CENTRE FOR MEDIA TRANSITION

Statutory Review of the Online Safety Act 2021

Issues Paper, April 2024

Submission from UTS Centre for Media Transition to the Department of Infrastructure, Transport, Regional Development, Communications and the Arts

Date: 28 June 2024

About the Centre for Media Transition

The Centre (CMT) was established in 2017 as an applied research unit based at the University of Technology Sydney (UTS). It is an interdisciplinary initiative of the Faculty of Arts and Social Sciences and the Faculty of Law, sitting at the intersection of media, journalism, technology, ethics, regulation and business.

Working with industry, academia, government and others, the CMT aims to understand media transition and digital disruption, with a view to recommending legal reform and other measures that promote the public interest. In addition, the CMT aims to assist news media to adapt for a digital environment, including by identifying potentially sustainable business models, develop suitable ethical and regulatory frameworks for a fast-changing digital ecosystem, foster quality journalism, and develop a diverse media environment that embraces local/regional, international and transnational issues and debate.

This submission was prepared by:

- Dr Michael Davis, Centre for Media Transition, Faculty of Law
- Dr Karen Lee, Faculty of Law
- Professor David Lindsay, Faculty of Law
- Professor Derek Wilding, Co-Director, Centre for Media Transition, Faculty of Law

CONTACT

Centre for Media Transition
Faculty of Law, University of Technology Sydney
Building 2, Level 15
UTS City Campus, Broadway
PO Box 123, Broadway NSW 2007

[redacted]

cmt.uts.edu.au

Executive Summary

Thank you for the opportunity to contribute to this consultation on the review of the *Online Safety Act 2021* (Cth) (OSA). In the main part of this submission we respond to selected questions from the Issues Paper. As there is necessarily some overlap in the topics covered by these questions, this Executive Summary provides a more thematic presentation of our main points.

Regulatory design

- We support an approach based on systemic regulation that shifts the regulatory focus from specific harms to platforms' systems and processes. This would include elements such as a statutory duty of care and safety by design.
- We support the introduction of a single, legally enforceable statutory duty of care that would place a general obligation on digital platforms to address the risks and harms arising from their systems and processes. However, as the proposed duty is not a magic bullet, we support a hybrid approach that retains *ex post* regulation in the form of notice and takedown provisions for illegal material.
- The duty of care should replace, and build on, the Basic Online Safety Expectations. Unlike the Basic Online Safety Expectations, however, the duty should be legally enforceable, with adequate penalties for non-compliance.
- The duty must be accompanied by adequate transparency measures. These would include obligations to undertake systemic risk assessments, to commission independent audits of platform risks, and actions taken to minimise risks and to provide sufficiently detailed information to the regulator (subject to confidentiality agreements) about the operation of algorithms to allow for assessment of algorithmic risk.
- Details of how to comply with a statutory duty of care could be set out in industry codes or standards, as in the UK; alternatively, as in the EU, codes could offer one way of meeting due diligence obligations, without precluding enforcement action by the regulator if it considers a provider has not met the overall duty or standard. In Australia, a statutory duty of care could build on the co-regulatory arrangements for code development under the current regime.
- The systemic approach with a duty of care should impose greater obligations on service providers that pose the greatest risk and have the greatest reach, with a clear understanding of the approach taken to defining and prioritising risks. This would facilitate regulation being focused on the systems and processes of high impact services, such as popular social media and search services.
- In accordance with this re-calibrating of regulation of online services to the risk and reach of those services, there is scope – over time – for some simplification and rationalisation of the service categories established under the Act. However, given the work that has already been undertaken in establishing codes and standards based on the existing industry sectors, we do not consider this to be an immediate priority.
- The OSA should be recast to expressly acknowledge that other rights and interests which, in this context, include the rights to freedom of expression and privacy, should be considered along with safety. Any balancing of fundamental rights and interests should incorporate the proportionality principle. The duty of care itself should incorporate the 'best interests of the child' principle, but it should be more broadly framed to serve the interests of the Australian community in general. Moreover, the formulation of the statutory duty must make it clear that it incorporates a duty to design algorithms in accordance with a binding safety by design framework.

- While technological neutrality in regulatory design is appealing, laws need to be framed in such a way that they effectively address the sources of the most serious online risks and harms; consequently, they must be able to be applied to existing technologies. Moreover, to the extent that specific technologies create specific risks or harms, laws should be designed to address those harms. In other words, laws should be as 'neutral' or 'specific' as is necessary to meet the regulatory objectives

Complaints

- We do not support a private cause of action for enforcing the proposed duty; instead, we favour public enforcement by the regulator coupled with complaints handling by an independent ombuds scheme. The scheme would be funded by industry.
- An external dispute resolution mechanism is required and it should take the form of an ombuds scheme. The ombuds would handle individual and systemic complaints; the office of the eSafety Commissioner would retain its existing functions associated with rule-making and compliance and enforcement of regulation.
- There is a need for a single ombuds scheme to deal with 'transactional' and 'social' complaints arising from the use of digital platforms. While this review only addresses matters dealt with under the Online Safety Act, the ACCC has separately recommended there should be an ombuds scheme for 'transactional' complaints. These complaints concern the conduct of the digital platforms involving customers' unmet contractual expectations and/or infringement of an amended Australian Consumer Law. In our research on options for digital complaints handling (referred to below) we said that the ACCC's recommendations for transactional complaints should not be considered in isolation from the question of how to address the 'social' complaints that arise under the OSA. We restate that position here: any recommendations coming out of this review of the online safety regime should be considered alongside the need for action on transactional complaints.
- In principle, we support the acceptance of 'third party' complaints. However, this approach could be difficult to administer in the high-volume environment of social media and other forms of publication and distribution by digital platforms. We suggest a reasonable accommodation could be made by allowing representative groups to make complaints.

