Submission for the statutory review of the *Online Safety Act 2021* (Cth)

Dr Martin Husovec, Associate Professor

Souha al-Fihri (PS), Research Assistant

London School of Economics and Political Science

[redacted]

Introduction

We welcome this opportunity to make a submission for the statutory review of the *Online Safety Act 2021* (Cth) (**OSA**). We understand the statutory review has been brought forward to more readily consider the operation and effectiveness of the OSA. Our submission will primarily focus on Number 5(a)[1] of the Terms of Reference by considering existing mechanisms within the OSA, opportunities to enhance these existing mechanisms and the introduction of new mechanisms in the OSA, with a particular focus on the proposals relating to a statutory duty of care (**DOC**) as set out in the policy briefing published by Reset Australia in April 2024.[2]

We will do so by exploring the following areas and proposing a related recommendation:

1. **Rationale for introducing a statutory duty of care (DOC) into the OSA**
2. **Differentiation of illegal content and legal (albeit harmful) content**
3. **Ensuring the new law distinguishes between duties relating to liability for content hosting and duties relating to systemic accountability**
4. **Enhancing the existing transparency and accountability regime**
5. **Introduction of procedural rights for users and collective redress**

1. Rationale for introducing a statutory duty of care (DOC) into the OSA

The terms of reference indicate a clear appetite for amendments to the OSA which focus on regulatory oversight over risk-producing systems of online platforms as opposed to content presented through these systems. We agree that the OSA could be *enhanced* by amendments which bolster the Commissioner's oversight at a systemic level but seek to consider the implications of adopting a statutory DOC to achieve this aim.

[1] Terms of Reference, Statutory Review of the *Online Safety Act 2021*: Whether the regulatory arrangements, tools and powers available to the Commissioner should be amended and/or simplified, including through consideration of (a) the introduction of a duty of care requirement towards users (similar to the United Kingdom's Online Safety Act 2023 or the primary duty of care under Australia's work health and safety legislation) and how this may interact with existing elements of the Act.

[2] Reset Australia, 'A duty of care in Australia's Online Safety Act', policy briefing, April 2024 <https://au.reset.tech/news/briefing-a-duty-of-care/#:~:text=This%20policy%20briefing%20reflects%20discussions,into%20Australia's%20Online%20Safety%20Act.>.

A statutory set of obligations, such as a DOC, should not be seen as a silver bullet. It should not be seen as the only solution to achieving systemic regulation. It is important to consider the purpose of a DOC as a statutory mechanism to achieve the aims of the OSA. This raises the question of whether a DOC is envisaged as a formal tool to impose liability on companies for the content produced through their systems, or whether a DOC is envisaged as a means of requiring companies to meet due diligence obligations. We strongly encourage the latter conception when designing a DOC in the OSA, particularly when also accompanied by clear due diligence requirements, such as risk assessments, risk mitigation measures and reporting obligations.

Should a DOC be adopted, it should be focused on ensuring companies adopt risk-based due diligence measures as a means of meeting their duty. In contrast, a DOC should not be used as a blunt tool to establish liability for individual content hosted by such services. Given the complexity of intermediary liability in Australia (discussed at point 3 below), it is worth exploring whether a potential DOC should be accompanied by explicit liability exemptions as is the case under the EU *Digital Services Act*[3] (**DSA**). There is a risk that a DOC that is not sufficiently contained as a form of 'systems regulation' would become, especially in the absence of clear rules about liability for third-party content, a catch-all clause for any type of civil liability claims for any acts of third parties online.

The UK's *Online Safety Act 2023* (**UK OSA**) imposes a statutory DOC on companies with respect to illegal content and activity.[4] Discussions around the imposition of a DOC in the UK were similarly based on the concept of a DOC in work health and safety legislation. Although it is important to recognise the shared legal history of the UK and Australia, it is important to note that the introduction of the UK OSA was met with considerable public debate and criticism from academic and civil society organisations regarding the proposal to regulate legal but harmful content.[5] Further, these discussions occurred despite the existence of federal civil rights protections in the UK, and clear liability exemptions based on the EU law, which are both currently lacking in Australia.

