanese's 'invalid' swearing in, election claims debunked', The Guardian, 30 May 2022, https://www.theguardian.com/australia-news/2022/may/30/from-voter-to-albaneses-invalid-swearing-in-election-claims-debunked

[12] Nick Bonyhady, 'Social media giants pull One Nation satire video over voter fraud claims', Sydney Morning Herald, 29 April 2022, https://www.smh.com.au/technology/social-media-giants-pull-one-nation-satire-video-over-voter-fraud-claims-20220429-p5ah4k.html

[13] Felicity Caldwell, 'Bots stormed Twitter in their thousands during the federal election', Sydney Morning Herald, 20 July 2019, https://www.smh.com.au/politics/federal/bots-stormed-twitter-in-their-thousands-during-the-federal-election-20190719-p528s0.html

[14] Katharine Murphy, Christopher Knaus and Nick Evershed, '"It felt like a big tide": how the death tax lie infected Australia's election campaign', The Guardian, 8 June 2019, https://www.theguardian.com/australia-news/2019/jun/08/it-felt-like-a-big-tide-how-the-death-tax-lie-infected-australias-election-campaign

[15] 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, [83]-[94]; UN Committee on the Rights of the Child, *General Comment No. 25 (2021) on Children's Rights in Relation to the Digital Environment* (2 March 2021) UN Doc CRC/C/GC/257.

[16] See eg UN Human Rights Council, *Role of States in countering the negative impact of disinformation on the enjoyment and realization of human rights*, 30 March 2022, UN Doc A/HRC/49/l.31/Rev.1; 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; UN Human Rights Committee, *General comment no. 34 – Article 19: Freedoms of opinion and expression*, 12 September 2011, UN Doc CCPR/C/GC/34; UN Committee on the Rights of the Child, *General Comment No. 25 (2021) on Children's Rights in Relation to the Digital Environment* (2 March 2021) UN Doc CRC/C/GC/257. See also: Access Now, Civil Liberties Union for Europe and EDRi, *Informing the disinfo debate: a policy guide for protecting human rights*, (December 2021) https://www.accessnow.org/cms/assets/uploads/2021/12/Informing-the-disinfo-debate-report.pdf; Amnesty International, *A human rights approach to tackle disinformation*, 14 April 2022, https://www.amnesty.org/en/wp-content/uploads/2022/04/IOR4054862022ENGLISH.pdf

[17] 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.

3. Current approaches are not working

3.1 Measures to address online harm must look beyond content moderation

Australia's primary existing mechanism for addressing disinformation online is the voluntary industry-drafted *Australian Code of Practice on Disinformation and Misinformation* (the **DIGI Code**).[18] The measures contemplated by the DIGI Code for reducing the spread of mis- and disinformation, and the transparency reports prepared by major platforms about the steps they are taking, reflect a strong emphasis on identifying and moderating content.[19]

Long-term approaches to addressing disinformation and other harmful content must go beyond reliance on content takedown and flagging. The sheer volume of disinformation and hate speech renders content moderation an endless game of whack-a-mole, while the lag typically involved in content moderation means action is often taken only after damage is done. In addition, defining and identifying disinformation and harmful content is difficult, especially in real time and across hundreds of languages and countless societal contexts, and platforms are poorly placed to arbitrate the appropriateness of political content.[20]

In Australia, victims of online hate speech have faced difficulty asserting their right to freedom from discrimination. For example, platforms have argued that they are beyond the reach of Australian privacy and anti-discrimination laws due to their corporate structure and incorporation in other countries.[21] The Committee may wish to investigate how gaps of this nature can be effectively addressed, while recognising that the burden of identifying, avoiding and responding to harm should be borne by platforms, and not the individuals and communities that are affected by it.

Mechanisms that rely on content moderation also pose a risk of censorship. Around the world, there is a growing body of evidence of excessively broad and vague laws becoming tools of governments to compel private companies to police communication in ways that unjustifiably limit public debate and freedom of expression.[22] Responses to the problem of disinformation, whether on the part of governments, regulatory bodies or platforms themselves, must not infringe upon the right to freedom of expression and the right to access information – two cornerstones of democratic discourse.