Decisions by the regulator

- While in general, eSafety is a responsive and transparent regulator that does a good job of explaining its role and its actions to the public, the failure to provide sufficient information about its action on the Wakeley stabbing video exposed a gap in public information about eSafety's decisions. There should be a public register of the decisions of eSafety.

Research

- We support the introduction into the OSA of an obligation similar to that in the EU Digital Services Act under which platforms have an obligation to provide properly vetted researchers with access to data for the purposes of assessing systemic risks and risk mitigation measures.

Response to Questions in Part 2: Australia's regulatory approach to online services, systems and processes

Q1: Are the current objects of the Act to improve and promote online safety for Australians sufficient or should they be expanded?

- We address this point in response to Question 26 below where we recommend that the protection and promotion of fundamental rights, including but not restricted to the rights of users, should be expressly incorporated into the OSA.

Q2: Does the Act capture and define the right sections of the online industry?

- The OSA establishes a comprehensive but complex regulatory regime based on eight industry sectors. As the Issues Paper correctly points out, there are potential problems with this approach: it adds to the complexity of drafting single industry codes for specific sectors and it may not keep pace with technological change, creating the potential for gaps in the regulatory regime. Moreover, unlike the UK and EU regimes, it is not primarily based on the risk and reach of online services.
- The UK Online Safety Act applies greater regulatory obligations to three categories of service provider (Categories 1, 2A and 2B), based on risk and reach. Similarly, the EU Digital Services Act applies the most stringent rules to very large online platforms and very large search engines, which are those with over 45 million users in the EU.
- As explained further in our response to Question 7, we consider there is scope for more precisely calibrating regulation of online services to the risk and reach of those services. We also consider that, over time, there is scope for some simplification and rationalisation of the service categories established under the Act. However, given the work that has already been undertaken in establishing codes and standards based on the existing industry sectors, we do not consider this to be an immediate priority.

Q4: Should the Act have strengthened and enforceable Basic Online Safety Expectations?

- On Question 4, see our response to Question 22 below which proposes that a new statutory duty of care should replace and build on the Basic Online Safety Expectations.

Q7: Should regulatory obligations depend on a service providers' risk or reach?

- Given the considerable challenges raised by the scale of regulating the online environment, it is natural that there has been a move towards risk-based regulation, where regulation is focused on those activities presenting the greatest risks. The adoption of risk-based regulation seems to reflect common sense, in that regulators with scarce resources can prioritise those issues that are most important. It is therefore hardly surprising that risk-based regulation has been expressly adopted in legislative instruments designed to regulate technologies at scale, such as the EU AI Act and Digital Services Act. We also note that considerable work has already been undertaken in building risk assessment into the online safety codes, such as the Social Media Services Online Safety Code.
- Despite its attractions, however, as regulatory experts such as Baldwin and Black have pointed out, experience indicates there are significant challenges in implementing risk-based regulation.[1] In particular, there are considerable challenges in both the selection and prioritisation of risks, including the extent to which this may be based on value assumptions. Moreover, as with any form of risk assessment, it is always necessary to take into account both the probability of the risk occurring and the degree of harm that may arise: it is necessary to provide for both low probability events that may result in catastrophic harms, and high probability events that lead to lesser, but still important harms. In the immediate context of the regulation of online content, and the examples set by the UK and EU, it is likely that there is a correlation between risk and reach, although this correlation may not be exact.
- In general, we support a regulatory regime that imposes greater obligations on service providers that pose the greatest risk and have the greatest reach. This would effectively conserve regulatory resources and minimise regulatory costs imposed on low risk service providers. If, as outlined in our response to Questions 21 and 22, a systemic approach is taken, this would facilitate regulation being focused on the systems and processes of high impact services, such as popular social media and search services. To be successful, however, this approach must be accompanied by a clear understanding of the approach taken to defining and prioritising risks, as well as the limitations of risk-based approaches.

[1] Robert Baldwin & Julia Black, 'Driving Priorities in Risk-based Regulation: What's the Problem?' (2016) 43(4) *Journal of Law & Society* 565

Response to Questions in Part 3: Protecting those who have experienced or encountered online harms

Q14: Should the Act empower 'bystanders', or members of the general public who may not be directly affected by illegal or seriously harmful material, to report this material to the Commissioner?