With these considerations in mind, it must be seriously questioned whether a DOC in itself would be sufficient to *simplify* the end goal of enhancing the effectiveness of the OSA with a more systemic regulatory focus. In other words, would imposing a DOC in itself enhance oversight over online systems? Or would the OSA be better enhanced by the adoption of clearer statutory due diligence requirements? As shown by the DSA, these two proposals are not mutually exclusive. The OSA can adopt both flexible due diligence standards when it comes to risk mitigation, and clear rules about specific forms of diligence, such as transparency, or content moderation. The adoption of such due diligence requirements may provide common ground between the OSA and the DSA, which may assist with the new partnership between the Commissioner and the European Commission which is concerned with areas including transparency and accountability of online platforms, risk assessment and mitigation.[6]

---

[3] *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* [2022] OJ L 277/1.

[4] *Online Safety Act 2023* (UK) ss 7 and 24.

[5] Summarised in House of Commons, 'The draft Online Safety Bill and the lawful but harmful debate: Government Response to the Committee's Eight Report' (24 March 2022) <committees.parliament.uk/publications/9408/documents/161169/default/>.

[6] Delegation of the European Union to Australia, 'European Commission Services sign administrative arrangement with Australian eSafety Commissioner to support the enforcement of social media

We consider that, ultimately, a DOC may not in itself be the most effective means of enhancing systemic regulation. An alternative would be to formulate similar obligations as accountability requirements that are framed in terms of due diligence already familiar to corporate regulatory culture.

Recommendation 1: instead of a single all-encompassing duty of care, we recommend the adoption of a combination of specific and broader but targeted accountability requirements in the OSA. If a DOC is adopted as a model, the DOC must properly distinguish between liability for hosting content of third parties and accountability for the design and functioning of online systems.

2. Differentiation of illegal content and legal (albeit harmful) content

Regardless of the mechanism adopted to enhance systemic regulation of the online environment, there is a need for the mechanism to distinguish between illegal and legal content. The OSA distinguishes between Class 1 and Class 2 material in accordance with the National Classification Scheme (**NCC**) however this does not adequately ensure classification with respect to legality and has been criticised for being "outdated and overly broad", particularly in how class 2 restricts legal content.[7] We note the NCC is scheduled for review.

We agree that multiple duties of care for different kinds of content or services can be counter-productive. However, any model must maintain a fundamental distinction between illegal and legal (albeit harmful) kinds of content. The legitimacy of regulators to act is much weaker when addressing expressions that are not unlawful. Such regulators lack the backing of the parliament which would prohibit certain expressions or types of conduct. There is a greater risk of regulatory overreach for legal content that is regulated merely because it is deemed "harmful", "objectionable", or "problematic" by some. Such material is not easily defined[8] and ultimately has no legal footing in Australian law.

This is not to say that the OSA should ignore online risks posed by legal content or behaviour. However, there must be caution when considering the concept of "risks" or "harms" without regard to the legality of content. The OSA must include extra limits to permissible regulatory intervention where legal (albeit harmful) speech is concerned to avoid regulators making new content rules. Otherwise, there is a risk of regulatory overreach over speech, which should only be restricted by Parliament. The adoption of a DOC where the duty with respect to illegal and legal content is not clearly delineated risks further encroachments on speech and expression.[9]

Distinguishing between illegal and legal content can be built into accountability requirements within the OSA, regardless of whether these are standalone accountability requirements (as in the DSA) or are required as part of a DOC. This can include explicit safeguards, such as the need to always clarify by a regulator whether the risks are posed by legal or illegal content or

---

regulations' (12 June 2024) <https://www.eeas.europa.eu/delegations/australia/european-commission-services-sign-administrative-arrangement-australian-esafety-commissioner-support_en>.

[7] Digital Rights Watch, 'Explainer: The Online Safety Bill' (11 February 2021) <digitalrightswatch.org.au/2021/02/11/explainer-the-online-safety-bill/>.

[8] Daphne Keller, 'Amplification and its discontents' *Knight First Amendment Institute* (8 June 2021) <knightcolumbia.org/content/amplification-and-its-discontents>.

[9] See Martin Husovec, 'The Digital Service Act's Red Line: What the Commission Can and Cannot Do About Disinformation' (10 January 2024). Available at SSRN: <https://ssrn.com/abstract=4689926>.

behaviour before requesting certain measures, and an obligation to not exercise authority in ways that amount to content-specific interventions over risks posed by legal content.