Regulation that relies on content moderation can lead to limits on these rights in circumstances where it was never intended. This is because penalising platforms for content-moderation failures incentivises

---

[18] DIGI, *Australian Code of Practice on Disinformation and Misinformation*, 22 December 2022, https://digi.org.au/disinformation-code

[19] See eg DIGI Code, section 5.9; See eg the reports of Meta, Twitter and TikTok: DIGI, *2021 Transparency Reports*, https://digi.org.au/disinformation-code/transparency

[20] According to one document released as part of the 'Facebook Papers' 87% of Facebook's operational budget for classifying misinformation goes towards the United States, while 13% is set aside for the rest of the world (despite the fact that North American users make up just 10% of its user base): Sheera Frenkel and Davey Alba, 'In India, Facebook grapples with an amplified version of its problems' New York Times, 23 October 2021 https://www.nytimes.com/2021/10/23/technology/facebook-india-misinformation.html. See also Julia Angwin and Hannes Grassegger, 'Facebook's secret censorship rules protect white men from hate speech but not black children', ProPublica, 28 June 2017, https://www.propublica.org/article/facebook-hate-speech-censorship-internal-documents-algorithms

[21] See eg the example of the complaints brought by the Australian Muslim community in Queensland: Australian Muslim Advocacy Network, Submission to the Senate Economics References Committee inquiry into the influence of international digital platforms, 28 February 2023.

[22] Amnesty International, *A human rights approach to tackle disinformation: Submission to the Office of the High Commissioner for Human Rights*, 14 April 2022, https://www.amnesty.org/en/wp-content/uploads/2022/04/IOR4054862022ENGLISH.pdf

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.

By focusing on content moderation alone, governments miss the opportunity to address the upstream drivers of disinformation and hate speech, which will be far more effective in the long term.

For all these reasons, content moderation ought to be seen as only one part of any comprehensive and effective framework for digital regulation.

3.2 Self-regulation and co-regulation are ineffective

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.[23] Co-regulation is inappropriate for such a powerful and high-risk sector, in which business models frequently come into conflict with community needs and the public interest. In the European Union, introduction of the *Digital Services Act* was driven by growing recognition that self- and co-regulatory models are inadequate and ineffective.

Regulator-drafted industry standards should be the norm. The Federal Government is expected to introduce legislation in 2023 that will grant the Australian Communications and Media Authority new powers to oversee digital platforms and to introduce an enforceable code in relation not misinformation and disinformation. It is essential that the regulation of digital platforms in Australia evolve toward an enforceable legal framework drafted by regulators and lawmakers and overseen by an appropriately resourced and empowered regulator.

The role of digital platforms in our lives and communities will continue to evolve in ways that require ongoing attention from policy makers and the Parliament. Accordingly, we encourage the Committee consider options for enhancing the Parliament's engagement with emerging tech issues and improving the coordination of tech policy across government. This could include the implementation of specific models for tech policy coordination,[24] and at the parliamentary level, a dedicated Parliamentary committee on digital matters similar to those in several European countries.

4. Better approaches to regulating digital platforms

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 *Digital Service Act* (**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*.[25] The DSA also marks a major shift away from co-regulation in Europe.

---

[23] 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

[24] See eg Tech Policy Design Centre, Australian National University, *Cultivating Coordination*, February 2023, https://techpolicydesign.au/wp-content/uploads/2023/02/TPDC_Cultivating_Coordination_2_20230221.pdf.

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

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.

4.1 Platforms should be subject to a comprehensive transparency 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.[26] 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. 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.

(A) RISK ASSESSMENTS, MITIGATION MEASURES AND INDEPENDENT AUDITS

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.[27] In their risk assessments, very large platforms are specifically required to consider the following systemic risks:[28]

- **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.

[26] 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

[27] 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 (**Digital Services Act**), 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.

[28] 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.

It is 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. It is also notable that risk assessments are required to consider the impact of moderation systems on the right to freedom of expression and opinion.

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.[29] 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.[30] Risk assessments and mitigation measures are also subject to annual independent audits at digital platforms' expense.[31]

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.[32] The European Commission has also recently launched the European Centre for Algorithmic transparency to support its enforcement efforts.

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.

(B) DATA ACCESS REGIME

The nature and impacts of disinformation, political polarisation, and harmful content ought to be open to scrutiny from civil society researchers who can communicate their findings to policymakers and the public. Currently, 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.[33] 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.

---

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

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

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

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

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

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

4.2 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.[34] 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.[35]

The DSA prohibits targeted advertising based on types of sensitive data, including a person's ethnicity and political views,[36] and prohibits any form of profile-based advertising that targets children.[37] 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.[38] 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.[39] 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.

---

[34] 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.

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

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

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

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

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