- In the Online Safety Act, complaints by members of the public (ie, not people directly affected) can only be made to the eSafety Commissioner under the online content scheme, not under the three other schemes. In principle, we support the acceptance of what are sometimes referred to as 'third party' complaints. There will be circumstances where there is a broader public interest in taking action against content that is harmful to an individual, even where that person is not themselves able or willing to submit a complaint. This is the reason that the Australian Press Council (APC), among other complaints handling bodies, allows such complaints (referred to by the APC as 'secondary complaints').[2]
- In practice, however, this approach could be difficult to administer in the high-volume environment of social media and other forms of publication and distribution by digital platforms, especially as consultation with a person directly affected is needed. To the extent that resources are available to commit to such action, we support it, but we would not place this proposed expansion of eSafety's functions above other essential reforms. We suggest a reasonable accommodation could be made by allowing representative groups to make complaints. The *Competition and Consumer Amendment (Fair Go for Consumers and Small Business) Act 2024* (Cth) does this by way of 'designated complainants' who are approved by the Minister to make a complaint to the ACCC. Under that scheme, approval may be granted after taking account of aspects such as 'the experience and ability of the applicant in representing the interests of consumers or small businesses (or both) in Australia in relation to a range of market issues that affect them' (see s 154ZQ(2)(a)).

[2] For the past decade, the APC has distinguished adjudications relating to these matters from those lodged by the person who is directly affect by using the term 'Complainant' in the title. For a recent example, see: Complainant / The Australian, Adjudication 1845, 15 March 2024.

Response to Questions in Part 5: International approaches to address online harms

Q21: Should the Act incorporate any of the international approaches identified above? If so, what should this look like?

- The OSA review is an ideal opportunity for Australia to learn from emerging international approaches to regulating online content. As the Issues Paper points out, in response to the challenges of regulating online content, internationally there has been a shift from *ex post* episode-based interventions – such as notice-and-takedown regimes – towards *ex ante* systemic regulation. Acknowledging the significant difficulties of regulating the speed, ubiquity and scale of online content distribution, and the importance of preventing harms rather than redressing specific harms after they occur, systemic regulation shifts the regulatory focus from specific harms to platforms' systems and processes. Systemic regulation includes elements such as a statutory duty of care and safety by design. That said, *ex ante* systemic regulation and *ex post* complaints-based regulation are not necessarily mutually exclusive: a hybrid regulatory regime may incorporate elements of both approaches. **In general, we support a hybrid regulatory approach**, with some of the details spelt out in our responses below.

Q22: Should Australia place additional statutory duties on online services to make online services safer and minimise online harms?

- In a series of publications and research reports, UK academics, Lorna Woods and William Perrin, proposed a systems-level approach to platform regulation centred on a statutory duty of care.[3] As Woods points out, this approach is directed at addressing the distinctive features of digital platforms, which are not content creators, but whose design decisions condition (and potentially manipulate) user choices and experiences.[4] In other words, platforms are not neutral in relation to content, as their design decisions and business imperatives may create or exacerbate online harms. Accordingly, an approach based purely on taking down or restricting access to specific content does not address the source of many of the risks and harms.
- The main arguments *in favour of* a statutory duty of care are as follows:

    1. It shifts the burden of addressing harms from individual users that incur harms to the entities – that is, digital platforms – that are best placed to take

[3] Their version of a statutory duty of care is commonly known as the 'Carnegie Proposal': see Lorna Woods and William Perrin, *Online Harm Reduction – A statutory duty of care and regulator,* A proposal for Carnegie UK Trust, April 2019, https://ssrn.com/abstract=4003986; Lorna Woods, 'The Duty of Care in the Online Harms White Paper' (2019) 11(1) *Journal of Media Law* 6.
[4] Lorna Woods, 'Introducing the Systems Approach and the Statutory Duty of Care' in Judit Bayer et al (eds), *Perspectives on Platform Regulation* (2021, Nomos, Baden Baden) 77-98.

steps to prevent harms. In general, this accords with the economic principle of 'least cost avoider'.[5]

2. By imposing a general standard rather than detailed rules, it avoids the costs involved with developing more detailed rules to apply to a wide diversity of circumstances that are subject to rapid change.[6] While less certain in application, a general standard is more flexible and adaptable to a range of circumstances than specific rules.[7] Moreover, as digital platforms have more information about the operation of their systems, they are better placed than legislators or regulators to know how a general standard might be operationalised in the context of rapidly changing technologies.
3. Appropriately designed, a statutory duty of care can be applied to systems and processes, which are the source of risks and harms, rather than on providing redress for specific harms.

- The main arguments *against* a statutory duty of care are that it could create an incentive for platforms to proactively block or filter content, potentially resulting in 'over policing' and threatening freedom of expression; or that it could result in generalised monitoring of users, threatening the right to privacy.[8]
- **On balance, we support the introduction of a legally enforceable statutory duty of care that would place a general obligation on digital platforms to address the risks and harms arising from their systems and processes**. That said, much depends on the way in which a proposed new duty of care is designed. We make the following observations in relation to some of the key issues that arise in designing a statutory duty of care.