For example, although the DSA does not specifically regulate 'lawful but harmful' content as a category, under the provisions relating to risk management obligations, Very Large Online Platforms (**VLOP**) and Very Large Online Search Engines (**VLOSE**) must assess systemic risks when undertaking risk assessments. This includes illegal content dissemination, as well as negative effects on fundamental rights and other societal concerns such as gender-based violence and the protection of minors.[10]

However, the DSA recognises that legal but harmful content should not be treated in the same way as illegal content so removal obligations only apply to illegal content.[11] In doing so, the DSA adopts a limited risk mitigation approach by preventing regulators from having power over what lawful individual behaviour should be banned and limiting regulator power to demanding content neutral solutions which preserve people's agency by empowering or re-designing user choice architecture.[12] Prioritising user agency requires a gradated scheme of action starting with platforms giving disclaimers or explanations about content moving to nudging tools then some types of visibility restrictions as a final stop.[13] This recognises that regulators do not have the same legitimacy as parliaments and thus should not have unfettered powers over lawful speech.[14]

In an Australian context, where there are no federal protections for speech and expression, albeit a very narrow implied freedom of political communication, this concern should be paramount when drafting legislation which inherently has the potential to significantly impact online speech and actions. Given this absence, Australia needs more explicit safeguards than other jurisdictions which have strong speech and other rights protections.

One way of balancing rights with respect to lawful speech is through the statutory requirement of risk assessments which require consideration of human rights, balancing certain rights including the right to freedom of speech and expression. This could balance concerns about not only systems which allow speech which breaches human rights but also systems which limit exposure to speech, such as shadowbanning.[15]

Although Australia does not have a federal human rights Act, it is a signatory to the International Covenant on Civil and Political Rights (**ICCPR**) and could require platforms to

---

[10] DSA, s 34(1).
[11] European Commission, 'Questions and Answers: Digital Services Act' (25 April 2023) <ec.europa.eu/commission/presscorner/detail/en/QANDA_20_2348>.
[12] See Martin Husovec, The Digital Service Act's Red Line: What the Commission Can and Cannot Do About Disinformation (10 January 2024). Available at SSRN: <https://ssrn.com/abstract=4689926>; Martin Husovec, 'Rising Above Liability: The Digital Services Act as a Blueprint for the Second Generation of Global Internet Rules', 38 Berkeley Technology Law Journal 3, available at <http://dx.doi.org/10.2139/ssrn.4598426>.
[13] Martin Husovec, 'How does the EU's Digital Services Act regulate content moderation' (Stanford Cyber Policy Centre, 11 April 2023) <www.youtube.com/watch?v=np05wM3h2mc&t=1750s&ab_channel=StanfordCyberPolicyCenter> 30:00-31:00.
[14] Martin Husovec, 'Rising Above Liability: The Digital Services Act as a Blueprint for the Second Generation of Global Internet Rules', 38 Berkeley Technology Law Journal 3, available at <http://dx.doi.org/10.2139/ssrn.4598426>.
[15] Kai Riemer and Sandra Peter, 'Algorithmic audiencing: why we need to rethink free speech on social media' (2021) 36(4) *Journal of Information Technology* 409, 418.

consider rights as set out in the ICCPR when conducting risk assessments. The Explanatory Memorandum to the Online Safety Bill 2021 recognises that the OSA has implications for human rights contained in the ICCPR. Further, the Commissioner is already required under s 24 of the OSA to have regard to the UN Convention on the Rights of the Child. Adopting an approach to online regulation which invokes the ICCPR could set a gold standard worldwide and encourage a path forward for other nations. The ICCPR has achieved nearly universal ratification. Invoking an international instrument of such status in the OSA may reduce concerns about the extraterritorial application of Australian law and concerns about the comity of nations where regulatory action is referable to internationally recognised human rights.

Recommendation 2: the OSA should delineate permissible regulatory action with respect to illegal and legal content, and treat the two differently. If a DOC is adopted, it should distinguish between requirements for illegal and legal content. For legal content, the OSA should require risk assessments to consider human rights as set out in the ICCPR, and potentially include explicit safeguards against mandating content-specific risk mitigation measures concerning lawful content that would amount to new content rules.

3. Ensuring the new law distinguishes between duties relating to liability for content hosting and duties relating to systemic accountability

The OSA is generally clear with respect to liability for content hosting for particular kinds of content, namely cyber-bullying of children, cyber-abuse of adults and non-consensual intimate images, and liability for failure to comply with industry codes of conduct and notice requirements relating to Basic Online Safety Expectations (**BOSEs**). The clarification of liability with respect to these kinds of content, distinct from liability for failure to meet regulatory expectations, is an important feature of the OSA and consistent with international trends in internet regulation.