As is the case with the UK Online Safety Act, details of how to comply with a statutory duty of care could be set out in industry codes or standards; alternatively, as is the case in the EU, codes could offer one way of meeting due diligence obligations, without precluding enforcement action by the regulator if it considers a provider has not met the overall duty or standard. **In Australia, a statutory duty of care could build on the co-regulatory arrangements for code development under the current regime.**

1. The proposed duty of care should replace, and build on, the Basic Online Safety Expectations. Unlike the Basic Online Safety Expectations, however, the duty should be legally enforceable, with adequate penalties for non-compliance.
2. At this stage, we do not support a private cause of action for enforcing the proposed duty but favour public enforcement by the regulator. To be effective, however, the regulator must be adequately resourced.
3. While noting that the content of a generalised duty of care is context-dependent, we agree with those arguing that Australia should introduce a single duty of care rather than the multiple, overlapping duties of care found

[5] See Guido Calabresi, *The Costs of Accidents: A Legal and Economic Analysis* (Yale University Press, 1970); Paul Rosenzweig, 'Content Moderation and the Least Cost Avoider' (2024) *Joint PIJIP/TLS Research Paper Series* 125, https://digitalcommons.wcl.american.edu/research/125.
[6] On the distinction between 'standards' and 'rules' see eg. Frederick Shauer, 'The Convergence of Rules and Standards' [2003] *New Zealand Law Journal* 303.
[7] As Woods and Perrin put it: 'The statutory duty of care approach is not a one-off action but an ongoing, flexible and future-proofed responsibility that can be applied effectively to fast-moving technologies and rapidly emerging new services': Woods and Perrin (2019) p. 13.
[8] See eg. Joint Civil Society Briefing on the Online Safety Bill for the House of Lords (January 2023), https://www.libertyhumanrights.org.uk/wp-content/uploads/2022/04/Joint-civil-society-briefing-on-private-messaging-in-the-Online-Safety-Bill-for-Second-Reading-in-the-House-of-Lords-January-2023.pdf.

in the UK Online Safety Act.[9] A single duty of care would prevent risks and harms from falling between the cracks and ensure that the focus remains on systems and processes rather than specific kinds of content.

4. The proposed duty must be accompanied by adequate transparency measures, such as a requirement to report regularly on compliance with the duty. In our view, the regulator should also have power to proactively initiate investigations.

As the proposed duty is not a magic bullet, we support a hybrid approach which retains *ex post* regulation in the form of notice and takedown provisions for illegal material. We think this gives appropriate recognition to the impactful work of the eSafety Commissioner in addressing the real harms caused to individuals in Australia in relation to cyber-bullying, adult cyber abuse, and image-based abuse. We do note, however, that there would be some change to the way in which the notification and initial consideration of these matters is handled if, as we propose below, an independent ombuds is established. In that case, the ombuds would make a decision on whether the content breached the standard; it would then seek a resolution with the service provider; and if a remedy was not achieved or a request was not complied with, the ombuds would refer the matter to the regulator for enforcement. Finally, we note that these actions by either the eSafety Commissioner or an ombuds constitute important elements in a scheme that does not provide – and, as noted above, we think this is appropriate – an individual cause of action.

- The proposed duty should incorporate the 'best interests of the child' principle, but this should not be the overriding principle that guides the legislation. The duty should be more broadly framed to serve the interests of the Australian community in general, incorporating safety by design principles. Moreover, the formulation of the statutory duty must make it clear that it incorporates a duty to design algorithms in accordance with a binding safety by design framework.[10]
- To prevent over-reach, and protect user rights, the proposed duty must be accompanied by adequate safeguards. To protect users' right to privacy, there should be a prohibition on generalised monitoring of users, such as that included in the EU Digital Services Act. Similarly, it should be clear that the duty does not include an obligation on platforms to access or decrypt encrypted messages. Further to this, we recommend that the difficult issue of access to encrypted content should be addressed separately from the proposed duty of care. Furthermore, acknowledging problems that may arise in automated content moderation, such as over-blocking of legitimate content, it is important to include safeguards for freedom of expression. As under the UK Online Safety Act, any general duty of care should incorporate duties to protect users' freedom of expression and privacy.[11]

[9] Rhys Farthing and Lorna Woods, 'The Dangers of Pluralism: A singular duty of care in the Online Safety Act', *The Policymaker, https://thepolicymaker.jmi.org.au/the-dangers-of-pluralisation-a-singular-duty-of-care-in-the-online-safety-act/*.

[10] See Esme Fowler-Mason, 'The Online Safety Bill Needs More Algorithmic Accountability to Make Social Media Safe', LSE Department of Media & Communications, 8 February 2023, https://blogs.lse.ac.uk/medialse/2023/02/08/the-online-safety-bill-needs-more-algorithmic-accountability-to-make-social-media-safe/.

[11] See *Online Safety Act 2023* (UK) s 22.

Q23: Is the current level of transparency around decision-making by industry and the Commissioner appropriate? If not, what improvements are needed?

Transparency of decision-making by platforms

- Taking into account developments in comparable jurisdictions, we believe there is a good case for imposing more transparency requirements on intermediaries, especially those responsible for the most significant risks, such as social media platforms and/or very large online platforms. The **additional transparency requirements** that could be imposed on platforms that create the most risks include:

    - Obligations to undertake systemic risk assessments and publish the risk assessments.
    - Obligations for independent audits of platform risks and actions taken to minimise risks and to provide unredacted audit reports to the regulator.
    - Obligations to provide sufficiently detailed information to the regulator about the operation of algorithms to allow for assessment of algorithmic risk.[12] This last point takes into account the need for greater algorithmic accountability. Such communications can be subject to appropriate confidentiality obligations.

Transparency of decision-making by the regulator

- In general, eSafety is a responsive and transparent regulator that does a good job of explaining its role and its actions to the public. This includes, for example, its informative summaries of reasons for refusing to register the draft Relevant Electronic Services Code and draft Designated Internet Services Code in May 2023.[13]
- For this reason, we were surprised that an explanation for its action taken against X Corp in the Wakeley stabbing video was suppressed. Until the Federal Court file was published online and primary sources were made available, the public was confronted with confusing commentary including speculation and errors in reports that were not corrected by eSafety. For example, on the weekend of 20/21 April, multiple news reports relying on a Reuters article appeared to suggest that eSafety was seeking the removal of content without specifying that content. More significantly, it was not clear why eSafety considered this video contravenes the law, and whether the regulator was seeking that access be blocked in Australia only (including via the use of VPNs) or across all jurisdictions.
- On the first aspect, it was the Minister who was left to explain the perceived problems with the video. For example, the Minister told Radio National: '"Class 1" depicts real violence, it has a very high degree of impact, in a way that's gratuitous and likely to cause offence to a reasonable person. In this case, the very high

[12] We note that Art. 40(3) of the EU Digital Services Act specifically states that 'providers of very large online platforms or of very large online search engines shall, at the request of either the Digital Service Coordinator of establishment or of the Commission, explain the design, the logic, the functioning and the testing of their algorithmic systems, including their recommender systems.'
[13] eSafety Commissioner, 'eSafety's Decisions on Draft Industry Codes', https://www.esafety.gov.au/industry/codes

degree of impact is reached by virtue of the terrorism designation that has been given to this particular event'.[14]

- On the second issue, an explanation from eSafety about the jurisdictional aspects was needed to counter the claims made by X Corp and by Elon Musk who asked: 'Should the eSafety Commissar (an unelected official) in Australia have authority over all countries on Earth?' The failure to provide sufficient information damaged the agency on this occasion but, more importantly, it exposed a gap in public information about eSafety's decisions. The Classification Board is itself reticent about publishing its decisions, but it does provide them upon request. This contrasts to the decisions of the Classification Review Board which are published online. With the commencement of the Online Safety Act, decisions on community standards are now made by public servants and office holders in eSafety rather than by the members of the Classification Board who are drawn from the community. The need for a faster, more responsive scheme for decisions on online content explains this shift in decision-making, but it should bring with it additional – rather than reduced – transparency. **We recommend this be addressed by the creation of a public register of the decisions of eSafety**.

Q24: Should there be a mechanism in place to provide researchers and eSafety with access to data? Are there other things they should be allowed access to?

- Under Article 40(4) of the EU Digital Services Act, platforms have an obligation to provide properly vetted researchers with access to data for the purposes of assessing systemic risks and risk mitigation measures. We consider that this is a potentially significant additional transparency and accountability requirement that provides the basis for expert independent evaluation of platform operations and, accordingly, support the introduction of a similar obligation into the OSA. That said, care is needed in the formulation of any proposed new obligation. For example, it should be made clear that the access provided will be sufficient to allow vetted researchers to properly assess platform operations, including the operation of algorithms. Needless to say, it is important for the right balance to be struck between transparency and protecting commercially sensitive information, which could generally be implemented through appropriate confidentiality agreements. Consideration will also need to be given to ensuring researchers and digital platforms adequately safeguard the protections of platform users found in the *Privacy Act 1988* (Cth) (as revised). The Draft Code of Conduct for Platform-to-Researcher Data Sharing and the Model Data Sharing Agreement, prepared by a multi-stakeholder working group of the European Digital Media Observatory, could provide a useful starting point in this regard.[15]

[14] The Hon Michelle Rowland MP, 'ABC Radio National - Interview with Patricia Karvelas', 24 April 2024, https://minister.infrastructure.gov.au/rowland/interview/abc-radio-national-interview-patricia-karvelas.
[15] See European Digital Media Observatory (2022), *Report of the European Digital Media Observatory's Working Group* and *Model Data Sharing Agreement*, 31 May 2022. https://edmo.eu/edmo-news/edmo-releases-report-on-researcher-access-to-platform-data/.

Q25: To what extent do industry's current dispute resolution processes support Australians to have a safe online experience? Is an alternative dispute resolution mechanism such as an Ombuds scheme required? If so, how should the roles of the Ombuds and Commissioner interact?