In laying these foundations, the OSA demonstrated a co-regulatory focus between government and industry, made explicit in the Explanatory Memorandum to the Online Safety Bill 2021. Given the existence of these provisions in the OSA, if a DOC were to be adopted, it would be beneficial to integrate a DOC into provisions relating to compliance with these regulatory measures. For example, non-compliance with an industry code or standard could be a factor relevant to determining a breach of a DOC.

With respect to BOSEs, we agree with the view expressed by Reset Australia that the requirement for platforms to take 'reasonable steps' relating to user safety without prescribing what this might involve or the means of measuring the effectiveness of these steps, including through risk assessment, risk mitigation measures and transparency measures, limits the effectiveness of this regulatory measure. However, we consider this to be a factor in favour of the OSA prescribing these accountability requirements, regardless of whether a DOC is adopted.

The imposition of a DOC risks blurring the lines of liability for hosting third-party content and accountability for the systems which interact with content unless a clear distinction is made when establishing a statutory DOC. The imposition of a DOC requires clear liability exemptions and a sense of what the DOC concerns. Provisions relating to enforcement action and damages would need to directly relate to systemic measures and not content hosting to avoid providers being the ultimate legal backstop for all harmful content online.

This is particularly important against the backdrop of scholarly concern about the status of Australian intermediary liability and the absence of a clear theory about when online entities will be liable for their systems which enable wrongful behaviour.[16]

Recommendation 3: if a DOC is adopted, the OSA must clearly delineate that the DOC is concerned with company responsibility for their systems and distinguish this from liability for hosting third-party content. The DOC should be linked to existing provisions in the OSA relating to company responsibility.

4. Enhancing the existing transparency and accountability regime

Through the BOSE framework and provisions relating to industry codes and standards, the OSA seeks to impose transparency and accountability measures on companies. These provisions in the OSA do not prescribe the form that transparency and accountability should take. For example, ss 49 and 56 of the OSA, relating to periodic and non-periodic reporting notices, provide the Commissioner with significant discretion about the content of the notice and what information may be required by a company to comply. In contrast, the DSA has a gradated accountability regime which imposes responsibilities on companies to address particular risks by setting out specific requirements for undertaking risk assessments, risk mitigation and meeting transparency reporting obligations and other due diligence obligations. VLOPs and VLOSEs have additional accountability obligations.

There are benefits to the flexibility in the OSA, such as allowing a notice to be tailored to the relevant company or service to which it applies and allowing the content of the notice to focus on the relevant issue at hand as opposed to the crux of the notice being bogged down with additional prescriptive requirements. However, we consider that the OSA should adopt more transparency and accountability measures, such as requiring risk assessments and risk mitigation measures, whilst still retaining some flexibility for the Commissioner to determine the content of notices which may relate to these assessments and measures.

The OSA does not impose risk assessment obligations on companies despite referring to risk assessments in policy documents relating to the Act. For example, the OSA could require companies to undertake risk assessments which consider particular systemic risks and suitable mitigation measures but still grant flexibility to the Commissioner in drafting the content of a notice so that she can ask targeted questions of services but still have the ability to also request risk assessments or reports relating to mitigation measures. If the OSA required companies to undertake risk assessments, this would enhance the effectiveness of the notice regime because if risk assessments were already required to be undertaken, it would be easier for companies to comply with notices.

We consider that these transparency and accountability requirements could be implemented into the OSA without the introduction of a DOC. Risk assessments and due diligence reporting obligations are a common feature of corporate regulation and are recognisable concepts within the industry. These obligations could be tailored to the Commissioner's priority areas. For

[16] See Kylie Pappalardo and Nicolas Suzor, 'The liability of Australian online intermediaries' in Giancarlo Frosio (ed), *Oxford Handbook of Online Intermediary Liability* (Oxford University Press, 2020) 237; Peter Leonard, 'Safe Harbors in Choppy Waters: Building a Sensible Approach to Liability of Internet Intermediaries in Australia' (2010) 3(2) *Journal of International Media and Entertainment Law* 221.

example, the OSA could impose risk assessment obligations which require companies to consider online harms not explicitly covered by the OSA, such as volumetric attacks or algorithmic amplification of content,[17] similar to the requirements for VLOPs and VLOSEs in articles 35(1)-35(2) of the DSA.