- **An external dispute resolution mechanism is required and it should take the form of an ombuds scheme**. The relationship of this scheme to the eSafety Commissioner should be similar to that of the Telecommunications Industry Ombudsman (the TIO) to the communications regulation, the ACMA: **the ombuds would handle both individual and systemic complaints; it would liaise with the office of the eSafety Commissioner on any perceived need for changes to the rules the ombuds applies; and it would refer to eSafety any aspects of non-compliance with its decisions and other matters that might require the exercise of the regulator's enforcement powers.** The two entities would presumably share the role of community education. As with the TIO scheme, the digital platform ombuds scheme would be funded by industry through fees charged to members.
- As the Discussion Paper notes, the ACCC recommended the creation of a new independent external ombud scheme to help address the market power imbalance that exists between consumers and digital platforms, as well as internal dispute resolution obligations.[16] In Interim Report No 5, the ACCC in fact changed its thinking about the body that should perform the role of an ombuds. Whereas previously it had suggested that the TIO could be considered, the ACCC later concluded that 'an industry-specific ombuds would be preferable given that an existing body may not have the capability and capacity to undertake this role'.[17] In addition, although it suggested further consideration should be given to the types of disputes the ombuds should handle, the ACCC indicated the scheme would primarily be expected to resolve user complaints concerning the conduct of the digital platforms involving customers' unmet contractual expectations (eg decisions to suspend services or terminate their accounts) and/or infringement of an amended Australian Consumer Law (ACL).[18]
- Meanwhile, obligations to provide complaints facilities were introduced into the online safety codes that were registered with the eSafety Commissioner in June 2023. For example, clause 7(4)(24)(c) of the Social Media Services Online Safety Code (Class 1A and Class 1B Material) requires that complaints tools must be 'easily accessible and simple to use'.[19] A complaint can be made, for example, if a user considers that a platform has not provided certain safety features required under a code. However, users do not have a right to complain more generally about the conduct of platforms and, while eSafety can exercise its enforcement powers in a limited set of circumstances, it cannot – as eSafety itself acknowledges – resolve disputes between platforms and their users.[20]

[16] ACCC, *Digital Platform Services Inquiry: Interim Report No 5 – Regulatory Reform* (September 2022) 16.
[17] Ibid 103.
[18] For information about the ACCC's proposal to amend the ACL, see ibid 64-71.
[19] See the Register of Industry Codes: https://www.esafety.gov.au/industry/codes/register-online-industry-codes-standards
[20] See eSafety, 'Industry Codes Complaints', 20 June 2024. https://www.esafety.gov.au/industry/codes/complaints

- And in December 2023, in its response to the ACCC's Report No 5, the Government said it would 'undertake further work to develop internal and external dispute resolution requirements by calling on industry to develop voluntary internal dispute resolution standards by July 2024'.[21]
- In the same period that the ACCC was preparing Report No.5, CMT was conducting its own research on digital platform complaint handling. Our report was published in July 2022. Drawing on this research and a round table consultation exploring options for an external dispute resolution scheme for digital platforms,[22] we concluded that, while an expanded TIO is preferable, the adoption of either option would be a significant, positive step forward for consumers. However, with its narrow focus on what can be characterised as 'transactional' complaints that users make against platforms, the proposed scheme would leave consumers and citizens without an external avenue to resolve complaints against platforms that are more 'social' in nature, as well as complaints that users make against each other (rather than against the platform itself). We concluded that attention needs be directed to the former in the medium term (if not sooner) and to the latter in the medium to longer term.
- It is this category of 'social complaints' – specifically, social complaints directed by users against platforms – that is the subject of this consultation on the review of the Online Safety Act. Our view that there is a need for an ombuds that deals with these complaints remains the same; however, government should now take a holistic approach to the resolution of these problems. Just as the ACCC's findings on transactional complaints should not be considered in isolation from the problems created by social complaints, we think that any recommendations coming out of this review of the online safety regime should be considered alongside the need for action on transactional complaints. **There is a need for a single ombuds scheme to deal with transaction and social (including online safety) complaints concerning digital platforms**.
- The reasons for our findings on the need for an ombuds are explained in the report. Here, we will just reproduce the typology we developed for classifying complaints involving social media platforms (Table 1 below) and refer to the overall conclusions.

21 Treasury, *Government Response to ACCC Digital Platform Services Inquiry* (8 December 2023) 3. See also Minister for Communications and Assistant Treasurer and Minister for Financial Services, 'Government's Response to the ACCC's Major Competition and Consumer Recommendations for Digital Platforms' (Media Release, 8 December 2023).

22 See Holly Raiche, Derek Wilding, Karen Lee, and Anita Stuhmcke, *Digital Platform Complaint Handling: Options for an External Dispute Resolution Scheme* (UTS Centre for Media Transition, 2022). See also UTS Centre for Media Transition, *Submission to ACCC Discussion Paper for Interim Report No. 5: Updating competition and consumer law for digital platform services*, April 2022 and Submission to The Treasury, *Digital Platforms: Government Consultation on ACCC's Regulatory Reform Recommendations*, 22 February 2023. The round table was held at UTS on 7 December 2022.

	Social	Transactional
User-to-platform Complaints	• Illegal content including terrorism, CSAM, instruction in criminal acts • Pornography and other offensive content • Disinformation and misinformation • Content moderation disputes • Sale of prohibited goods or services • Election advertisements • Proliferation of fake accounts and other inauthentic behaviour	• Unfair digital platform business practices • Complaints about digital platform service, charges etc. • Privacy / other personal violations by digital platform • Failure to protect user eg, account hacking • Complaints about service disruption eg, account suspension • Dispute over terms of service / account suspension etc. • Complaints involving digital platform failure to comply with dispute resolution obligations
User-to-user Complaints	• Abuse, harassment and discrimination and other personal harms with an online dimension • Damage to reputation • Identity theft, impersonation • Disclosure of confidential or protected information • Other privacy breaches by third parties • Advertising content eg, community standards, offensive material • News content eg, accuracy and fairness	• Scams and fraudulent transactions • Misleading advertising and product claims, unfair terms, product defects, other sales disputes • Breach of copyright • Comments in reviews of products and services • Spam and unwelcome notifications or communications

Table 1: Types of complaints made about content and conduct on digital platforms

- As noted above, we found that the examples of complaints cited by the ACCC in DPSI Report No 5 were mostly 'transactional disputes'. They largely encompassed user-to-platform complaints, but also included some user-to-user complaints (eg, reporting and removal of scams and fake reviews). We concluded that the development of formal complaint-handling requirements, including an external scheme, for these transactional disputes would certainly be a step forward. However, without a more expansive design for the ombud scheme, there will be no external means of resolving many types of user-to-platform social complaints that arise on social media platforms. Existing regulators and industry schemes do not have jurisdiction over these types of complaints. Examples of such complaints include the failures of social media platforms to discharge their obligations in relation to: disinformation and misinformation (apart from the narrow category of complaints under the Australian Code of Practice on Disinformation and Misinformation[23] that amount to failure to implement systems and processes); news content and breaches of community standards in advertising content (where the complaint is about how the platform itself treats that content); election advertisements (except for the narrow category of actions covered by some electoral laws); content removal and moderation disclosure of confidential or protected information; and damage to reputation (apart from the narrow class of actions against platforms that might succeed, at great expense, via the law of defamation – see comments below).
- In reaching this position, we also said that consideration should be given to how internal dispute resolution standards could be used to encourage platforms to provide effective means of resolving disputes between users (eg, online dispute resolution)

[23] Digital Industry Group Inc, *Australian Code of Practice on Disinformation and Misinformation* (22 December 2022).

over matters that arise as a result of the use of the platform, apart from the schemes administered by Ad Standards and the Australian Press Council which provide a forum for the resolution of complaints about the content of advertising and news. As social disputes are likely to increase, there is a strong public policy argument for encouraging and possibly mandating social media providers to fund easily accessible and no-cost dispute resolution mechanisms. However, we believe a reasonable approach would be to consider mechanisms to address these user-to-user complaints as a second stage of regulatory reform, with attention focussed initially on user-to-platform social complaints.

- We note that since we made these findings, the parliaments of News South Wales and the ACT have passed legislation to introduce the latest round of defamation law reforms addressing the challenges presented by digital intermediaries. These reforms include a new defence for social media providers which essentially comprises an adaptation of the statutory defence of innocent dissemination.[24] To take advantage of the defence, social media services must establish a complaints mechanism for users and then take action to remove, block or otherwise prevent access to the content within seven days of a complaint being submitted. This approach should, in principle, help to address the problem of defamatory content posted online, without creating an unsustainable and undesirable increase in complainants seeking outcomes via the legal system. However, it may well lead to substantial restraints on speech if platforms remove or otherwise block access to content as a result of unsubstantiated claims of defamation. An ombuds scheme would help to address genuine user complaints while avoiding over-zealous content blocking, leaving the regulator (the eSafety Commissioner) to dedicate resources to its other important regulatory obligations.
- Finally, we note that the establishment of an ombuds scheme would help to achieve the kind of effective non-judicial grievance mechanisms anticipated under Principle 31 of the UN Guiding Principles on Business and Human Rights. This includes that such mechanisms should be legitimate, accessible, predictable, equitable, transparent, rights-compatible, a source of continuous learning and based on engagement and dialogue.[25]

Q26: Are additional safeguards needed to ensure the Act upholds fundamental human rights and supporting principles?

- The objects set out in s 3 of the OSA are narrowly focused on ‘online safety’, with ‘online safety for Australians’ defined in s 5 as ‘the capacity of Australians to use social media services and electronic services in a safe manner’. While protecting and promoting the online safety of Australians are important objectives, the implications of the regulatory regimes in the OSA extend beyond the ‘safety’ paradigm that currently underpins the Act. As expressly recognised by comparable laws in the UK and EU, regimes regulating online content have significant implications for fundamental rights, especially the rights to freedom of

[24] For example, see *Defamation Act 2005* (NSW), new s 31A Defence for publications involving digital intermediaries, which commences on 1 July 2024.
[25] UN, *Guiding Principles on Business and Human Rights*, 2011. https://www.business-humanrights.org/en/big-issues/un-guiding-principles-on-business-human-rights/.

expression and privacy. For example, regimes that may restrict access to lawful material can interfere with the right to freedom of expression, while regimes permitting unconstrained monitoring of users may infringe the right to privacy. Additionally, the regulation of online content has implications for other significant rights, such as the right to non-discrimination.[26] As observed by the UN Special Rapporteur on Freedom of Expression in its 2020 research report on content moderation, the UN Guiding Principles on Business and Human Rights (UNGPs) 'indicate how companies should respect rights — including … through policy, due diligence, implementation and remedy'.[27] While the UNGPs are not binding, signatory states, including Australia, have an obligation to promote private sector compliance with the UNGPs as part of their obligation to protect and promote rights under treaties such as the ICCPR. .