These tools could also be complemented by enhancing the Commissioner's investigative and information-gathering powers, in line with other regulators. The question remains as to whether a DOC would assist in achieving systemic regulation beyond these requirements. As mentioned above, there could be normative value in doing so, provided a DOC is accompanied by accountability requirements and clearly delineated in terms of scope.

Recommendation 4: the OSA should enhance its transparency and accountability regime by requiring companies to undertake risk assessments and appropriate mitigation measures while still providing flexibility for the Commissioner to determine the content of reporting requirements relating to transparency.

5. Introduction of procedural rights for users and collective redress

To enhance liability and accountability requirements in the OSA, the avenues for redress available to users could be enhanced by the introduction of procedural rights for users and collective redress under complaint mechanisms. This recognises what can be described as an 'ecosystem of actors' in the online environment.[18]

The OSA could establish procedural rights for users to assist with the internal resolution of complaints, such as requirements relating to internal review of decisions and explanations for decisions. This would reduce the regulatory burden on the Commissioner and enhance user agency. To illustrate, the DSA imposes procedural requirements on platforms which set a minimum standard for users accessing their internal complaint mechanisms.[19] Additionally, the DSA makes provision to require platforms to prioritise notifications received by 'trusted flaggers' who act within an area of expertise. The UK OSA has provisions which allow entities to make 'super-complaints' to Ofcom about a service feature which causes significant concern.[20]

User agency would also be enhanced by providing for collective action within the OSA. This recognises that the online environment can involve collective harms and that representative bodies should be able to exercise rights on behalf of individuals.[21]

Australia has a strong tradition of collective action in enforcing consumer rights. Cross-regulatory working groups and projects already exist to facilitate regulatory cooperation and recognise the intersection of consumer rights, competition law and privacy online.[22] The OSA should provide collective redress for individuals. The DSA, for instance, includes a collective

[17] Terms of Reference Number 4.
[18] Martin Husovec, 'Rising Above Liability: The Digital Services Act as a Blueprint for the Second Generation of Global Internet Rules', 38 Berkeley Technology Law Journal 3, available at <http://dx.doi.org/10.2139/ssrn.4598426>.
[19] DSA, Chapter 2, and standardisation in Article 44(1).
[20] UK, OSA s 169.
[21] This is recognised in Recital 149 of the DSA.
[22] For example, the Digital Platform Regulators Forum.

redress for qualified consumer organisations.[23] Implementing this in the OSA could include enhancing the public-facing complaints mechanism to include complaints from individuals and consumer groups regarding systemic risks and, if adopted, perceived breaches of a DOC. These could be open to government agencies as well as NGOs. Although the Commissioner should be the focal point for administering the OSA, dispersing responsibility for actions under the OSA would be a more effective means of redress.

Recommendation 5: amendment of the OSA to include provisions for procedural rights for users and collective redress within complaint mechanisms.

Conclusion

We are mindful of the challenges which accompany the need to enhance systemic regulatory measures in the OSA. This is an opportunity to strengthen the regulation of the systems that govern the online environment and avoid regulatory resources being wasted on a never-ending game of whack-a-mole in a world of never-ending content creation. How this is done may take different forms but we consider the above considerations to be relevant to any proposed measure.

We are grateful for your consideration of our submission.

- **Recommendation 1:** instead of a single all-encompassing duty of care, we recommend the adoption of a combination of specific and broader but targeted accountability requirements in the OSA. If a DOC is adopted as a model, the DOC must properly distinguish between liability for hosting content of third parties and accountability for the design and functioning of online systems.
- **Recommendation 2:** the OSA should delineate permissible regulatory action with respect to illegal and legal content, and treat the two differently. If a DOC is adopted, it should distinguish between requirements for illegal and legal content. For legal content, the OSA should require risk assessments to consider human rights as set out in the ICCPR, and potentially include explicit safeguards against mandating content-specific risk mitigation measures concerning lawful content that would amount to new content rules.
- **Recommendation 3:** if a DOC is adopted, the OSA must clearly delineate that the DOC is concerned with company responsibility for their systems and distinguish this from liability for hosting third-party content. The DOC should be linked to existing provisions in the OSA relating to company responsibility.
- **Recommendation 4:** the OSA should enhance its transparency and accountability regime by requiring companies to undertake risk assessments and appropriate mitigation measures while still providing flexibility for the Commissioner to determine the content of reporting requirements relating to transparency.
- **Recommendation 5:** amendment of the OSA to include provisions for procedural rights for users and collective redress within complaint mechanisms.

[23] DSA, Article 90.