- As currently drafted, the OSA does not adequately recognise the range of rights and interests that should be taken into account in regulating online content: it does not sufficiently recognise and protect fundamental rights, which include the rights to freedom of expression and privacy, and which are not limited to the rights of users. In our response to Question 22 we have recommended that the proposed duty of care should incorporate duties to protect users' freedom of expression and privacy. Over and above this, we consider that the OSA should more clearly set out the importance of taking into account fundamental rights, which include users' rights, but extend to the rights of all Australians, in decisions relating to online content. This would bring the OSA regimes more into line with both international approaches and existing Australian content regulation, which expressly recognises a complex range of rights and community interests. For example, clause 1 of the National Classification Code (NCC), which sets out the principles that guide classification decisions, includes the important principle that 'adults should be able to read, hear, see and play what they want'.
- We therefore recommend that the protection and promotion of fundamental rights, including but not restricted to the rights of users, should be expressly incorporated into the OSA. This could be done through either a restatement of the objects set out in s 3 of the Act, or by a statement of regulatory principles, such as that in clause 1 of the NCC. At the least, we consider that the OSA should be recast to expressly acknowledge that other rights and interests which, in this context, include the rights to freedom of expression and privacy, should be considered along with safety. In relation to 'supporting principles', any balancing of fundamental rights and interests necessarily incorporates the proportionality principle.[28] In the context of the Australian legal system, which lacks a bill of rights, it may be desirable for legislation to incorporate the proportionality principle, such as by expressly providing that any interference to the fundamental rights to freedom of expression and privacy should be proportionate.

[26] See eg. International Covenant on Civil and Political Rights (ICCPR), Art 26.
[27] Special Rapporteur on the Promotion and Protection of the Right to Freedom of Opinion and Expression, *Freedom of Expression and Oversight of Online Content Moderation*, 2020, para. 8.
[28] See eg, Ofcom, Protecting People from Illegal Harms Online, Volume 4, 'How to mitigate the risk of illegal harms – the illegal content Codes of Practice (9 November 2023).

Response to Questions in Part 6: Regulating the online environment, technology and environmental changes

Q29: Should the Act address risks raised by specific technologies or remain technology neutral? How would the introduction of a statutory duty of care or Safety by Design obligations change your response?

- In our response to Question 7, we outlined the benefits and limitations of a risk-based approach to regulation. The issue question of technology specificity or technology neutrality raises different issues. The concept, or principle, of 'technology neutrality' is more complex than is sometimes thought.[29] It includes the concepts, first, that regulation should not discriminate between technologies, such as that the same laws should apply online as apply offline; and, secondly, that laws should be drafted in such a way as they are not limited to particular technologies but allow for technological change. While there are advantages with these principles, the limitations and challenges of implementing technology neutral laws and regulations should not be overlooked.
- For example, where purportedly 'technology neutral' laws are drafted at too high a level of generality, there are considerable uncertainties and ambiguities in applying the laws to actual technologies (Greenberg refers to this as 'the problem of the penumbra'). Moreover, as it is impossible to accurately predict whether, or to what extent, a future technology should be regulated (Greenberg refers to this as 'the problem of prediction'), there are limits on the extent to which laws can (or should) be future-proofed.[30]
- In our view, the challenge of designing laws to effectively apply to rapidly changing technologies is predominantly a practical issue. In the context of the OSA, laws need to be designed in such a way that they effectively address the source (or sources) of the most serious online risks and harms and, consequently, they must be able to be applied to existing technologies. Moreover, to the extent that specific technologies create specific risks or harms, laws should be designed to address those harms. In other words, laws should be as 'neutral' or 'specific' as is necessary to meet the regulatory objectives.[31]
- As explained in our response to Question 22, we support the introduction of a statutory duty of care and safety by design obligations, as these obligations effectively focus regulatory attention on the systems and processes of digital platforms. At first glance, the introduction of a statutory duty of care and safety by design obligations might, by virtue of establishing high level standards or

[29] See Essi Puhakainen & Karin Elisabeth Väyrynen, 'The Benefits and Challenges of Technology Neutral Regulation – A Scoping Review', Twenty-Fifth Pacific Asia Conference on Information Systems. Dubai, 2021, https://oulurepo.oulu.fi/bitstream/handle/10024/31036/nbnfi-fe2021081843548.pdf?sequence=1; Brad A Greenberg, 'Rethinking Technology Neutrality' (2016) 100 *Minnesota Law Review* 1495.
[30] Greenberg, ibid.
[31] This generally accords with Greenberg's concept of 'technological discrimination', which combines technology neutrality and technology specificity.

principles, appear to resolve the dilemma of drafting ‘technology neutral’ laws. In our view, however, in practice this assumption would be misguided. In relation to safety by design obligations, the practical implementation of the obligations necessarily differs depending upon the targeted technology or service. The problem of designing laws that are both sufficiently general to apply to a range of technologies but sufficiently specific to effectively apply to existing technologies is not easily solvable. That said, the introduction of general principles, in the form of a statutory duty of care and safety by design obligations into the legislation means that the issue does not need to be specifically resolved in the legislation. The details of the requisite standard of care and of the implementation of safety by design can be dealt with in subsidiary instruments, such as codes of practice; this allows for flexible responses to changes in technologies and practices.